ELECTRONIC GOVERNMENT AND ELECTRONIC PARTICIPATION

Innovation and the Public Sector

The functioning of the public sector gives rise to considerable debate. Not only the efficiency and efficacy of the sector are at stake, but also its legitimacy. At the same time we see that in the public sector all kinds of innovations are taking place. These innovations are not only technological, which enable the redesign of all kinds of processes, like service delivery. The emphasis can also be put on more organizational and conceptual innovations. In this series we will try to understand the nature of a wide variety of innovations taking place in the public sector of the 21st century and try to evaluate their outcomes. How do they take place? What are relevant triggers? And, how are their outcomes being shaped by all kinds of actors and influences? And, do public innovations differ from innovations in the private sector? Moreover we try to assess the actual effects of these innovations, not only from an instrumental point of view, but also from a more institutional point of view. Do these innovations not only contribute to a better functioning of the public sector, but do they also challenge grown practices and vested interests? And what does this imply for the management of public sector innovations?

Volume 22

Recently published in this series

Electronic Government and Electronic Participation

Joint Proceedings of Ongoing Research, PhD Papers, Posters and Workshops
of IFIP EGOV and ePart 2015

Edited by

Efthimios Tambouris
University of Macedonia, Greece

Hans Jochen Scholl
University of Washington, USA

Marijn Janssen
Delft University of Technology, The Netherlands

Maria A. Wimmer
University of Koblenz-Landau, Germany

Konstantinos Tarabanis
University of Macedonia, Greece

Mila Gascó
Escuela Superior de Administración y Dirección de Empresas (ESADE), Spain

Bram Klievink
Delft University of Technology, The Netherlands

Ida Lindgren
Linköping University, Sweden

Michela Milano
Bologna University, Italy

Panos Panagiotopoulos
Queen Mary University of London, UK

Theresa A. Pardo
Center for Technology in Government, University at Albany, SUNY, USA

Peter Parycek
Danube University Krems, Austria

and

Øystein Sæbø
Agder University, Norway

IOS
P r e s s

Amsterdam • Berlin • Tokyo • Washington, DC

Preface

Electronic government and electronic participation are continuously transforming the public sector and society worldwide and are constantly being transformed themselves by way of emerging information and communication technologies.

The International Federation for Information Processing Working Group 8.5 Information Systems in Public Administration (IFIP WG 8.5) organizes every year two leading conferences in the field.

The international conference on electronic government (EGOV) is one of the first conferences focusing on electronic government. Since 2001, it provides a forum for scholars to present and discuss their research work.

The international conference on electronic participation (ePart) is probably the only scientific conference focusing exclusively on electronic participation. Since 2009, ePart has emerged from the strand of EGOV as a sister conference.

Through the years both conferences established themselves as leading scientific events in their fields. At the same time, they continue innovating in an attempt to increase the value each provides to its attendees. In this respect, this year, for the first time, the conferences are organised around tracks.

More specifically, five tracks were introduced, namely:

- The General E-Government Track
- The General eParticipation Track
- The Open Government & Open and Big Data Track
- The Policy Modelling and Policy Informatics Track
- The Smart Governance, Smart Government, & Smart Cities Track

The five tracks of the dual IFIP EGOV and ePart conference highlight core areas of importance for the domains of study. The overall objective of the dual conference remains to attract scholars coming from different academic disciplines to present and discuss their latest research and to shed light on advancements in the field from different, sometimes even diverse, perspectives. With the introduction of a new organisational model of the dual conference, we are also happy to have in the co-chairing a number of distinguished scholars who provide fresh insights to the conferences and who bring in new relevant communities.

Papers at IFIP EGOV and ePart aim to combine scientific credibility and rigor with high relevance to practice. Likewise, the keynote speakers come from both practice and academia, which presents a fruitful combination as practice can drive research, and research is needed by practice.

The dual IFIP EGOV and ePart 2015 "Call for Papers" for all five tracks attracted a wide range of topics with 117 submissions, which included 37 accepted completed research papers (published in separate EGOV and ePart volumes of Springer Lecture Notes in Computer Science), 31 accepted ongoing research papers, 5 accepted posters, 3 workshops, 6 PhD colloquium papers and 1 PhD presentation. The latter are being included for the first time in the IOS Press proceedings along with the ongoing research papers, workshops and posters. The papers of this volume were grouped under the following headers:

- eParticipation
- Policy Modelling
- Open Government and Smart Cities
- General e-Government
- E-Government Services

The Paper Awards Committee of IFIP EGOV and IFIP ePart was again led by committee chair Olivier Glassey of IDHEAP, Lausanne/Switzerland. The organizing Committee carefully reviewed the accepted papers and granted outstanding paper awards to the winning authors. The winners were awarded in the ceremony during the conference dinner, which has become a highlight of each year's conference. The names of the award winners can be found on the conference web page: http://www.egov-conference.org/egov-conf-history/egov-2015/.

This year, EGOV and ePart were organised in Thessaloniki, Greece under the aegis of the University of Macedonia. The University of Macedonia has long been active in research in the areas of eGovernment and eParticipation. However, the success of a conference takes much more. We would therefore like to thank the team of the University of Macedonia and particularly Eleni Panopoulou but also Maria Zotou, Elina Nanopoulou and Eleni Kamateri for their efforts in the excellent organisation of the dual conference.

The University of Macedonia is a relatively new and small university. Its departments of Applied Informatics and Business Administration have worked together for these conferences. The University of Macedonia is located in Thessaloniki, Greece; a city with 2,500 years of history and at the same time a lively, artistic city and one of the largest student centres in South-Eastern Europe. The conference dinner was held at the Byzantine museum and was preceded by an exclusive museum tour especially organised for conference participants. It could not have been more appropriate!

August/September 2015

Efthimios Tambouris
Hans Jochen Scholl
Marijn Janssen
Maria A. Wimmer
Konstantinos Tarabanis
Mila Gascó
Bram Klievink
Ida Lindgren
Michela Milano
Panos Panagiotopoulos
Theresa A. Pardo
Peter Parycek
Øystein Sæbø

Organization

Conference Chairs
Efthimios Tambouris, University of Macedonia, Greece
Hans Jochen Scholl, University of Washington, USA
Marijn Janssen, Delft University of Technology, The Netherlands
Maria A. Wimmer, University of Koblenz-Landau, Germany
Konstantinos Tarabanis, University of Macedonia, Greece

General E-Government Track Chairs
Hans Jochen Scholl, University of Washington, USA
Marijn Janssen, Delft University of Technology, The Netherlands
Maria A. Wimmer, University of Koblenz-Landau, Germany

General eParticipation Track Chairs
Efthimios Tambouris, University of Macedonia, Greece
Øystein Sæbø, Agder University, Norway
Panos Panagiotopoulos, Queen Mary University of London, UK

Open Government & Open and Big Data Track Chairs
Marijn Janssen, Delft University of Technology, The Netherlands
Bram Klievink, Delft University of Technology, The Netherlands
Ida Lindgren, Linköping University, Sweden

Policy Modelling and Policy Informatics Track Chairs
Maria A. Wimmer, University of Koblenz-Landau, Germany
Michela Milano, Bologna University, Italy
Theresa A. Pardo, Center for Technology in Government, University at Albany, SUNY, USA

Smart Governance, Government and Cities Track Chairs
Hans Jochen Scholl, University of Washington, USA
Mila Gascó, Escuela Superior de Administración y Dirección de Empresas (ESADE), Spain
Peter Parycek, Danube University Krems, Austria

Chair of Outstanding Papers Award
Olivier Glassey, IDHEAP, Switzerland

Program Committee
Suha Alawadhi, Kuwait University, Kuwait
Vincenzo Ambriola, University of Pisa, Italy
Kim Normann Andersen, CBS, Denmark
Renata Araujo, UNIRIO, Brazil
Karin Axelsson, Linköping University, Sweden
Lasse Berntzen, Vestfold University College, Norway
Laurence Brooks, Brunel University, United Kingdom
Paul Brous, Delft University of Technology, The Netherlands
Wojciech Cellary, Poznan University of Economics, Poland

Bojan Cestnik, Temida d.o.o., Jožef Stefan Institute, Slovenia
Yannis Charalabidis, National Technical University Athens, Greece
Soon Ae Chun, CUNY, USA
Wichian Chutimaskul, King Mongkut's University of Technology Thonburi, Thailand
Flavio Corradini, University of Camerino, Italy
Peter Cruickshank, Edinburgh Napier University, United Kingdom
Todd R. Davies, Stanford University, USA
Sharon Dawes, Center for Technology in Government, University at Albany/SUNY, USA
Ahmed Darwish, Ministry of State of Administrative Development, Egypt
Fiorella De Cindio, Università di Milano, Italy
Anna De Liddo, The Open University, United Kingdom
Paulo Depaoli, LUISS "Guido Carli", Italy
Yogesh Dwivedi, Swansea University, United Kingdom
Annelie Ekelin, Linneaus University/BTH, Sweden
Elsa Estevez, United Nations University, Macao
Enrico Ferro, Istituto Superiore Mario Boella, Italy
Sabrina Franceschini, Regione Emilia-Romagna, Italy
Ivan Futo, National Tax and Customs Administration, Hungary
Andras Gabor, Corvinno technology Transfer Center Nonprofit Public Ltd., Hungary
Rimantas Gatautis, Kaunas University of Technology, Lithuania
Christos Georgiadis, University of Macedonia, Department of Applied Informatics, Greece
J. Ramon Gil-Garcia, Centro de Investigación y Docencia Económicas, Mexico
Olivier Glassey, Lausanne University, Switzerland
Dimitris Gouscos, University of Athens, Greece
Johann Höcht, Danube University Krems, Austria
Helle Zinner Henriksen, Copenhagen Business School, Denmark
M Sirajul Islam, Örebro University, Sweden
Tomasz Janowski, UNU Operating Unit on Policy-Driven Electronic Governance, Portugal
Arild Jansen, University of Oslo, Norway
Luiz Joia, FGV, Brazil
Nikos Karacapilidis, University of Patras, Greece
Bram Klievink, Delft University of Technology, The Netherlands
Roman Klinger, University of Stuttgart, Germany
Ralf Klischewski, German University in Cairo, Egypt
Katarina Lindblad-Gidlund, Midsweden University, Sweden
Euripidis Loukis, University of the Aegean, Greece
Rui Pedro Lourenço, INESC Coimbra / FEUC, Portugal
Luis Luna-Reyes, University at Albany, SUNY, USA
Cristiano Maciel, Universidade Federal do Mato Grosso, Brazil
Yuri Misnikov, University of Leeds, United Kingdom
José Maria Moreno-Jiménez, Universidad de Zaragoza, Spain
Gregoris Mentzas, National Technical University of Athens, Greece
Michela Milano, DEIS Universita' di Bologna, Italy
Carl Erik Moe, Agder University, Norway

Adegboyega Ojo, United Nations University - International Institute for Software
Technology, Macao
Eleni Panopoulou, University of Macedonia, Greece
Theresa A. Pardo, Center for Technology in Government, University at Albany,
SUNY, USA
Peter Parycek, Danube-University Krems, Austria
Vassilios Peristeras, European Commission, DG DIGIT, Interoperability Solutions
for European Public Administration (ISA), Belgium
Rimantas Petrauskas, Mykolas Romeris University, Lithuania
Marco Prandini, Universita di Bologna, Italy
Michael Räckers, European Research Center for Information Systems (ERCIS),
Germany
Barbara Re, University of Camerino, Italy
Nicolau Reinhard, University of São Paulo, Brazil
Andrea Resca, Cersi-Luiss "Guido Carli" University, Rome, Italy
Øystein Sæbø, University of Agder, Norway
Rodrigo Sandoval-Almazan, Universidad Autonoma del Estado de Mexico,
Mexico
Sabrina Scherer, University of Koblenz-Landau, Germany
Margit Scholl, TH Wildau, Germany
Toramatsu Shintani, Nagoya Insutitute of Technology, Japan
Henk Sol, University of Groingen, Netherlands
Mauricio Solar, Universidad Tecnica Federico Santa Maria, Chile
Maria Sokhn, University of Applied Sciences of Switzerland, Switzerland
Maddalena Sorrentino, University of Milan, Italy
Witold Staniszkis, Rodan Systems, Poland
Yao-Hua Tan, Delft University of Technology, Netherlands
Ella Taylor-Smith, Edinburgh Napier University, United Kingdom
Peter Teufl, IAIK/Graz University of Technology, Austria
Dmitrii Trutnev, e-Government Center of ITMO University, Russian Federation
Jörn Von Lucke, Zeppelin Universität Friedrichshafen, Germany
Petra Wolf, TU Muenchen, Germany
Adam Wyner, University of Aberdeen, United Kingdom
Chien-Chih Yu, National ChengChi University, Taiwan
Anneke Zuiderwijk, Delft University of Technology, The Netherlands

This page intentionally left blank

Contents

Open Government and Smart Cities

General e-Government

e-Government Services

PhD Colloquium Papers

Posters

Workshops

eParticipation

This page intentionally left blank

The Role of Trust in E-Participation: Predictors, Consequences, and Design

Maria A. WIMMER [a,1], Sabrina SCHERER [a], Markus APPEL [b]

[a] *University of Koblenz-Landau, Faculty of Computer Science,*
Institute for Information Systems Research, Germany
[b] *University of Koblenz-Landau, Faculty of Psychology;*
Institute of Communication Psychology and Media Education, Germany

Abstract. E-participation has become subject of considerable research over the past decade. However, trust as a pre-condition and result of e-participation has not yet been extensively investigated in e-participation research. In literature, trust is perceived as a complex construct studied in distinct research disciplines. To identify and implement measures for increasing trust as well as for minimising distrust in e-participation endeavours, a trust model helps to explain the full scope of the trust construct in the context. This contribution introduces a research design that aims on the one hand to analyse predictors and consequences of trust in e-participation based on a trust model for e-participation. On the other hand, a 'trust-by-design' approach for designing and implementing e-participation projects is aimed at. The approach combines empirical research to 'understand' trust factors with design science research for 'innovating' in regards to improving the design of e-participation endeavours by the lessons and insights of the empirical research. Both strands of research also contribute to theory building of trust in e-participation. This paper aims to set the scope of the research, to introduce the research questions and to present the research design with the multidisciplinary setting.

Keywords. E-participation, Trust, Trust model, Multidisciplinary research, Research design

1. Introduction

Citizen participation is a key to a healthy democracy [2], [7]. Roberts defines citizen participation "as the process by which members of a society [...] share power with public officials in making substantive decisions and in taking actions related to the community" [24]. With the wide diffusion of the Internet, the concept of e-participation arose, which refers to citizen participation by means of innovative Information and Communication Technologies (ICT) (see e.g. [15], [16], [26]). ICT provides intriguing opportunities for citizen participation ([12], [15], [18], [22]), yet e-participation projects are not always successful [27]. Panopoulou et al investigated and identified 21 success factors in designing e-participation initiatives, ranging from strategic aspects over capacity building and organisational aspects to integration, project management, value propositions and sustainability [23]. Interestingly, trust is mentioned only in relation to security and privacy, derived from [1] and [17].

---

[1] Corresponding Author, University of Koblenz-Landau, Faculty of Computer Science, Universitaetsstr. 1, 56070 Koblenz, Germany; E-mail: wimmer@uni-koblenz.de

In literature, citizens' trust is conceived as a variable that predicts whether or not, and to what extent citizens engage in participatory initiatives more generally, and in e-participation specifically (e.g. [31]). To understand trust, several models exist, such as the integrative model of trust by Mayer et al, which conceives trust as a predictor of a risk-taking relationship between the trustor (the citizen as the trusting person or entity) and the trustee (to whom or to which trust is addressed, e.g. an individual, a group, an institution or even an ICT tool). In this model, the link between trust and the risk-taking relationship is influenced (moderated) by the perceived risk that is accompanied by the risk-taking relationship. The risk-taking relationship may yield more or less desired outcomes, which in turn influence the future trustworthiness perceived by the trustor towards a trustee [19]. Another trust model is proposed by McKnight & Chervany. Their interdisciplinary model of trust construct combines different perspectives and interrelations between these perspectives, which classify five trust types (disposition to trust, institution-based trust, trusting beliefs, trusting intentions, and trust related behaviour) [20]. In his conceptualisation, Blind identifies five types of trust in e-government, which are also relevant in e-participation contexts: political trust, social trust, technological trust, moral trust and economic trust [4]. An extensive literature review on trust and e-participation is already carried out by the authors, and the findings are summarised in [29]. The literature review shows that the concept of trust is not extensively researched in the context of tools and processes to support citizen participation. Empirical research is needed to better understand the predictors and consequences of trust in e-participation contexts. Furthermore, research needs to go beyond sole investigation of understanding influence factors. Multidisciplinary research is needed to combine the analytical and empirical investigations with design sciences, i.e. the insights and lessons of empirical research should be incorporated in the design of future e-participation solutions to increase trust and citizen participation.

In this paper, we present a multidisciplinary research design to combine investigations on 'understanding' phenomena of trust in e-participation with research on 'innovating' solutions based on the insights provided by the empirical research. Both strands of research are complemented with research to embark on and to contribute to theory building in the field.

The paper is structured as follows: Section 2 introduces the underlying trust model informing the research. The research design with research questions and the methods to be applied are described in section 3. Section 4 concludes with a discussion of some challenges of such kind of research as well as an outlook on current and further work.

2. Trust model for e-participation

As already outlined in the introduction, different models exist to explain the concept of trust. In an earlier publication, we developed a trust model for e-participation [28], which embarks on the *Integrative Model for Trust in Organisational Settings* by Mayer et al [19] and extends this model with facets of the *Interdisciplinary Model of Trust Constructs* by McKnight and Chervany [20]. It also adds elements to the integrative model of Mayer et al that are based on findings of own experiences in e-participation research and on literature review [29], including insights from information systems research (e.g. [13]), e-participation research (e.g. [17]), e-government research (e.g. [4]) and information technology research (e.g. [32]).
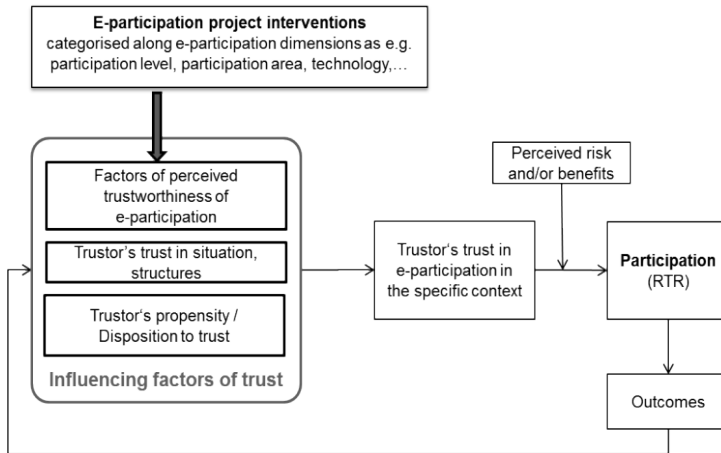
Figure 1. Trust model for e-participation (abstract view) as foundation for the research work

The trust model for e-participation is the foundation for the research work introduced in this paper. It is presented in Figure 1 in an abstract view. Subsequently, the main elements of the model are outlined:

Influencing factors of trust suggested in literature such as:

- *Trustor's propensity / disposition to trust* refers to the willingness of the trustor to trust others, as „some parties are more likely to trust than are others" ([19] pp. 714-715). McKnight and Chervany ([20], p. 38f) suggest to incorporate 'faith in humanity' and 'trusting stance'.
- *Trustor's trust in situation and structures* covers 'institution-based trust' suggested by McKnight & Chervany [20], which includes trustors' trust in protective structures "that are conducive to situational success" such as legal and technological measures to protect from privacy loss ([20], p. 37).
- *Factors of perceived trustworthiness of e-participation* covers trustworthiness of other actors, of ICT tools and other relevant concepts in e-participation contexts (e.g. processes). Mayer et al. propose ability, benevolence, and integrity as variables to measure trustworthiness of actors ([19], p. 717ff). Lankton & McKnight propose functionality, reliability, and helpfulness as the factors to assess trustworthiness of tools [13]. Own e-participation research indicates that trusting the process of participation might be another facet of trust to measure. Söllner et al suggest performance and purpose of e-participation projects as factors [32].

Trustor's trust in e-participation in the specific context is a function of all above influencing factors of trust.

Perceived risks and/or benefits involves *"trustor's belief about likelihoods of gains or losses outside of considerations that involve a relationship with the particular trustee"* ([19], p. 726). It also includes trustor's belief about advantages and positive outcomes.

Participation refers to the action taken by the trustor as a result of trust, i.e. whether or not a trustor will engage in a RTR (Risk-taking-relationship, [19], p. 726).

Outcomes correspond to the concept in Mayer et al.'s model ([19], p. 728) that reflect positive or negative impact on influencing factors of trust in a next iteration of trust and RTR; in which form this impact can happen is still subject of research.

E-participation project interventions refer to characteristics that might influence factors of perceived trustworthiness of e-participation due to different design and implementation choices, settings of an e-participation project (cf. "*Web Vendor Interventions"* cited in [20], p. 44), or key e-participation dimensions ([15], p. 6).

The trust model for e-participation as presented here provides the foundation for our research, whose research design is introduced next.

3. Research Design

3.1. Research objectives and research questions

The overall scope and research design is depicted in Figure 2, showing in the grey box an abstract view of the trust model for e-participation introduced in the previous section.
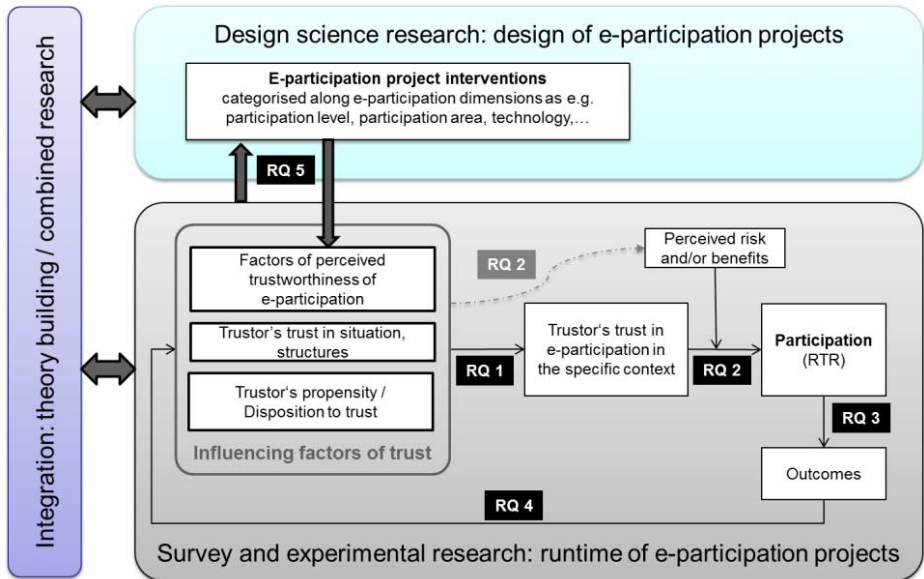
Figure 2. Research design with scope, types of research and research questions reflected along an abstract trust model for e-participation

The research involves four specific objectives and related research questions; the first three objectives support "understanding" of phenomena through empirical studies and incorporate four research questions (RQ 1 to RQ 4 within the grey box labelled "Survey and experimental research"). The fourth objective contributes to "innovating" through design guidelines for e-participation solutions and is reflected with RQ 5 (in the turquoise box labelled "Design science research"). The objectives are informed by - and likewise feed back findings to - theory (the bluish box) as is indicated with the arrows in the figure. The interplay between "understanding" and "innovating" is displayed with arrows, too. Next, the four specific objectives and research questions are introduced.

The first objective is to identify predictors of trust and risk-taking relationships in the e-participation context. These predictors reflect variables of the potential users (the

trustors) such as general propensity to trust, attitudes towards e-participation, self-efficacy, digital literacy, media use, etc. (cf. [6]; [33]). They also reflect variables of the e-participation actors, processes and ICT tools that may affect perceived trustworthiness and - in turn - trust and actual behaviour of a trustor in e-participation offers. The research question (indicated with RQ 1 in the figure) is: *Which factors influence trust of potential participants (trustors) in e-participation in a specific context?*

The second objective is to examine the interplay between trust and participation given variations in citizens' perceptions of risks and benefits of engaging in e-participation. The research question (RQ 2) is: *Which risks and benefits influencing the decision to participate are perceived by trustors, based on the trust in e-participation in the specific context?* The dotted grey arrow from the 'influencing factors of trust' box to the 'perceived risks and/or benefits' box is also labelled with RQ 2 as the influencing factors may directly moderate perceived risks and/or benefits outside the trust function.

The third objective is to examine the consequences of taking (or not taking) part in an e-participation. We aim to get an understanding and shed light on how patterns of a participation experience (the RTR) relate to immediate outcomes (e.g., evaluations of the experience) and more long-term changes in trustor's variables (like attitudes towards e-participation, attitudes towards the trustee, self-efficacy). Two research questions embody this objective: RQ 3 - *What potential outcomes may the trust relation of a participation (RTR) deliver?* RQ 4 - *How does the (non-)participation of trustors in e-participation endeavours affect the influencing factors of trust? (trustor's propensity, trust in situation and structures, and perceived trustworthiness of e-participation (tools, actors, process, etc.)).*

Based on a better understanding of the antecedents and consequences of e-participation, the fourth objective is to develop guidelines for designing and implementing e-participation tools and processes that receive higher trust (i.e. contributing to realise trust-by-design). The research question (RQ 5) is: *Based on insights of RQs 1-4 determining causal relationships within the e-participation initiatives, how should e-participation initiatives be designed to positively influence the factors of perceived trustworthiness of e-participation?*

3.2. Methods employed in the research

To investigate the five research questions, a three-fold methodological approach is chosen, based on the three types of research as indicated in Figure 2:

1) *Theory building* by studying and reviewing existing theory and empirical literature on trust, trust and e-participation, and related contexts from different disciplinary perspectives. The study of theory contributes to extend the body of theoretical knowledge through findings of the review and of the other two types of research (the dependencies are in both directions and are realised over time). Based on such review, the underlying trust model for e-participation (cf. section 2) will be validated and developed further. The literature review will unveil a number of variables already studied conceptually or empirically. Relationships between variables that have attracted substantial attention by empirical research will be evaluated by means of a quantitative meta-analysis. For the meta-analysis, empirical literature on e-participation (e.g. [11]) as well as literature on other risk-taking online behaviours (e.g., in the field of e-commerce, [3]) and on participatory behaviour in the political realm will be inspected (without violating a need for homogeneity among studies, [14]). Theory

building will support the investigation of the five research questions by providing underlying theoretical models and by identifying variables influencing trust that have been studied in literature before (empirically, conceptually and theoretically).

2) *Empirical research (longitudinal and experimental studies)* to gather insights and evidence to better understand which factors influence trust in e-participation endeavours, and under what conditions. The integrative model for trust [19] as well as the trust model for e-participation (cf. section 2) include relationships that can be examined empirically. A longitudinal study using surveys allows to investigate the perception of trust along citizen participation in an online participatory budgeting project. Participatory budgeting involves citizens in the planning of the annual budget of a municipality or city. Citizens who engage in such an e-participation project will be followed at different points in time. Psychometrically pre-tested scales to assess the key variables will be used. The data will be analysed using structural equation modelling (SEM, [5]; [9]). Citizens who do not participate in the participatory budgeting project will be surveyed through paper-based questionnaires and/or interviews. These data will supplement the findings of the online surveys. To strengthen the assumption of causal relationships, selected relationships will be examined in a laboratory environment, where experimental methods will be used. The research in this strand will help to answer research questions 1 to 4, which focus on "understanding" phenomena that influence trust in e-participation contexts.

3) *Design science research* to develop guidelines and a trust-by-design development framework / engineering approach. Insights derived from the previous research activities will be used to inform guidelines to design and implement trustworthy e-participation processes and tools. These processes and tools are supposed to help overcoming current obstacles in e-participation success by increasing trust in tools, processes and actors. The guidelines will enrich the reference framework for e-participation projects [30] and will support the development of a "trust by design" approach for e-participation similar to Keen [10] who put forward such a concept for e-commerce. With this, RQ 5 will be answered. *Design science research* [8] guides this research. In line with Hevner et al ([8], p. 77), the design and evaluation of guidelines as well as the extension of the reference framework and the development of the "trust-by-design" approach will rely on findings generated by activities in 1) and 2) above and on experiences collected in previous research activities. Likewise, activities in this strand of research will provide new grounds for empirical research to understand influence factors changing after application of a new "trust-by-design" approach of design science research, and it will also contribute to theory-building as aimed at in 1).

The ultimate principle in this research design is that teams from social sciences, communication studies and psychology collaborate with information systems and e-government researchers to ensure an effective complementarity and transition of research findings from one strand of research to the other. In the figure, these interrelations across disciplines and demands for collaboration are indicated with the dark-grey shaded arrows.

4. Discussion and Outlook

As outlined before, the nature of the research questions raised in the research design requires cooperation across different research disciplines. First results of a fertile inter-disciplinary cooperation have already been demonstrated in the drafting of the trust

model for e-participation. In the current project, researchers from psychology and communication studies collaborate with information systems and e-government researchers. The research presented in this paper will continue this successful collaboration in a more structured, systematic and three-fold methodological approach as presented before.

Different challenges have to be borne in mind and resolved in such an inter-disciplinary research cooperation as e.g. different understandings of terms, different schools and methodical approaches, different publication outlets, or different value propositions and perceptions of what rigorous research is. A major challenge has been so far the difference in the research methods applied. Exploratory and descriptive qualitative research is identified as valuable research to construct "definitions and productions of typologies, driving subsequent quantitative research" ([25], p. 552). Difficulties of evaluating e-participation initiatives are e.g. that only few quantitative results are available that can be compared with each other ([26], p. 420). Traditionally, psychology bases on quantitative research approaches ([21], p. xvii) – even if mixed-method approaches are also applied in psychology nowadays (see e.g. [21]) – while in design science research and in research closer to technical disciplines, engineering approaches and conceptual modelling are crucial and well accepted methods, besides empirical methods (both, qualitative and quantitative). The tree-fold approach suggested in this paper tries to overcome this challenge by integrating different research methods. It also better ties theory and empirical / conceptual research together.

Current work of this research investigates literature of different disciplines on trust theories and empirical studies as described in the first methodical outlet of the research design section. Findings will inform the empirical research and design sciences research. Besides the literature analysis, we will elaborate a questionnaire for the longitudinal study of trust in participatory budgeting. The research work in the multidisciplinary team is planned to continue for several years. Besides the objectives of the research as presented in this paper, insights will be gained on factors impacting the success of multidisciplinary collaborations as well as on the triangular composition of the research design, involving theory building as well as empirical research for gathering understanding of trust in e-participation and design science research for innovating e-participation designs towards more trustworthy solutions.

References

[1] Aichholzer, G., & Westholm, H. (2009). Evaluating eParticipation projects: Practical examples and outline of an evaluation framework. *European Journal of ePractice*, 7, 27–44

[2] Barber, B. (1984). *Strong Democracy: Participatory Politics for a New Age* (Vol. 1986). Berkeley: University of California Press.

[3] Beldad, A., De Jong, M., & Steehouder, M. (2010). How shall I trust the faceless and the intangible? A literature review on the antecedents of online trust. *Computers in Human Behavior*, *26*, 857-869.

[4] Blind, P. K. (2006). Building trust in government in the twenty-first century: review of literature and emerging issues. In 7th Global Forum on Reinventing Government (pp. 26–29).

[5] Bollen, K.A. (1989). Structural Equations with Latent Variable. John Wiley & Sons, Inc., North Carolina

[6] Campbell, S. W., & Kwak, N. (2010). Mobile communication and civic life: Linking patterns of use to civic and political engagement. *Journal of Communication*, *60*, 536-555.

[7] Habermas, J. (1992). *Faktizität und Geltung: Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats*: Surkamp, Frankfurt am Main.

[8] Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS quarterly, 28*(1), 75–105.

[9]   Hooper, J., Coughlan, M.R., & Mullen. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. *Electronic Journal of Business Research Methods*, 6 (1), 53-60

[10] Keen, P. G. W. (2000). *Electronic commerce relationships: Trust by design*. Upper Saddle River, N.J.: Prentice Hall.

[11] Kim, S., & Lee, J. (2012). E-participation, transparency, and trust in local government. *Public Administration Review, 72*(6), 819–828. doi:10.111/j.1540-6210.2012.02593.x

[12] Kubicek, H. (2014). Staatliche Beteiligungsangebote im Internet – Ein Überblick. In K. Voss (Ed.), *Bürgergesellschaft und Demokratie. Internet und Partizipation* (pp. 263–298). Springer Fachmedien Wiesbaden. Retrieved from http://dx.doi.org/10.1007/978-3-658-01028-7_15

[13] Lankton, N. K., & McKnight, D. H. (2011). What does it mean to trust facebook? Examining technology and interpersonal trust beliefs. The DATA BASE for Advances in Information Systems, 42(2), 32–54. doi:10.1145/1989098.1989101

[14] Lipsey, M. W. & Wilson, D. B. (2001). *Practical meta-analysis.* Beverly Hills, CA: Sage.

[15] Macintosh, A. (2004a). Characterizing e-participation in policy-making. In *Proceedings of the 37th annual hawaii international conference on system sciences (HICSS)*. Los Alamitos, CA, USA: IEEE Computer Society.

[16] Macintosh, A. (2004b). Using information and communication technologies to enhance citizen engagement in the policy process. In (OECD 2004).

[17] Macintosh, A., & Whyte, A. (2008). Towards an evaluation framework for eParticipation. Transforming Government: People, Process and Policy, 2(1), 16–30. doi:10.1108/17506160810862928

[18] Maier, J. (2007). Web 2.0 - Moderatorenrechte für alle? Gibt es eine E-Partizipation 2.0 im Web 2.0? In *Beiträge zur Demokratieentwicklung. E-Partizipation: Beteiligungsprojekte im Internet* (pp. 282–296). Bonn: Stiftung Mitarbeit.

[19] Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organization trust. *Academy of Management Review, 20*, 709–734

[20] McKnight, D. H., & Chervany, N. L. (2001). Trust and distrust definitions: one bite at a time. In R. Falcone, M. Singh, & Y.-h. Tan (Eds.): *LNAI, Trust in Cyber-societies* (pp. 27–54). Springer-Verlag Berlin Heidelberg.

[21] Mertens, D. M. (2014). Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods. Sage Publications.

[22] OECD. (2004). *Promises and problems of e-democracy, challenges of online citizen engagement*. Retrieved from http://www.oecd.org/dataoecd/9/11/35176328.pdf

[23] Panopoulou, E., Tambouris, E., & Tarabanis, K. (2014). Success factors in designing eParticipation initiatives. *Information and Organization*, 24 (4), 195-213

[24] Roberts, N. (2004). Public Deliberation in an Age of Direct Citizen Participation. *The American Review of Public Administration,* 34 (4), 315-353. doi: 10.1177/0275074004269288

[25] Rowe, G., & Frewer, L. J. (2004). Evaluating public-participation exercises: a research agenda. Science Technology Human Values, 29(4), 512–556. doi:10.1177/0162243903259197

[26] Sæbø, Ø., Rose, J., & Flak, L. S. (2008). The shape of eParticipation: characterizing an emerging research area. *Government Information Quarterly, 25*(3), 400–428. doi:10.1016/j.giq.2007.04.007

[27] Sæbø, Ø., Flak, L. S., Sein, M. K. (2011). Understanding the dynamics in e-Participation Initiatives: Looking through the Genre and Stakeholder lenses. *Government Information Quarterly, 28*(3), 416–425.

[28] Scherer, S.; Wimmer, M. A. (2014): Conceptualising Trust in E-Participation Contexts. In: Electronic Participation. 6th IFIP WG 8.5 International Conference, ePart 2014. Springer Berlin / Heidelberg. Nr. 8654. S. 64-77.

[29] Scherer, S.; Wimmer, M. A. (2014): Trust in e-participation: Literature review and emerging research needs. In: ICEGOV2014 8th International Conference on Theory and Practice of Electronic Governance. ACM. S. 61-70.

[30] Scherer, S.; Wimmer, M. A. (2011): Reference Framework for E-participation Projects. In: Tambouris, E.; Macintosh, A.; de Bruijn, H.: Proceedings of Third International Conference on eParticipation (ePart 2011). Springer Berlin / Heidelberg. Nr. 6847. S. 145-156.

[31] Shah, D., Schmierbach, M., Hawkins, J., Espino, R., & Donavan, J. (2002). Nonrecursive models of Internet use and community engagement: Questioning whether time spent online erodes social capital. *Journalism & Mass Communication Quarterly*, *79*, 964-987.

[32] Söllner, M., Hoffmann, A., Hoffmann, H., Wacker, A., & Leimeister, J. M. (2012). Understanding the Formation of Trust in IT Artifacts. In Proceedings of the International Conference on Information Systems (ICIS), Orlando Florida, USA.

[33] Xenos, M., & Moy, P. (2007). Direct and differential effects of the Internet on political and civic engagement. *Journal of Communication*, *57*, 704-718.

The Communities of Practice model for understanding digital engagement by hyperlocal elected representatives

Peter CRUICKSHANK[a,1] and Bruce RYAN[b]
a, b Edinburgh Napier University, Scotland, UK

Abstract. There has been much research into citizens' engagement with their representatives. This paper offers an approach to understanding sustained take-up of internet technologies by these representatives in a (hyperlocal) democratic context using Community Councils in Scotland a case study. A Community of Practice model was developed and initial data collected to evaluate whether the model can be adapted for contexts where community boundaries are not clear. The focus is the community of users of technology: representatives as primary content creators as a necessary first stage before higher levels of engagement and participation are possible. The CoP model is found to have potential, even in a context of weak, dispersed and non-self-aware communities. The importance of understanding transitions and level of engagement is highlighted and another avenue for further research identified.

Keywords. Hyperlocal government; Digital engagement; Communities of Practice; e-participation; knowledge management; Scotland

1. Introduction

1.1. Motivation

Research has generally focused on citizens' online engagement with government and their elected representatives, e.g. [1]–[4], and has focused on success, though the occasional failure is acknowledged [5]. But there has to be something to engage with. This article is about the other end of engagement – the (hyperlocal) governments that citizens engage with at community level. At larger government levels there may be resources and paid professionals; at the hyperlocal or community level there is a dependency on the representatives themselves to create and curate content, and to carry out the engagement as part of their role in the participative process. There is some evidence that at this level of democracy, engagement is erratic, inconsistent and often short-lived [6]. Many local groups rely on small groups of volunteers, leading to dependency and vulnerability. Yet some groups succeed nevertheless.

The motivation for this paper is to explore a framework for understanding representatives' use of technology. It evaluates the extent to which knowledge management approaches can contribute to an understanding of why some actors at the

---

[1] Corresponding Author: Peter Cruickshank, Edinburgh Napier University, 10 Colinton Road, Edinburgh EH10 5DT, Scotland, UK. E-mail: p.cruickshank@napier.ac.uk

lowest level of representation succeed in creating sustained engagement with citizens while others fail, by applying a Communities of Practice (CoP) model. It reviews some of the relevant literature and contrasts different models of hyperlocal government to contextualize the Scottish experience that was the basis for this study.

1.2. Hyperlocal government in Scotland: Community Councils

The UK is currently experiencing a cycle of constitutional change, with forms of government under scrutiny: in particular the balance of centralized/local control (the so-called localism agenda) and even after the independence referendum of 2014, the extent to which the power should or could be devolved to its constituent nations, including Scotland. (Much has been written on this subject; [7] provides one overview).

In the Scottish context, the tiers of government are the UK Government, the Scottish Government, 32 Local Authorities and potentially 1370 Community Councils. Community Councils (CCs) are the smallest, most local units of democracy in Scotland. They consist of unpaid, nominally elected citizens who live in the communities they represent. Their statutory duty is to ascertain, co-ordinate and express community opinions. CCs also have the right to be consulted on licensing and spatial planning.

The first three tiers, along with government-associated public bodies, are responsible for almost all service provision in Scotland: we therefore use 'hyperlocal government' [8] to refer to the smallest elected units of government because it fits better than 'municipality' which implies responsibility for service delivery. Another issue is the low level of interest in CCs as democratic structures. Currently, 16% of potential CCs do not exist [9], [10] while there is a paucity of candidates for those that do exist: two thirds of elections are uncontested, meaning that CC membership is essentially self-selecting. This has contributed to a history of challenges starting with establishing their legitimacy in reflecting public opinion in their own areas [8].

Despite increasing use of online communications by other tiers of government in the UK, recent research has shown that very few CCs effectively use online techniques: just 27% have up to date online presences. Further, the proportion using social media is very small (less than 10%) [9], in great contrast to the Austrian and Norwegian examples discussed below. Worse, there was significant churn between on- and off-line status between 2012 and 2014: 1129 CCs existed in both years but 34% changed status in this period [9]. For instance 68 CCs moved from 'exists, not online' to 'online, out of date': they must have gone online since summer 2012 but let their presences lapse in under two years. Overall, 45% of viable presences have failed: at the least, this is likely to lead to significant self-efficacy issues to overcome if these CCs are to return to the internet.

Previous research into the drivers and inhibitors of individual CC internet use suggested that the major barriers were cost (particularly time-costs), factors related to the digital divide and lack of support for CC 'digital engagers' by their peers [9]. On the other hand, the support of community volunteers who carry out hyperlocal news reporting can provide can be vital.

1.3. Government digital engagement: the European contrast

The general pattern across Europe is for a bottom tier of elected government to represent small areas: villages and environs, individual neighborhoods and suburbs and similar. In contrast to Scotland, they can provide services, enabled through local taxes and charges supplementing government grants. For example, Austria's *Gemeinden* provide services

such as water, sewerage and recreation facilities and have consistently punched above their weight [11]. Similarly, Norwegian *kommuner* provide services from education to transport, and work on agricultural and environmental issues [12].

Table 1. Summary of population and municipal data for the three countries discussed

Country	Population	Hyperlocal governments		Population / unit
Austria	8,375,000	*Gemeinden*	2,354	3,558
Norway	4,986,000	*kommuner*	432+	11,542
Scotland	5,538,000	community councils	1,369	3,892

These local government units use the internet to communicate with their citizens. In 2009, 98% of *Gemeinden* had websites. As early as 2003, 90% of *kommuner* had websites[13] and by 2011, 58% had social media presences[14].

To summarize, Scottish Community Councils are an edge case – officially representative bodies of comparable size to equivalents across Europe (see Table 1) but with negligible budgets and powers. This raises interesting and relevant questions around how the low level of online engagement[2] arose: is it purely down to the lack of powers? How much can be attributed to socio-technical challenges?

2. Literature review

2.1. E-participation and digital engagement

There is a large body of published research on the effectiveness of e-engagement, often showing that its importance has been overstated [1]; even so it is clear that digital communication is growing in importance, at least as one channel of many for reaching and interacting with citizens [15]. Much e-participation research has focused on citizen learning or engagement [16], rather than learning by the content creators. An online presence is more than simply about marketing – it is potentially about two-way communication, but the reality is that levels of citizen engagement are low [10] and having a simple but current online presence is a start. More positively, at the community level, there is some evidence that online conversations can support deliberative democracy in the medium to long terms, based around single communities or a dispersed network of sites [17], but that is only possible once a critical mass of participants is present. There has been some work in the past on the technology needs of community councils [18] but in the context of the development of specialist digital tools to support engagement.

2.2. Communities of Practice

Communities of Practice (CoPs) are characterized by Wenger [19] by their collective learning in a shared domain. They have three defining characteristics: a *domain* of interest (collective knowledge and competence is valued by members); a *community* (relationships involving joint activities and knowledge-sharing, even though members may work alone); a *practice* (including shared repertoires of experiences, tools and methods). A CoP is a social construction and social learning system which drives mutual

---

[2] Defined here as conversations and human interaction via the internet – posting a document to a hyperlocal government or community website would not be included but disseminating links about it would, as would online conversations about its content.

learning and knowledge exchange, and as such CoP studies are most significantly found in Knowledge Management literature [20]. As the idea has evolved, it has broadened to include learning mechanisms and using a social dimension [21].

A CoP is also characterized by a clear boundary and by *boundary objects,* which are defined as the entities that can link communities together as they allow different groups to collaborate on a common task [19]. More simply, boundary objects can communicate to large numbers of people, potentially in different ways. As such, they have different meanings to CoP members and the general public [22]. Knowledge-transfer occurs within CoPs [23] but explicit knowledge may cumulate in an ad-hoc fashion [24], so that significant knowledge remains tacit.

Motivation to join CoPs may come from expectation of access to knowledge and rewards of various kinds [24]. The presence of a 'cognitive pressure' (i.e. knowledge needs experienced within an organization) is a necessary condition for the emergence and survival of CoPs [20]. In a business context, CoPs often need to be intentional: that is they can require management to plan make the learning points explicit [20] and have to be deliberately designed, managed and cultivated [25]. Legitimate peripheral participation is also an important concept as it recognizes that the boundary of a community may be present, but unclear – and that peripheral participation may be denied by existing members who feel disadvantaged by new entrants[26].

2.3. Conclusion

The relevance of CoPs to e-participation has been noted before [27] but the concept has not often been used to evaluate actual behaviors or elected representatives: it could be used as an analytic tool to understand the role of knowledge sharing networks in creating sustainable online presences. The expectation is that self-efficacy and sustainability of online presences would be higher and churn in online presences lower for CC members who are embedded in a CoP.

One possible challenge to the formation of a CoP here is that the barriers to exit are so low: in other cases there is an assumption that there is a cost of leaving the community (e.g. loss of salary or access to information) – this may not be the case for a CC member. If CoPs are characterized by conflict while learning (or learning to become a CoP member is inherently problematized) [26], the temptation is to leave rather than work through the conflict (and learn). Related to this is the question of whether non-CC members can be considered members of a CoP here at all.

3. Pilot study and research subjects

3.1. Background and use of digital engagement

A pilot project ran in 2014 with the aim of validating a methodology for identifying, creating and supporting a CoP for digital engagement by Scottish CCs. Objectives included characterizing the digital presence and potential boundary objects acting as markers of a CoP and identifying the impact of interventions.

The project investigated the potential existence of a CoP around three neighboring CCs in a distinct area of a Scottish city. These CCs are connected by a similar social context (notably highly multicultural communities, drug abuse and planning blight). CC1

has long had an active blog and a Twitter account; its blog had been auto-tweeting links to new posts for about 18 months before this project began. However, 'organic' tweets were rare until it ran a debate on the Scottish Independence Referendum in summer 2014. Thereafter, CC1 started to use its Twitter account actively to communicate with citizens. CC2 also had a long-standing blog. CC3's blog, modelled on CC1's, was created about a year before this research began.

3.2. Data gathering

A complete social network analysis (SNA) was judged infeasible, most significantly because at this stage a boundary could not be identified: for instance it was known that CC activity is often supported by non-members. Instead, the three neighboring CCs were studied using mixed methods focused on qualitative data, using interviews with members who undertook their CCs' digital communication to gather data on how knowledge around digital engagement was acquired, shared and managed. Several types of knowledge were considered, including (i) the technical skills of creating online content (ii) skills around digital communication and engagement, namely how to write and work out what to say (iii) knowledge around how CCs work and what they do, (iv) knowledge of local news and developments. The focus for research was technical and digital knowledge.

A sociogram of the knowledge sharing was used to visualize and understand the knowledge sharing ties between the identified actors. This data was augmented with personal knowledge and reviews of online presences. Information about citizens who communicate digitally with CCs was not gathered unless citizens were explicitly part of knowledge-sharing links. These methods yielded rich data on the relationships between actors dispersed between units of hyperlocal government.

4. Results

4.1. Impact on participants

As would be expected in a project such as this, the research made an impact on the participants. The research in CC1 had been structured around addressing the perceived needs of participants, and their skills and availability. However, many interviews turned into ad-hoc one-to-one coaching sessions, where explicit and tacit knowledge was shared. This may best be summed up by a quote from an interviewee: *'serendipity and discovery happen when not working in isolation'.*

The rewards for participation were most often personal satisfaction at fulfilling 'democratic duties' and helping neighbors, although satisfaction also arose from successful use of new software and building personal relationships. New and existing CC digital engagers increased self-efficacy thanks to emotional and technical support from other community members. For newer members, there was increased sense of their roles as representative of their CCs, learning more about what is 'out there', conveying this back to CCs, and understanding that others are interested in CC's work. Other interviewees confirmed these benefits, acknowledging that their role includes emotional support as well as teaching know-how.

4.2. Identifying communities

The data gathered showed that there are citizens within each CC who are interested in and do CC digital engagement, and who communicate with each other to share relevant knowledge. There are a number of citizens who contribute knowledge and have interests in CC digital engagement but do not currently tweet or post on behalf of CCs, that is they facilitate but do not directly create boundary objects. Recalling Wenger's defining characteristics, they could be classed as members of a Community of Learning around the putative CoP. Others, including representatives of bodies who have interests in CC digital engagement, could be members of a Community of Interest. Figure 1 below illustrates how the main communities identified may interrelate.

Creation of boundary objects such as tweets might be used to distinguish those who practice and hence 'truly' are in the CoP from those who do not practice. Of course there may well be people who have the skills to produce tweets and posts but currently do not do so; although they are CC members and help or advise on tweet/post creation it still remains to be established whether they should be counted as being part of the CoP.
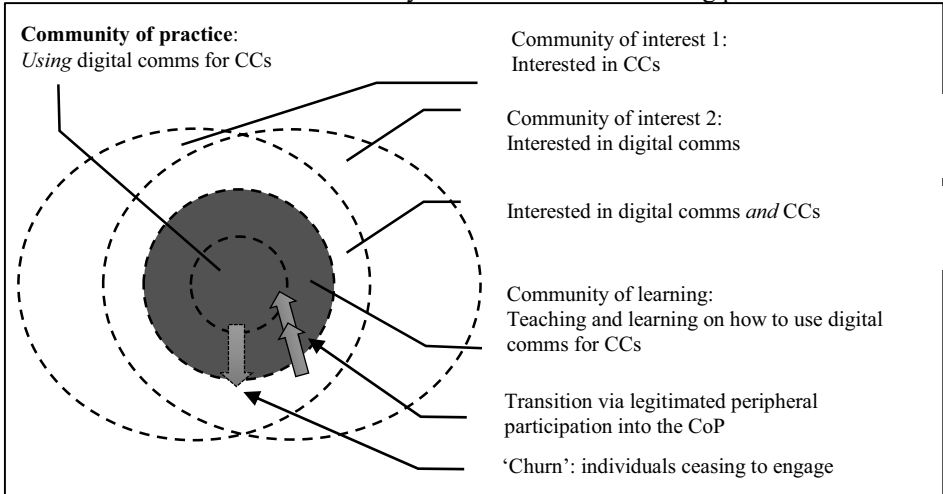
Figure 1: Sharing skills around CC digital engagement

4.3. Revisiting the method

This analytical approach has the potential to enable practitioners to move from analysis to interventions supporting vulnerable peers. A challenge is to keep the work focused on creators of digital content. It is also important to be clear about what level of knowledge is being exchanged (technical, communications skills, domain procedural knowledge or news on local developments).

As is often the case with pilot projects one of the aims was to refine the data collection process; we can draw the following lessons for improving the method. *First:* The importance of being clear about whether technical, communications or procedural knowledge is being exchanged (and being clear whether the relationship is perceived as teaching, sharing or learning). *Second:* Being clear about the relationship between roles and individual: those who communicate digitally do so in dialogue with their offline colleagues. *Third:* The most interesting data might relate to the people who are not (yet)

part of any CoP. It is therefore important to ensure the research method is open to identifying boundary objects *and* 'boundary people' as part of the characterisation of the community. *Fourth:* The interview format is essential because respondents may not understand the distinctions between *technical* skills and *writing content skills.* Also, semi-structured interviews can be used to gather important qualitative data about *how* knowledge-sharing takes place, what inspires it, whether it is valued, how and why the various actors became involved and so on. It is important that the learning impact of this research on participants is acknowledged.

5. Conclusions and issues for further research

As a small, time-limited pilot project, the results are constrained in a number of ways, particularly the small number of people who could be interviewed or provide data. However, the tentative results show that there is value in using the CoP model as a lens for analyzing the sustainability of online activity: it provides a framework for putting sustained technological acceptance by hyperlocal democratic practitioners into a social context. Refining the data has given an abstract model (Figure 1) for visualizing the sharing the technical skills of creating online content and skills around digital communication. Membership of this CoP, whether conscious or not, does appear to increase self-efficacy in the participants.

This research has uncovered a number of overlapping communities in one city. However, this project was not able to conclusively establish whether there is a core CoP of community councilors practicing digital engagement, or whether one could be intentionally constructed. A crucial test for further research would be to evaluate the extent (legitimated) peripheral participation can be observed and to do more to identify boundary objects. This is challenging because it seems a boundary between any CoP and the peripheral actors has not (yet) been formed, defined or recognized: members may not be aware they are in a CoP. A related challenge is understanding transitions into (and out of) a community such as this with its blurred or transient boundaries – where individuals move between practice, learning and interest. It would be interesting to investigate whether and how community councils with stable and churning online presences differ.

The work can be extended through the identification of similar communities elsewhere: geographically and on other platforms, Facebook in particular.

The question remains as to whether this will lead to sustained engagement: this would require a sustained study.

Finally, we look forward to applying this approach in a larger context and to evaluating the method in similar contexts across the UK and Europe. As well as allowing for wider sharing of good practice, this would allow a more rigorous model to be constructed.

Acknowledgement

References

[1] S. Vissers and D. Stolle, "The Internet and new modes of political participation: online versus offline participation," *Information, Commun. Soc.*, vol. 17, no. February 2015, pp. 937–955, 2013.

[2] N. Ellison and M. Hardey, "Social Media and Local Government: Citizenship, Consumption and Democracy," *Local Gov. Stud.*, vol. 40, no. February 2015, pp. 21–40, 2014.

[3] I. Susha and Å. Grönlund, "eParticipation research: Systematizing the field," *Gov. Inf. Q.*, vol. 29, no. 3, pp. 373–382, Jul. 2012.

[4] Ø. Sæbø, J. Rose, and L. Skiftenes Flak, "The shape of eParticipation: Characterizing an emerging research area," *Gov. Inf. Q.*, vol. 25, no. 3, pp. 400–428, Jul. 2008.

[5] A. Chadwick, "Explaining the Failure of an Online Citizen Engagement Initiative: The Role of Internal Institutional Variables," *J. Inf. Technol. Polit.*, vol. 8, pp. 21–40, 2011.

[6] J. Saglie and S. I. Vabo, "Size and e-Democracy: Online Participation in Norwegian Local Politics," *Scan. Polit. Stud.*, vol. 32, no. 4, pp. 382–401, Dec. 2009.

[7] S. Davidson and S. Elstub, "Deliberative and participatory democracy in the UK," *Br. J. Polit. Int. Relations*, vol. 16, no. 3, pp. 367–385, 2014.

[8] P. Cruickshank, B. M. Ryan, and C. F. Smith, "'Hyperlocal e-democracy'? The experience of Scotland's Community Councils," in *CeDEM Proceedings*, 2014, no. 2011.

[9] B. M. Ryan and P. Cruickshank, "Scottish Community Councils online: the 2014 survey," Edinburgh, 2014.

[10] P. Cruickshank, B. M. Ryan, and C. F. Smith, "Disconnectd Democracy? A survey of Scottish Community Councils' online presences," *Scottish Aff.*, vol. 23, no. 4, pp. 486–507, 2014.

[11] Österreichischer Gemeindebund, "Strong municipalities – Europe's driving force," 2003. [Online]. Available: http://www.gemeindebund.at/rcms/upload/7_60_1_eng.pdf. [Accessed: 09-Jun-2015].

[12] Ministry of Local Government and Regional Development, "Local government in Norway," 2003.

[13] A. V. Haug and A. Jansen, "The window of opportunity for e-democracy is wide open. Some take advantage of it, some don't ! Why?," pp. 1–15, 2003.

[14] I. Volan, "Norway: almost 40% of municipalities use Facebook," *Social Media Nordic*, 2011. [Online]. Available: http://socialmedianordic.com/2011/09/09/norway-almost-40-percent-of-municipalities-use-facebook.

[15] N. Ellison and M. Hardey, "Developing political conversations?," *Information, Commun. Soc.*, vol. 16, no. 6, pp. 878–898, Aug. 2013.

[16] S. Rummens, "Staging Deliberation: The Role of Representative Institutions in the Deliberative Democratic Process," *J. Polit. Philos.*, vol. 20, no. 1, pp. 23–44, Mar. 2012.

[17] P. Matthews, "The longue duree of community engagement: New applications of critical theory in planning research," *Planning Theory*, vol. 12, no. 2. pp. 139–157, Aug-2012.

[18] A. Macintosh and A. McKay-Hubbard, "Renewing democracy with 'e-community councils,'" in *e-Society*, 2004, vol. II, no. July, pp. 808–811.

[19] E. Wenger, *Communities of Practice: Learning, Meaning and Identity*. Cambridge: Cambridge University Press, 1998.

[20] E. Bolisani and E. Scarso, "The place of communities of practice in knowledge management studies: a critical review," *J. Knowl. Manag.*, vol. 18, no. 2, pp. 366–381, Apr. 2014.

[21] E. Wenger, "Communities of practice and social learning systems: the career of a concept," in *Social Learning Systems and Communities of Practice*, C. Blackmore, Ed. London: The Open University, Springer, 2010, pp. 179–198.

[22] S. Leigh Star, "This is Not a Boundary Object: Reflections on the Origin of a Concept," *Sci. Technol. Human Values*, vol. 35, pp. 601–617, 2010.

[23] S. Denning, *A leader's guide to storytelling: mastering the art and discipline of business narrative*. San Francisco: John Wiley, 2005.

[24] E. Davenport and H. Hall, *Knowledge Discovery Organizational Knowledge and Communities of Practice*, vol. 36. Medford, NJ, NJ: Information Today, 2002.

[25] E. Wenger, R. McDermott, and W. M. Snyder, *A guide to manageing knowledge: Cultivating Communities of Practice*. Boston: Harvard Business Press, 2002.

[26] J. Lave, "Situated learning and changing practice," in *Community, economic creativity, and organization*, Oxford: Oxford University Press, 2008, pp. 283–296.

[27] S. Coleman and J. Gøtze, *Bowling Together : Online Public Engagement in Policy Deliberation*. London: Hansard Society, 2001.

Collaborative Problem Solving for Regional Development? – A New Approach of Public Participation

Dana Mietzner[a,1], Ralf Vandenhouten[b] and Frank Hartmann[a]

[a] *Technical University of Applied Sciences Wildau, Research Group Innovation and Regional Development, Hochschulring 1, D-15725 Wildau*

[b] *Technical University of Applied Sciences Wildau, Research Group Telematics, Hochschulring 1, D-15745 Wildau*

Abstract. In this paper, we present a concept prototype for a socio-technical assistant (s.TAS) as a collaborative problem-solving approach to support the regional development in a selected rural district. The presented concept considers insides from citizen participation as an open government instrument and strongly incorporates collaborative value creation to address regional development issues. The presented approach incorporates citizens and other regional stakeholders (e.g., companies, organizations, associations) as problem-solvers and strongly enables regional stakeholders and the public administration to develop solutions jointly as an ongoing learning process. In this sense, s.TAS addresses participation approaches and elaborates participation to collaboration. For the development of this participatory problem-solving approach, we conducted (1) a broad literature review on participatory approaches and (2) we implemented a qualitative, collaborative research approach with participants from different stakeholder groups. In several workshop rounds, supported by interviews and surveys we developed a concept prototype for the socio-technical assistant s.TAS, which is a technical, smart system, which supports information, communication, participation and collaboration in the selected region.

Keywords. Participatory approach, socio-technical assistant, open government, regional development, collaboration

1. Motivation

This paper presents an ongoing research project about the development of a collaborative problem-solving tool, which should support rural areas, regions or communities to deal with challenges provoked by the demographic change and its consequences for rural areas. The research project is conducted collaboratively with a selected rural district in Germany, research institutions, companies and citizens.

"Demographic change in Germany is marked by low birth rates and a declining population size. Increased life expectancy, the resulting ageing of the population and the growing proportion of the population with an immigrant background affect Germany more than other industrial countries. Demographic change impacts almost every area of life and will significantly influence our society and economy in the coming decades."[1]

---

[1] Corresponding Author.

With regard to demographic change, we very often problematize higher costs in terms of care and mobility, the necessity of technical assistant systems to compensate physical detracting, the (social) isolation of rural areas and the overpopulation of urban cities. In our research, we address the aging society with its professional and life experiences as a chance and potential to participate in the solution of demographic change induced challenges. We address a generation of elderly people, who are longer in good health conditions, active and already experienced with technical systems. From our perspective, this generation is able and willing to participate in a collaborative development of the region they live in by using a socio-technical assistant. Furthermore, we aim to introduce a participatory instrument, which supports collaborative problem solving, which is not limited to information delivery or communication tools for public administrations only.

2. Theoretical Considerations

In order to explore new instruments of participation and collaborative problem-solving approaches a common understanding of open government, and public participation as an open government instrument is necessary. The term "open government" contains a variety of nuances of meanings [2]. Open government is used as synonym for open data (see IT-Planungsrat 2010), other authors rather link open government to open innovation (Noveck 2011) or emphasize citizen participation (Bauer/Seckelmann 2012) or a comprehensive integration of all societal stakeholders (Krabina/Lutz/Prorok 2012) [2]. Other sources link the "open government" movement to the memorandum of the US-President Barack Obama (McDermott, 2010) [3], who defines in the memorandum in 2009 open government as a triad of transparency, participation and collaboration [4]. In our research, we follow this understanding of open government in terms of transparency, participation and collaboration. In this sense, transparency includes the delivery of data and public operations, while participation and collaboration address the integration of citizens in decision-making and service provision processes of the public administration [5]. Through this, on the one hand, the public administration operations are legitimized; on the other hand, information, knowledge and citizens competencies are incorporated into the administration.

2.1. Citizen Participation as an Open Government Instrument

The participation of citizens allows partaking in political decision-making processes in several forms and on different political levels. The instruments of citizen participation have been developed continuously over the last years. Beside rather traditional instruments, which are incorporated in administrative proceedings or building legislations and obligatory, a variety of new facultative participatory instruments can be implemented [6]. Two forms of citizen participation can be distinguished: (1) direct-democratic and (2) deliberative forms of citizen participation. In terms of direct-democratic forms of participations, citizens vote directly for or against certain issues. In terms of deliberative forms of participation, citizens participate by exchange of arguments in discussions. Citizen participations are also characterized by the elements information, communication, and participation and consulting. Nowadays, there is a variety of different citizen participation instruments (e.g., wikis, scenario workshops, forums, open space conferences, bar camps, deliberative polling, citizens' panels, citi-

zens' jury), with different aims, number and kind of participants, duration, costs, expected outcomes and not every instrument suits certain issues and conditions. Each citizen participatory instrument must be chosen carefully due to the intended aims, costs and benefits [7]. Nevertheless, there are obstacles to overcome when it comes to proper implementation of citizen participation on a regular basis. Opinions, arguments and participations of citizens have to be taken seriously and there is a need for proper conflict moderation. Following decisions need to be transparent and comprehensible, especially in terms of contrary decisions. The value of citizen participation for public administrations and politics is obvious; decisions are legitimized and information can be delivered in an appropriate way. If public administration take citizens seriously, positive image effects for the public administration can be a consequence [8]. To sum it up, the discussion of citizen participation instruments indicates that information, communication and participation are key elements of existing participatory instruments while the problem-solving itself by citizens is not in the centre of participatory instruments. There is a need for collaborative problem-solving approaches, which allow a joint development of solutions by all stakeholder groups as well as the implementation of solutions. Furthermore, in our approach problem-solving is an ongoing process or cycle rather than a selective implementation of individual participatory instruments.

2.2 Technical Assistant Systems for Problem Solving

Technical Assistant Systems as a human-technical interaction approach, with interdependent effects of technical processes and social behavior [9], are still addressed as future issues, which needs to be further explored [10]. In some learning, working and living environments, assistant systems are already implemented, e.g., in the area of welfare-technologies [10], in terms of systems for the preservation of qualifications and competencies of employees [11], systems for driver assistance or systems for decisions making support, and the development of alternatives [12]. The status of marketability and implementation differs significantly between the different areas of application. While driver assistance is widely accepted, welfare technologies are still confronted by a low level of marketability [10]. In our study, we address a technical assistant system, which should not be focusing on the support of individuals or organizations, like driver assistance. The system aims to support the solving of comprehensive, societal problems with regard to demographic change on the level of regions.

3. Object of Investigation

In order to further explore new participatory approaches, the role of citizens and their ability to solve regional development problems collaboratively with the public administration and other regional stakeholders (e.g., companies in the regions, nonprofit and private organizations), we address a rural district in Germany confronted with specific regional development problems. The selected district with 160.000 inhabitants is highly affected by demographic change and its consequences, like an aging society with 43.2% of inhabitants over the age of 50 [13], the isolation of rural areas, corresponding new requirements for infrastructure, mobility concepts and logistics. On the other hand, the selected region also encompasses "hot spot areas" characterized by a significant growth in terms of society and economic development with a high demand for a skilled workforce, living spaces and other needs of infrastructure. Overall, there

are several burning issues to address and to solve by the affected stakeholder groups and a need for collaborative problem-solving approaches.

4. Methodological Approach

For the development of a participatory problem-solving approach, we conducted (1) a broad literature review on participatory approaches and (2) we implemented a qualitative, collaborative research approach with participants from different stakeholder groups. In several workshop rounds, supported by interviews and surveys we developed a concept prototype for a socio-technical assistant (s.TAS), which enables and supports citizens to solve regional development problems collaboratively with the public administration and other regional stakeholders. In this iterative process, we included two research institutes, representatives from the public administration, citizens as well as companies and organizations with different perspectives on the topic. During idea workshop rounds, we involved 29 people (17 from the rural district administration; two persons as city representatives; 5 company partners with technical background and 5 members from the responsible research group). We implemented creativity techniques, design thinking [14], [15] and other user centric approaches for the generation and assessment of ideas for the development of a participatory tool. The aim of the workshop rounds was to explore requirements and use cases for a new problem-solving approach. For this purpose, we introduced and explored different personas (user-centric approach) and the workshop participants developed jointly ideas to deal with the persona's circumstances and requirements.

The idea generation was supported by key questions in order to trigger and to guide the idea generation during several brainstorming sessions. Sub-groups collected 147 rough ideas by answering several key questions: e.g. "Which issues has s.TAS to address in order to encourage persona 1 to participate enthused?" or "What has s.TAS to address in order to make Persona 2 telling his friends about s.TAS?" or "How can the public administration supported by s.TAS help Persona 2 to develop leisure facilities in his home town?" or "How can s.TAS support Persona 3 to develop a small and desolated village into a mobility hot spot with attractive new transport solutions?"

In further rounds, the workshop participants and the facilitator clustered, evaluated and consolidated the outcome of the idea generation. As a key result, the discussion was centered around three main topics for s.TAS. (1) easy handling, (2) contact, information and participation and (3) rules. A multitude of participants emphasized the need for an easy-to-use intuitive approach for s.TAS. Under this point, the participants discussed certain requirements like multilingualism, easy access possibilities via internet, mobile devices and other forms like terminals in appropriate areas. Furthermore, participants emphasized the necessity for s.TAS to act as a contact and knowledge broker. In this sense, s.TAS should deliver information, provide solutions, and enhance citizen participation and joint problem solving. Workshop participants also agreed that s.TAS has to have rules. For example, the users of the system should have profiles, clearly indicating the competencies, skills or experiences of the user with explicit rights to read and write or administrate. In another workshop round the s.TAS system was roughly sketched by the involved participants. In that way, first ideas for its technical implementation could be collected and discussed as a fundament for the development of a first concept prototype.

5. New Problem-Solving Approach with s.TAS

In this section, we present the derived conceptual prototype and the key functions of a socio-technical assistant (s.TAS) for collaborative problem solving. S.TAS is conceptualized as an action system, in which different human and technical subsystems are integrated comprehensively [16]. In general, s.TAS is a technical, smart system, which supports information, communication, participation and collaboration in the selected region. Furthermore, s.TAS is supported by "face-to-face" participatory instruments with linkages to the technical system and vice versa. s.TAS is open for all stakeholder groups and has to meet the needs for several stakeholders ("design for all") in terms of accessibility and usability. s.TAS will be supported by an internet website, mobile applications via smartphone or tablet, but s.TAS should be also available on an appropriate terminal in public areas. s.TAS allows three key functions (1) information and communication platform, (2) participation and (3) collaboration.

As an information and communication platform, s.TAS allows the public administration to present information in an appropriate way; citizens, companies, and other stakeholder groups receive information; questions and answers; ongoing discussions will be supported as well.

As a participation assistant s.TAS will coordinate participatory processes and moderate collaborative approaches between different stakeholder groups. In terms of collaboration s.TAS allows the joint development and implementation of solutions between relevant stakeholder groups. A key element within the s.TAS approach is the learning, dynamic knowledge base (see figure 1), which contains data and information (statistics, videos, data etc.). Furthermore, knowledge regarding certain issues, best practice solutions or any problem-solving information will be contained as well. In this sense, the system learns continually. This knowledge is based on the knowledge of all stakeholder groups who are willing to support and enrich a certain issue. Solutions can be delivered by experts (public administration, companies, and research institutions) but also by laypersons with creative input. Each stakeholder has its own profile and competencies. Supported by a messenger system, stakeholders receive enquiries or invitations to participate in the problem-solving process.

For each problem-solving process, a task force is formed and moderated. This procedure allows to collect and present necessary information for a certain problem-solving process and it allows to bring together an appropriate set of people as a task force. The problem-solving process itself will be supported by the s.TAS but also by offline instruments.

The technical system provides tools for orchestrating the collaboration process as well as for editing and managing the various kinds of artefacts arising out of the solution-finding process, like sketches, text documents, tables, technical specifications, audio or video files, mathematical models, construction plans etc. s.TAS will also provide methods for a collaborative evaluation of contributions and (partial) solutions by means of innovative interactive visualization techniques as well as collaborative virtual and augmented reality technologies. Generated solutions will be implemented in the real world and the developed solutions will enrich the knowledge base for similar or related issues.
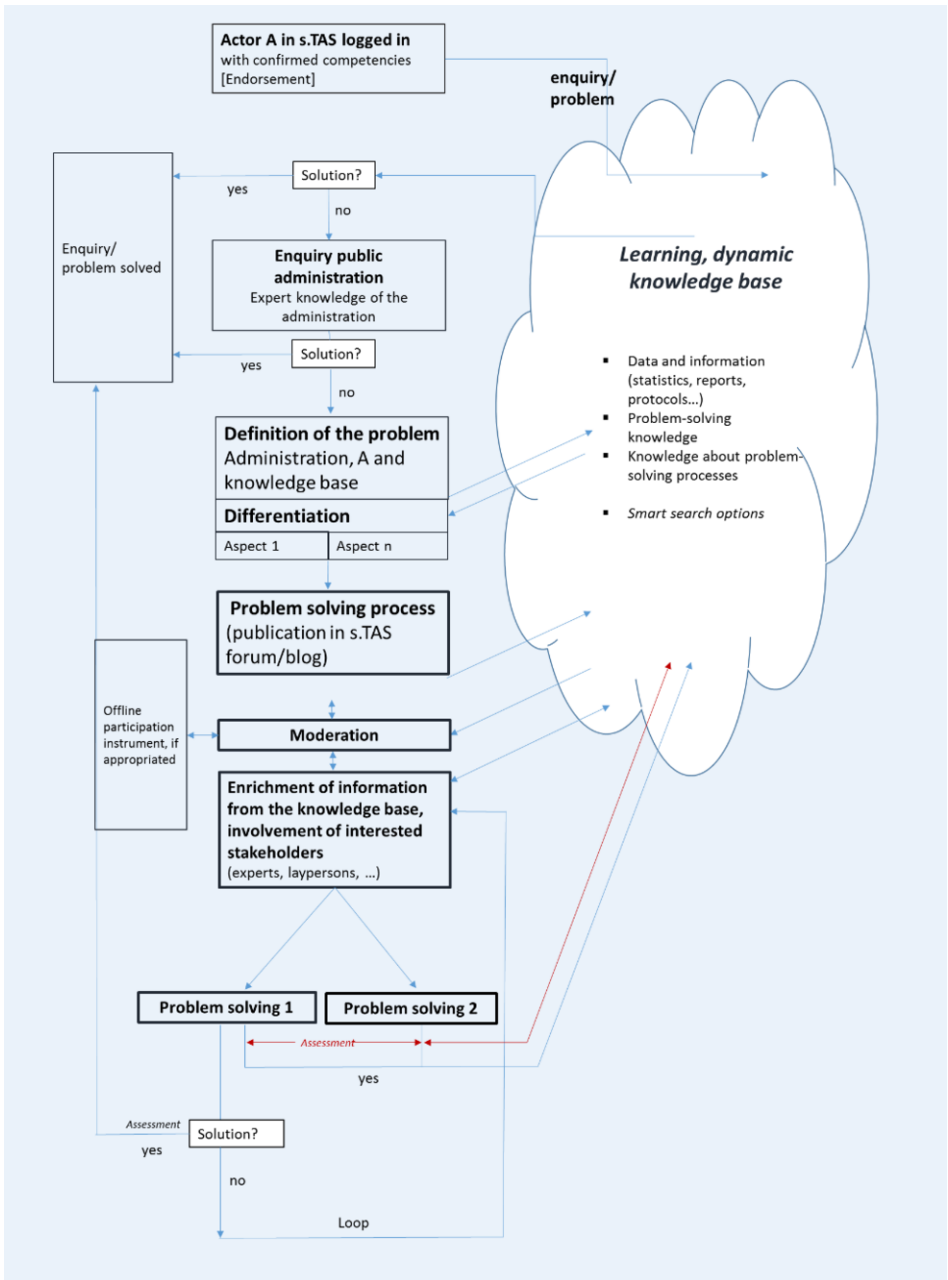
Figure 1. Concept Prototype

6. Conclusion and Further Research

In this paper, we present a concept prototype for a socio-technical assistant as a collaborative problem-solving approach to support the regional development in a se-

lected rural district. The concept considers insights from citizen participation as an open government instrument and strongly incorporates collaborative value creation to address regional development issues. This approach incorporates citizens and other regional stakeholders (e.g., companies, organizations, associations) as problem-solvers and strongly enables regional stakeholders and the public administration to develop solutions jointly as an ongoing learning process. In this sense, the approach addresses participation approaches and elaborates participation to collaboration. In this paper, we shortly present key results of ongoing research, which has to be deliberated further in the future. Future research will emphasize several research questions. Firstly, we have to explore how stakeholders can be motivated to contribute in the problem-solving process. Secondly, we have to consider possible barriers within the public administration and corresponding measure to address and conquer identified barriers. Thirdly, we have to design a development process, which allows regional stakeholders to accompany the technical development of s.TAS in a systematic way in order to ensure implementation and usage.

References

[1] GERMAN FEDERAL MINISTRY (Bundesminsterium des Innern). 2011. Federal Government Report on the Demographic Situation and Future Development of Germany (Demography Report). Avail-able: http://www.bmi.bund.de/SharedDocs/Downloads/DE/Themen/Politik_Gesellschaft/DemographEntwick lung/demografiebericht_kurz_en.pdf;jsessionid=09766EA98190CF689758E9DE95D1D3B3.2_cid364?__blob=publicationFile (Accessed 30.03.2015).

[2] HERZBERG, J. 2013. "Open Government"-Versuch einer Begriffsbestimmung. Verwaltung und Management, 19. Jg., 40-44.

[3] HILGERS, D. 2012. Open Government: Theoretische Bezüge und konzeptionelle Grundlagen einer neuen Entwicklung in Staat und öffentlichen Verwaltungen. Zeitschrift für Betriebswirtschaft, 82, 631-660.

[4] OFFICE OF MANAGEMENT AND BUDGET (OMB). 2009. Open Government Directive, M-10-06 [Online]. Washington, DC: U.S. Office of the President. Available: http://www.whitehouse.gov/omb/assets/memoranda_2010/m10-06.pdf [Accessed 17.12. 2014].

[5] SCHULZ, S. E., JANDA, T. C. & TISCHER, J. 2013. Offene Schnittstellen in der Kommunikation zwischen Bürger und Staat - OPEN GOVERNMENT Fact Sheet#1. In: INTERNET & GESELLSCHAFT COLLABORATORY (ed.). Berlin.

[6] NANZ, P. & FRITSCHE, M. 2012. Handbuch Bürgerbeteiligung: Verfahren und Akteure, Chancen und Grenzen, Bonn, Bundeszentrale f. Politische Bildung. OFFICE OF MANAGEMENT AND BUDGET (OMB). 2009. Open Government Directive, M-10-06 [Online]. Washington, DC: U.S. Office of the President. Available: http://www.whitehouse.gov/omb/assets/memoranda_2010/m10-06.pdf [Accessed 17.12. 2014].

[7] HIERLEMANN, D., VEHRKAMP, R. B. & BERTELSMANN, S. 2010. Politik beleben, Bürger beteiligen: Charakteristika neuer Beteiligungsmodelle, Gütersloh, Verl. Bertelsmann-Stiftung.

[8] BERTELSMANN STIFTUNG & STAATSMINISTERIUM BADEN-WÜRTTEMBERG. 2014. Partizipation im Wandel Unsere Demokratie zwischen Wählen, Mitmachen und Entscheiden [Online]. Gütersloh: Verlag Bertelsmann Stiftung. Available: http://public.eblib.com/choice/PublicFullRecord.aspx?p=1784165FÖHR, T. 2014. Intensität der Beteiligung [Online]. Available: https://www.flickr.com/photos/wissenstransfer/13040113985/ [Accessed 02.02. 2015].

[9] JAHNKE, I., HERRMANN, T. & METZ-GÖCKEL, S. 2006. Organisation als soziales und soziotechnisches System. In: JAHNKE, I., HERRMANN, T. & METZ-GÖCKEL, S. (eds.) Dynamik sozialer Rollen beim Wissensmanagement. DUV.

[10] BUNDESKANZLERAMT 2012. Dialog über Deutschlands Zukunft - Ergebnisbericht des Expertendia-logs der Bundeskanzlerin. Berlin.

[11] MEIER, H. & KRÜCKHANS, B. 2012. Innovatives technisches Assistenzsystem zur Optimierung der Arbeitsgestaltung. In: MÜLLER, E. (ed.) Demographischer Wandel: Herausforderung für die Ar-beits- und Betriebsorganisation der Zukunft. Berlin.

[12] BLUTNER, D., CRAMER, S., KRAUSE, S., MÖNKS, T., NAGEL, L., REINHOLZ, A. & WITTHAUT, M. 2009. Assistenzsysteme für die Entscheidungsunterstützung. In: BUCHHOLZ, P. & CLAUSEN, U. (eds.) Große Netze der Logistik. Springer Berlin Heidelberg.

[13] STATISTISCHES BUNDESAMT (2011): Zensus 2011. Bevölkerung nach Alter und Geschlecht für Gemeinden.

[14] DUNNE, D. & ROGER, M. 2006. Design Thinking and How it will Change Management Education: An Interview and Discussion." Academy of Management Learning and Education 5 (4): 512-523.

[15] GROTS, A. & PRATSCHKE, M. 2009. Design Thinking — Kreativität als Methode. Marketing Review St. Gallen, 26, 18-23

[16] ROPOHL, G. 2009. Allgemeine Technologie. Eine Systemtheorie der Technik. Universitätsverlag Karlsruhe, Karlsruhe. SCHULZ, S. E., JANDA, T. C. & TISCHER, J. 2013. Offene Schnittstellen in der Kommunikation zwischen Bürger und Staat - OPEN GOVERNMENT Fact S heet#1. In: INTER-NET & GESELLSCHAFT COLLABORATORY (ed.). Berlin.

Who are the Internet Voters?

Uwe SERDÜLT, Micha GERMANN, Maja HARRIS, Fernando MENDEZ and Alicia PORTENIER[a]

[a] *All authors affiliated to Centre for Democracy Studies Aarau (ZDA) at the University of Zurich, Aarau, Switzerland*

Abstract. Assessing the influence that socio-economic characteristics have on the division between traditional voters and those who choose to vote via the internet is crucial to political debate as well as for the future development of democracies. Does the introduction of internet voting technology simply widen the divide between voters and non-voters, further isolating the part of the electorate already underrepresented in the political process? We address these issues by reviewing the current state of research in 22 empirical studies relating internet voting to socio-economic variables. The results are not homogeneous but suggest that although socio-economic factors do play an important role in explaining the choice of voting channel, they are strongly moderated by the general use of and trust in the internet.

Keywords. Digital divide, internet voting, voting technology, socio-demographic variables

1. Introduction

Internet voting (i-voting) has been considered, discussed and implemented in several countries since the late 1990s [26]. The development and success of i-voting in each country differs greatly. For instance, while the experience in the USA has been limited to individual trials in primary elections [30], [38], in Estonia, or 'E-stonia' as it has been named, the entire electorate can vote online in national elections [39]. In Switzerland, pilot projects have been carried out in Cantons Geneva, Zurich and Neuchâtel, beginning in 2003 in the Geneva municipality of Anières. Since then, other cantons have also introduced i-voting [36]. In some countries such as Norway, i-voting projects have been stopped completely after several trials [4].

Although the introduction of this additional voting channel clearly has some advantages, such as reducing costs and simplifying the voting and vote counting processes [22], i-voting undeniably raises various technical, legal and especially political questions. One such question is that of the socio-economic characteristics of i-voters. I-voting is perceived as offering a new way to reach the population and involve those underrepresented in the electorate in the political process. This argument is based primarily on the hope for increased participation of younger voters as they have a higher affinity for the internet compared to other age groups [20], but a generally low turnout rate [22].

Increasingly, there is also a fear that socio-economic groups that are already underrepresented in the electorate would be further disadvantaged in the political process. It is argued that i-voting would predominantly be utilised by those groups who

already exhibit a high participation rate without the introduction of this additional voting method. Well-educated, affluent citizens are more likely to have access to the internet and the knowledge required to navigate it than the less well-educated, low income socio-economic groups [23], a phenomenon often referred to as the *digital divide*. What is pivotal is that according to [28] the digital divide runs alongside existing socio-economic cleavages and can therefore exacerbate existing inequalities. It must be noted that in the long term the existing distortion can be changed or at least softened, and in the best case scenario corrected altogether [6]. Using data from the USA, [1] illustrates that the digital divide tends to diminish over time. Among other things, the data refers to the percentage of households in certain income brackets that own a computer. While in households with an income of under US$25,000 p.a. there was an increase of 80.3 percent between 1977 and 2000, the increase was only 15.7 percent in households with a yearly income of over US$75,000.

Aside from the much discussed and argued *digital divide*, the sociological explanatory model for voting behaviour, developed by researchers at the University of Columbia in the 1940s, remains one of the main theories in this field. It is based on an analysis of interviews of 600 people during the 1940 presidential election campaign by a research team led by Paul F. Lazarsfeld to document the opinion-forming process of the interviewees. The principal argument of their analysis was that a person's vote is highly dependent on their socio-economic characteristics. The ensuing criticism, that the theory lacked a causal connection or had its own social determinism, led to the advancement of the ideas of the Lazarsfeld group in the following decades. For example, [24] developed a macro-sociological approach explaining the party systems as well as the party affiliation of individuals on the basis of four conflict structures in society arising from national revolutions and the industrial revolution: centre and periphery as defined by ethnicity and language; State and church; city and country; and labour and capital. Today the socio-economic composition of the electorate is still considered relevant to the election result [32].

For various reasons the socio-economic profile of i-voters is also of interest. Compared to existing voting channels like ballot or post, are certain population or age groups over or underrepresented among i-voters? Do certain socio-economic characteristics have a detectable influence on the choice of voting channel? I-votes conducted so far in Switzerland and Estonia, but also in the USA, United Kingdom, Norway and Canada, have allowed for these questions to be empirically explored.

This paper summarises the current state of research by analysing and comparing the 22 known studies with reference to underlying methods and results. The focus is on the socio-economic dimensions of age, gender, income and education and how these dimensions influence the choice of voting channel. Though several studies examined a multitude of additional information, these four variables were included in enough of the studies to compile a meaningful comparison.

The fundamental methodology and variables examined vary from analysis to analysis. In most of the studies, i-voters are compared with voters using other channels. Only a few isolated studies limit their analysis to the profile of i-voters without any comparison. Central to the interpretation and structuring of the results is the methodology used in each study. This paper differentiates between bivariate and multivariate analyses, each made up of two categories respectively. Not all studies refer to the models using our terminology, however every analysis can be grouped into one

of the four categories. The bivariate studies are limited to the relationship between a specific socio-economic variable and the choice of voting channel and in most cases also describe the frequency distribution. Rarely, an additional chi-squared test is carried out to test the significance of the differences found. The multivariate models consider multiple independent variables. They therefore answer questions such as: Does age have a significant influence on the choice of voting channel when the influence of gender, income and education is taken into account at the same time? Again, this category of analyses differentiates between socio-economic and global models. In the first, only socio-economic information is included.

Variables such as marital status or social class are also examined in some studies. The global models consider above all variables that relate to politics and IT such as political orientation, computer knowledge or trust in the internet. Although this paper does not focus explicitly on these variables, the corresponding results are included as they can provide indications for further explanatory approaches for the choice of a specific voting channel. Certain coefficients that show significant values in socio-economic models can lose their significance in global models. For example, it is obvious that better computer knowledge and a higher trust in the internet are typical for the younger population. This means that they do not use i-voting because they are young but rather because compared to other age groups they have a higher affinity to the internet. The global models can therefore put the significance of socio-economic dimensions into perspective and provide new explanatory approaches.

2. Switzerland

Who uses i-voting in Switzerland? What are their socio-demographic characteristics? Six studies have been conducted so far for the Swiss case. They give a first indication as to whether the variables age, gender, income and education have an impact on the use of i-voting. Table 1 gives an overview of these studies.

2.1. Age

The studies reveal by and large a uniform picture regarding age. According to bivariate measures, younger voters tend to use i-voting in greater numbers than older voters. Interestingly, it is not the youngest voters (18-29 years) who use i-voting most often but rather voters in the 30-49 years age group [10, 34]. The oldest age group, in turn, only rarely makes use of i-voting. Likewise, a multivariate analysis [34] shows that it is the second youngest electorate group that uses i-voting most frequently. Crucially, age loses its statistical significance when controlling for the frequency of internet usage and trust in the internet. This indicates that younger voters tend to use i-voting more frequently because they currently show a greater affinity for the internet. Consequently, differences between age groups with regard to i-voting usage could fade over time as it depends not on age itself but rather on affinity for the internet.

2.2. Gender

All analyses which are limited to the description of frequencies conclude that men are overrepresented among i-voters [10, 11], [33, 34], [37]. [37] shows that 35.7 percent of

men use i-voting compared to merely 16.4 percent of women, whereas [34] demonstrates with 36.7 percent to 30.2 percent respectively, a less stark contrast between the genders. Performing a chi-squared test, however, [18] finds no statistically significant effect for gender. Similarly, multivariate analyses [11], [33] show no significant effect of gender once the variables computer skills and frequency of internet usage are controlled for.

Table 1. Overview of studies examining socio-demographic characteristics of i-voters in Switzerland.

Study	Results
Christin and Trechsel [10]	Based on *bivariate* analysis **younger** (under 40, although not the youngest), **male** and **higher educated** voters are overrepresented among i-voters.
Christin and Trechsel [11]	Based on *bivariate* analysis i-voting is used more frequently by **younger**, **male** voters with a **high education** level and a relatively **high** household **income**. The *multivariate* model including socio-demographic and ICT variables produces significant results only for the latter variables.
Serdült and Trechsel [37]	The following results are found based on *frequency distributions*: voters aged **40-49** followed by voters aged **18-39** are overrepresented among i-voters. **Men** and voters with a **higher education** and an **above average income** are also overrepresented.
Serdült [34]	Based on *bivariate* analysis, **30-39** year old voters most often choose to use i-voting, followed by the **40-49** year old group. **Men** and the **better educated** are overrepresented among i-voters.
Sciarini et al. [33]	*Frequency distributions* show that voters between **25 and 34** more frequently use i-voting. 7% more **men** than women chose to vote online. **Better educated** (university degree) voters and voters living in **high income households** are overrepresented. *Multivariate* models show significant effects of **age, gender** and **education** on i-voting. However, the model including frequency of **internet use** and **trust in the internet** explains away effects of all socio-demographic variables.
Germann et al. [18]	Based on *chi-squared tests* **age** appears unrelated to i-voting use. **Men** are more likely to vote via the internet. I-voters are significantly more tech-savvy (**IT-skills**).

2.3. Income

Bivariate results show that higher income voters more frequently use the internet to vote [11], [33], [37]. As [37] shows, the share of i-voters grows as income increases, but at the same time the share of voters who use the two established voting channels (postal and ballot voting) decreases. Similarly, [33] observes greater shares of i-voters compared to postal or ballot voters among middle and higher income households. Although [11] finds a general positive association between income and i-voting, no clear impact of income on the use of the traditional voting channels is observed. In a multivariate analysis, income proves to be positively related to the probability to vote via the internet, but loses its significance as soon as IT-variables such as frequency of internet usage or trust in the internet are controlled for.

2.4. Education

Similar to age and income, education is found to be positively related to i-voting in bivariate analyses. Among i-voters, the highly educated are clearly overrepresented.

For instance, [37] found only 2.8 percent of i-voters indicated compulsory education as their highest education level, whereas 29.5 percent indicated they had a higher vocational education and 36.1 percent a university degree. In a comparison of voting channels (internet or postal) used by Swiss living abroad from the cantons Aargau, Basel-City, Graubünden and St. Gallen, a chi-squared test produces no significant results for education [18]. Finally, a multivariate analysis results in a familiar pattern: as soon as the frequency of internet usage and trust in the internet are introduced into the analysis, education loses its effect [11], [33]. When looking at frequency distributions therefore, i-voters display distinct socio-demographic characteristics. They are generally more highly educated, between 30 and 49 years old, male and have an above average income. However, as the multivariate models show, possible effects of socio-demographic variables on i-voting are all explained away by measures of internet usage and attitudes towards the internet.

3. Estonia

Since 2005, i-voting has been available to the entire Estonian electorate, not only at a local level but also for national and European parliamentary elections [26], [39]. Five studies were conducted: [39] and [6] examined data from the 2007 parliamentary election; [8] focused on the 2005 local elections; [2] considered both of the abovementioned elections; [40] relied on data from five different elections: local elections from 2005 and 2009, the 2007 and 2011 national elections and the 2009 European parliamentary election. [2] was limited to bivariate analyses, [6] and [40] to multivariate analyses. [8] and [39] first undertook bivariate analyses, followed by multivariate analyses. Table 2 gives an overview of these studies. The studies all compared i-voters with traditional voters. Only [2], in a section examining electoral roll data, limited the analysis to i-voters.

3.1. Age

The results for age support those produced in the Swiss studies - that it is generally younger voters, though not the youngest, who choose to use i-voting. The bivariate studies, which describe the distribution of this variable, point out that younger voters tend to be overrepresented amongst i-voters. [8] and [2] show that the percentage of i-voters increases from the youngest age category (18-29) to the second youngest (30-39 (25-29 for [2])), and then decreases. [39] also describes the relationship between i-voting and age as a curve. The turning point appears to be the 40-49 age group. Both [8] and [39] show that the distribution of ballot voters is different – here the percentage increases consistently with age.

The multivariate analyses confirm these results for the most part. [6] not only considers age as a variable on its own but uses the square value of the age. The results that support the bivariate reports give a positive coefficient for the single variable and a negative coefficient for the squared one. This means that the relationship between age and i-voting is non-linear. Initially, the probability of using i-voting increases with age. There is then a turning point when the probability reduces as age continues to increase. [8] and [39, 40] all generated a multivariate socio-economic model and a second multivariate global model. In the first model, all results produced a negative

coefficient; that the probability of using i-voting decreases with age. The global models give conflicting results: [8] and [40] found a positive coefficient. [40] deduces for all five models in the global evaluations (one for each election studied), that age, as soon as control variables such as trust in the internet are considered, has no significant influence on the choice of voting channel. The studies conclude that although i-voters are predominantly young, the reason for the increased use of i-voting is not due to age but rather to a greater affinity for the internet that is typical for the younger generation.

3.2. Gender

While the bivariate analyses [2], [8], [39] identify that men are overrepresented (slightly) amongst i-voters, the multivariate study results [6], [8], [39], [40] conclude that gender has no significant influence over the choice of voting method. The results from [8] and [39] show that the percentage of men voting online is significantly greater than those voting by ballot. [2] first limited its study to i-voters and found that 48.24 percent were female and 52.76 percent were male. [2] compares traditional voters with i-voters and confirms that the gender distribution for the two voting channels is identical. Socio-economic and global multivariate analyses [6], [8] and [39] conclude that gender has no statistically significant influence on the choice of voting method. While for [40] gender is insignificant overall, there are some exceptions. The authors apply both a socio-economic and global model for each of the analysed elections. In the first model, two of the five elections show that men tend to use i-voting more. The application of the global model for the 2011 parliamentary election likewise shows that men are overrepresented among i-voters.

In most cases there is therefore either no or very minimal indication of a gender gap. Although the bivariate analyses often determine a varying distribution, the multivariate models (with few exceptions) do not confirm this significance. In those studies where there are slight differences between the genders, it is always men who are overrepresented.

Table 2. Overview of studies examining socio-demographic characteristics of i-voters in Estonia.

Study	Results
Trechsel et al. (2007) [39]	*Bivariate* results show that i-voters are **young** (but not the youngest), **well-educated** and **high earning**. It is the **tech-savvyness** of the **young** that explains their use of i-voting. **Males** are overrepresented but the **gender bias** is not significant. Based on the *multivariate* socio-economic model, **young**, **well-educated**, **high earning** voters are overrepresented among i-voters. **Trust in the internet** increases the likelihood of i-voting. In the *multivariate* global model **income** and **education** cease to be significant. **IT skills**, rather than **age** increase the likelihood of i-voting. **Gender** plays no significant role in either of the multivariate results.
Bochsler (2010) [6]	Based on *multivariate* analysis **middle-aged** (**40-50** years), **well-educated**, **high earning** voters are overrepresented among i-voters. **Gender** is not significant.
Breuer and Trechsel (2006) [8]	*Bivariate* results show that **younger** (**under 40**, although not the youngest), **male**, **well-educated** voters with **very high incomes** are overrepresented among i-voters. It appears to be the **tech-savvyness** of the **young** that explains their use of i-voting. **Trust in the internet voting procedure** has a strongly significant impact in both *bivariate* and *multivariate* results.

	Based on the *multivariate* socio-economic model, i-voters are **younger**, **high earning** and **well-educated**. **Gender** is not significant. Based on the *multivariate* global model **gender**, **education** and **income** lose their significance. **Young** voters are overrepresented due to their **IT skills**.
Alvarez et al. (2009) [2]	Based on *bivariate* analysis **young**, **tech-savvy** voters are overrepresented among i-voters. **Trust in the internet voting procedure** is highly significant. **Gender**, **income** and **education** have little or no significance.
Trechsel and Vassil (2011) [40]	Based on *multivariate* socio-economic analysis, i-voters are **young**, **well-educated** and **high earning**. **Gender** is not significant. Based on *multivariate* global analysis, **age** (as soon as control variables such as **trust in i-voting** are considered), **education** and **gender** have no significant influence on choice of voting channel.

3.3. Income

[2] is the only study to find no difference in the choice of voting method between high and low income earners. The general tone of the other studies is that the probability of using i-voting increases as income increases. [8] and [39] find a similar distribution among ballot voters and i-voters in relation to income, though very high earners are overrepresented among i-voters. Interestingly, [39] finds that voters in the highest income bracket (over 10,000 Estonian crowns) are not overrepresented. The multivariate socio-economic models consistently show a positive correlation between income and i-voting [6], [8], [39], [40]. As shown with the age and gender variables however, this trend changes as soon as political or IT-variables are included. Only in [40] are positive significant coefficients delivered in three of the five models. The remaining two studies do not provide any statistically significant values with regard to income.

3.4. Education

The education variable delivers similar results to income. [2] again finds that there is no difference between i-voters and traditional voters. Bivariate analyses [8], [39] highlight that well-educated voters are overrepresented among i-voters. [8] finds that the most well-educated make up 60 percent of i-voters but only 30 percent of ballot voters. Results from [39] show a somewhat lower percentage for well-educated i-voters (50 percent). The multivariate studies that limit their models to socio-economic variables find a positive correlation between education and i-voting [6], [8], [39, 40]. The multivariate global models give, as in previous cases, less significant results. In [8] and [39] education loses all significance. It is the same for [40] in four of its five models, though in respect of the model based on the 2011 election, there is a negative coefficient for compulsory schooling and i-voting.

4. United States of America

The United States has had limited experience with i-voting. The four studies analysed here relate to two Democratic Party primary elections. One took place in Arizona in 2000 and [3] and [38] based their analysis on data from this election. [30] and [31] examined the profile of voters in the Democratic Party primary in Michigan in 2004.

Table 3 gives an overview of these studies. All four studies incorporate multivariate models. Only [38] also examines bivariate relationships. [3] defines the dependent variables as a percentage of the electronic votes in 15 districts. The remaining studies compare i-voters with traditional voters.

4.1. Age

Age produces significant results in all studies. [3] divides the variable into two categories: under and over 60 years. The results show that the probability of using i-voting decreases in the higher age group. [38] first undertakes a bivariate analysis through a chi-squared test to test the significance of the recognised differences. This shows that the 36-55 age group (68 percent), followed by the 56-65 age group (55 percent), then the 18-35 age group (44 percent) exhibit the highest percentage of i-voters. It is therefore by far not the youngest who use i-voting most. The chi-squared test and the results from the multivariate analyses that were carried out subsequently, produce significant results. The multivariate model from [30] shows a negative correlation between age and i-voting (compared to ballot box voting). [31] created a two-part model, comparing i-voters with postal voters. A negative coefficient is produced for internet and postal voting methods in the first part of the model and solely for internet in the second part of the model. I-voters therefore tend to be younger than postal or ballot voters.

4.2. Gender

All studies except [31] examine gender but only [30] finds significant results. According to [30] men exhibit a higher probability of voting online. [3] finds no statistically significant values. [38] is the same, although bivariate results show that only 42 percent of women, but 51 percent of men who took part in the study use i-voting. However the chi-squared test produces an insignificant result and the multivariate analysis therefore does not include this variable.

Table 3. Overview of studies examining socio-demographic characteristics of i-voters in the United States of America.

Study	Results
Alvarez and Nagler (2001) [3]	Based on _multivariate_ analysis, i-voters are **young**. **Gender** has no influence.
Solop (2001) [38]	Based on _bivariate_ analysis (chi-squared tests) i-voters are **young** (though not the youngest), **male** and **well-educated** with **high incomes**. Based on _multivariate_ analysis, **young**, **well-educated** voters are overrepresented among i-voters. **Income** is not significant.
Prevost and Schaffner (2008) [31]	Based on _multivariate_ selection model analysis, i-voters are **young** and **well-educated** with **high incomes**. Based on _multivariate_ outcome model analysis, **young** voters are overrepresented among i-voters. **Income** and **education** had minimal effect.
Prevost (2008) [30]	Based on a two-part _multivariate_ analysis, **young**, **male** voters with **high incomes** are overrepresented among i-voters. **Education** is not significant.

4.3. Income

Three of the four studies examined income. According to the bivariate analysis [38], in which the chi-squared test also produces significant results, high earners are overrepresented among i-voters. While i-voters make up only 21 percent of voters with the lowest income, the percentage is 69 percent for the top earners. The multivariate logistical regression, that includes age, income and education variables, does not confirm this result however. Only age and education provide significant values. [30] also finds a positive correlation between income and i-voting compared to ballot voters who tend to earn less. [31] finds, in the first part of their model, that i-voters and postal voters have higher incomes than ballot voters. However, a comparison of internet and postal voters produces a negative coefficient for i-voting.

4.4. Education

Three of the four studies examine this variable. According to [30], education has no influence on the choice of voting method, while [31] produces a positive correlation. [38] also confirms the significance of this relationship. According to the bivariate study of [38], well-educated voters are overrepresented among i-voters. While lowest income earners make up 17 percent of i-voters, 69 percent come from the highest income earners. The chi-squared test confirms the significance of this difference. As previously mentioned, the multivariate analysis also produces a significant result. According to [31], a university degree increases the probability of voting via the internet or by post (compared with ballot voting). If postal voting and i-voting are compared with one another, a university degree increases the probability that voting will be online.

5. Canada

In the last decade, i-voting has been offered for local elections in approximately 60 municipalities in the provinces of Ontario and Nova Scotia [29]. Four studies examine the socio-economic composition of i-voters and refer to elections in 2003, 2006 and 2010 in Markham, Ontario [12, 13, 14], [19]. Table 4 gives an overview of these studies. The study analyses are bivariate and limited to providing a description of the distribution of the different variables. The statistical significance of the differences found in the distribution is not provided. Only [12] compares ballot voters with i-voters; the other studies focus exclusively on i-voters.

5.1. Age

Age is examined in all the studies. [12], which provides a comparison between different voting channels, shows that i-voters are slightly younger that ballot voters. The percentage of 18-24, 35-44 and 45-54 year olds is about three percentage points higher among i-voters than traditional voters. In contrast, 19 percent of ballot voters and only 8 percent of i-voters are over 65. Since none of the other studies make comparisons to traditional voters, the results are less meaningful for this report. The general tone, however, is that middle-aged voters use i-voting most frequently.

5.2. Gender

Gender is examined in two of the four studies. [12] observes no significant gender bias when it comes to voting method. 51 percent of i-voters and 52 percent of ballot voters are male. [13] likewise concludes that of the sample voters for the 2006 election, approximately the same number of male (52 percent) and female (48 percent) were i-voters.

Table 4. Overview of studies examining socio-demographic characteristics of i-voters in Canada.

Study	Results
Delvinia Report (2004) [12]	I-voters are **young** and **well-educated** with **high incomes**. No significant **gender** bias amongst i-voters.
Delvinia Report (2007) [13]	**Young**, **well-educated** voters with **high incomes** are overrepresented among i-voters. No significant **gender** bias amongst i-voters.
Delvinia Report (2011) [14]	**Middle-aged**, **well-educated**, **high earning** voters are overrepresented among i-voters. **Gender** not significant.
Goodman (2014) [19]	**Middle-aged** and **older** voters are overrepresented among i-voters. No information regarding significance of other variables provided.

5.3. Income

All three Delvinia Reports deal with income. [13] regards CAD$99,300 as the average yearly income of i-voters. [14] concludes that i-voters typically earn between CAD$55,000 and CAD$84,999 per annum. Only [13] makes a comparison with traditional voters. It concludes that high earners are overrepresented amongst i-voters when compared with traditional voters. The percentage of i-voters is consistently higher than ballot voters in the income categories above CAD$85,000.

5.4. Education

Education is considered in [12] and [13]. In a comparison between i-voters and ballot voters in [12], it is observed that i-voters are slightly better educated than traditional voters. 42 percent of i-voters and only 36 percent of ballots voters have a university degree. 7 percent of ballot voters and 4 percent of i-voters are in the lowest education category. [13] only examines i-voters and finds that the distribution of the education variable is very similar to [12] and there is therefore a high probability that in this case i-voters are also slightly better educated than traditional voters. However, this is an assumption based on a comparison of the two studies and cannot be conclusively confirmed due to a lack of data regarding traditional voters from 2006.

6. Norway

In Norway i-voting trials were conducted during local elections in 10 municipalities in 2011 and during parliamentary elections in 12 municipalities in 2013. Recently the decision was made not to pursue i-voting [4]. For both studies, data on the socio-economic composition of the electorate was compiled and analysed [5] and [7]. The studies examine electoral roll and survey data. Table 5 gives an overview of these studies. Although multivariate analyses are carried out in the studies, their results have

been excluded from this report because they only compare i-voters to non-voters and ballot voters to non-voters, and make no direct comparison between the two voting methods. The following analysis is therefore limited to bivariate results. It must be noted that the bivariate studies not only compared i-voters with traditional voters, but also often included data from non-voters.

6.1. Age

In relation to age, one can say that i-voters are more evenly distributed across the age groups than ballot voters. The percentage of i-voters is relatively constant up to age 60 and then decreases markedly. In contrast, the percentage of traditional voters increases steadily until age 75 before falling away. A comparison of both voting methods determines that i-voters are most strongly represented among the younger age groups. According to [7] the 21-34 year old group has the same percentage of i-voters and ballot voters. In all the other age categories the percentage of ballot voters exceeds that of i-voters. Likewise in [5], the difference between i-voters and ballot voters is smallest among the youngest voters.

6.2. Gender

Both studies conclude that men are slightly overrepresented among i-voters. [5] shows that the percentage of voters who vote online is 26 percent on average; 25 percent female and 28 percent male. Whether this difference can be seen as statistically significant is not tested.

Table 5. Overview of studies examining socio-demographic characteristics of i-voters in Norway.

Study	Results
Bergh and Christensen (2012) [5]	Based on *bivariate* analysis, **younger**, **male**, **well-educated**, **high earning** voters are overrepresented among i-voters. The *multivariate* analysis delivers the same results.
Bock Segaard et al. (2014) [7]	Based on *bivariate* analysis, i-voters are **younger**, **male** and **well-educated** with **high incomes**.

6.3. Income

According to [5], the percentage of i-voters and ballot voters increases as income increases. Compared with non-voters, voters generally have higher incomes. [7] concludes that the percentage of ballot voters up to the third income category (of a total six) increases and thereafter decreases slightly. With i-voters, the percentage decreases from the first to the second income category and then steadily increases. The percentage of non-voters steadily decreases as income increases.

6.4. Education

The results in [5] for education are very similar as for income. While the percentage of voters for both voting methods (ballot and internet) increases as the education level improves, it decreases for non-voters. [7] confirms this, though in the analysis the

increase in the percentage of i-voters is much more prominent than for ballot voters. One can conclude that the well-educated are overrepresented among i-voters.

7. United Kingdom

Initially it seemed that the United Kingdom would be one of the pioneers in the implementation of i-voting. For a number of years pilot projects were carried out at the local level. After a time however, i-voting was no longer pursued [21]. Only one study was carried out that examines the socio-economic composition of i-voters. It relates to local elections in Swindon and in a bivariate analysis compares the profile of ballot voters and i-voters [21]. Of the four variables of interest in this report, [21] only examines age and gender. Table 6 gives an overview of the study.

Table 6. Overview of studies examining socio-demographic characteristics of i-voters in the United Kingdom.

Study	Results
Henry (2003) [21]	Based on *bivariate* analysis i-voters are **young**, **male** and belong to the **upper social classes**.

7.1. Age

According to [21] i-voters are younger than ballot voters. While the over 60 years category has the highest number of ballot voters (42.6 percent), it only accounts for 2.6 percent of i-voters. 35.9 percent of i-voters belong to the 18-24 years category and 48.7 percent (the highest) to the 25-44 years category. Other than for the over 60 years category, the study does not provide any comparative data concerning ballot voters.

7.2. Gender

[21] finds that 43.3 percent of i-voters and 52.3 percent of ballot voters are female. Whether this difference is statistically significant is not tested in the study.

8. Conclusion

The common assumption of the typical i-voter being relatively young, male, well-educated and having a higher than average income can be partly confirmed by the studies reviewed in this article. Overall, the *age* variable is the most significant one. All bivariate analyses and multivariate socio-economic models as well as a high proportion of more global multivariate analyses show the importance of age as an explanatory factor for the selected voting channel. In most studies it is the second youngest age group rather than the youngest that proves to be the most frequent user of i-voting. Certainly, the oldest age groups make use of i-voting the least. As a preliminary first result, after i-voting trials in several countries in the last fifteen years, we should therefore refrain from raising the expectation that the i-voting channel will bring the young, underrepresented voters back to the ballots.

The socio-demographic variable with the least clear effect is *gender*. Although men are overrepresented among i-voters in most cases, the differences are usually not statistically significant. However, it should be noted that in none of the presented studies are women choosing this voting channel more often than men. Regarding *income* and *education* we can observe certain similarities in their effect on the choice of the voting channel. Most bivariate and multivariate socio-demographic models show that voters with a higher education and a high income tend to use i-voting more often. Global models integrating variables other than just socio-economic ones still tend to deliver significant effects for income.

In general, the socio-economic analyses show that the significance of the four variables focussed on tends to decrease as soon as the models include the variables of computer literacy and trust in the internet. The strongest moderating effect can be observed for gender and education. As long as variables outside of the socio-economic ambit are not included, most analysis would report that a higher income and a good education would very likely lead to a higher propensity to use i-voting. According to the more encompassing models, the real cause for the divide between i-voters and traditional voters is their affinity to the internet. With time, i-voters will get older and the age gap will disappear. We should therefore consider that socio-economic factors such as age, gender, income and education are only relevant to a certain degree when explaining the choice to vote online and are rather strongly moderated by knowledge and use of the internet.

References

[1]   M.R. Alvarez and T.E. Hall, *Point, Click, and Vote: The Future of Internet Voting*, Brookings Institution Press, Washington DC, 2004.

[2]   M.R. Alvarez, T.E. Hall and A.H. Trechsel, Internet Voting in Comparative Perspective: The Case of Estonia, *PS: Political Science and Politics* **42(3)** (2009), 497-505.

[3]   M.R. Alvarez and J. Nagler, *The Likely Consequences of Internet Voting for Political Representation*, Loyola Law School, Los Angeles, California, 2001.

[4]   BBC, E-Voting Experiments End in Norway amid Security Fears. http://www.bbc.com/news/technology-28055678

[5]   J. Bergh and D.A. Christensen, Valgdeltakelse og bruk av internettstemmegivning. Har hvordan velgerne stemmer betydning for om de stemmer?, in S. Bock Segaard and J. Saglie (eds.), *Evaluering av forsøket med e-valg 2011*, Institutt dor Samfunnsforskning, Oslo, 2012, 29-157.

[6]   D. Bochsler, *Can Internet Voting Increase Political Participation? Remote Electronic Voting and Turnout in the Estonian 2007 Parliamentary Elections*, Presentation at Internet and Voting Conference Fiesole, 2010. http://www.eui.eu/Projects/EUDO-PublicOpinion/Documents/bochslere-voteeui2010.pdf

[7]   S. Bock Segaard, D.A Christensen, B. Folkestad and J. Saglie, *Internettvalg: Hav gjør og mener velgerne?* Institutt for Samfunnsforskning, Oslo, 2014.

[8]   F. Breuer and A.H. Trechsel, *E-voting in the 2005 Local Elections in Estonia*. Report for the Council of Europe, Robert Schuman Centre for Advanced Studies, European University Institute, Florence, 2006.

[9]   Bundeskanzlei, Bericht des Bundesrates zu Vote électronique. Auswertung der Einführung von Vote électronique (2006-2012) und Grundlagen zur Weiterentwicklung, 2013. http://www.bk.admin.ch/themen/pore/evoting/07977/index.html?lang=en

[10] T. Christin and A.H. Trechsel, *Qui vote par internet? - Une approche scientifique des scrutins de Carouge et Meyrin*, Chancellerie d'Etat, Geneva, 2004. http://ge.ch/vote-electronique/media/site_vote-electronique/files/imce/doc_rapports/200409_rapport_carouge_meyrin.pdf

[11] T. Christin and A.H. Trechsel, *Analyse du scrutin du 26 septembre 2004 dans quatre communes genevoises (Anières, Carouge, Cologny et Meyrin)*, E-Democracy Center, University of Geneva, 2005. http://ge.ch/vote-electronique/media/site_vote-electronique/files/imce/doc_rapports/rapport_26sept_final.pdf

[12] Delvinia Report: Internet Voting and Canadian E-Democracy in Practice. The Delvinia Report on Internet Voting in the 2003 Town of Markham Municipal Election, Delvinia Interactive, Toronto, 2004. http://www.delvinia.com/egov/Delvinia_Voting_Report_04.pdf

[13] Delvinia Report: Understanding the Digital Voter Experience. The Delvinia Report on Internet Voting in the 2006 Town of Markham Municipal Election, Delvinia Interactive, Toronto, 2007. http://www.delvinia.com/egov/Delvinia_Voting_Report_07_full2.pdf

[14] Delvinia Report: eDemocracy and Citizen Engagement. The Delvinia Report on Internet Voting in Town of Markham, Delvinia Interactive, Toronto, 2011. http://de.slideshare.net/delviniainteractive/delvinia-dig-issue3internetvotingreport201109

[15] A. Driza-Maurer, O. Spycher, G. Taglioni and A. Weber, *E-voting for Swiss Abroad: a joint project between the Confederation and the cantons*, in M.J. Kripp, M. Volkamer and R. Grimm (eds.), Electronic Voting 2012, Proceedings of the 5th conference on electronic voting (EVOTE2012), Gesellschaft für Informatik, Bonn, 2012.

[16] Federal Chancellery, Bericht über die Pilotprojekte zum Vote électronique, 2006. http://www.admin.ch/opc/de/federal-gazette/2006/5459.pdf

[17] G. Gerlach and U. Gasser, Three case studies from Switzerland: e-voting, *Berkam Center Research Publication* **3(1)** 2009, 1-17.

[18] M. Germann, F. Conradin, C. Wellig and U. Serdült, *Five Years of Internet Voting for Swiss Expatriates*, in P. Parycek and N. Edelmann (eds.), CeDEM 14. Conference for E-Democracy and Open Government, Danube University Krems, Krems, 2014, 127-140.

[19] N. Goodman, Internet Voting in a Local Election in Canada, in B. Grofman, A.H. Trechsel and M. Franklin (eds.), *The Internet and Democracy in Global Perspective. Voters, Candidates, Parties, and Social Movements*, Springer, Switzerland, 2014, 7-24.

[20] K. Grönlund, Cyber Citizens: Mapping Internet Access and Digital Divides in Western Europe, in N. Kersting and H. Baldersheim (eds.), *Electronic Voting and Democracy: A Comparative Analysis*, Palgrave Macmillan, Basingstoke, 2004, 20-38.

[21] S. Henry, Can Remote Internet Voting Increase Turnout? Aslib Proceedings, *Journal of Information Management* **55(4)** (2003), 193-202.

[22] N. Kersting and H. Baldersheim, Electronic Voting and Democratic Issues: An Introduction, in N. Kersting and H. Baldersheim (eds.), *Electronic Voting and Democracy: A Comparative Analysis*, Palgrave Macmillan, Basingstoke, 2004, 3-19.

[23] W. Linder, E-Voting - eine Belebung der direkten Demokratie? *LeGes* **1** (2003), 103-124.

[24] S. M. Lipset and S. Rokkan, Cleavage Structures, Party Systems, and Voter Alignments: An Introduction in S.M. Lipset and S. Rokkan (eds.), *Party Systems and Voter Alignments: Cross-National Perspectives*, The Free Press, Toronto, 1967.

[25] S. Lüchinger, M. Rosinger and A. Stutzer, The impact of postal voting on participation: Evidence from Switzerland, *Swiss Political Science Review* **13(2)** (2007), 167-202.

[26] F. Mendez and U. Serdült, From Initial Idea to Piecemeal Implementation: Switzerland's First Decade of Internet Voting Reviewed, in D. Zissis and D. Lekkas (eds.), *Design, Development and Use of Secure Electronic Voting Systems*, IGI Global, Hershey, Pennsylvania, 2014, 115-127.

[27] F. Mendez, U. Serdült and J. Wheatley, *The growth of mechanisms of e-governance and e-democracy in Europe: an overview*, Smart City: Social Participation and Urban Governance International Conference, Shanghai, China, 2014.

[28] P. Norris, *The Worldwide Digital Divide: Information Poverty, the Internet and Development*, Annual Meeting of the Political Science Association of the UK, 2000. http://www.hks.harvard.edu/fs/pnorris/Acrobat/psa2000dig.pdf

[29] J.H. Pammett an N. Goodman, *Consultation and Evaluation Practices in the Implementation of Internet Voting in Canada and Europe*, Elections Canada, 2013. http://www.elections.ca/res/rec/tech/consult/pdf/consult_e.pdf

[30] A. Prevost, *Assessing Internet Voting as an Early Voting Reform in the United States*, in R. Krimmer (ed.), Electronic Voting (EVOTE 2008), 3rd International Conference, Gesellschaft für Informatik, Bonn, 2008, 63-80.

[31] A. Prevost and B. Schaffner, Digital Divide or Just Another Absentee Ballot? Evaluating Internet Voting in the 2004 Michigan Democratic Primary, *American Politics Research* **36(4)** (2008), 510-29.

[32] D. Schloeth, Vor die Wahl gestellt. Erklärungen des Wahlverhaltens bei den Eidgenössischen Wahlen 1995, *Selects - Swiss Electoral Studies* **3**, Verlag Paul Haupt, Bern, Stuttgart, Vienna, 1998.

[33] P. Sciarini, F. Cappelletti, A. Goldberg, A. Nai and A. Tawfik, *Étude du vote par internet dans le canton de Genève*, Rapport final à l'intention de la Commission externe d'évaluation des politiques publiques, Département de science politique et relations internationales, University of Geneva, 2013. http://ge.ch/vote-electronique/media/site_vote-electronique/files/imce/doc_rapports/rapport-final-sciarini-vote-par-internet.pdf

[34] U. Serdült, Internet voting for the Swiss abroad of Geneva: First only survey results, in J-L. Chappelet (ed.), *Electronic government and electronic participation: joint proceedings of ongoing research and projects of IFIP EGOV and ePart 2010*, Trauner Verlag, Linz, 2010, 319-325.

[35] U. Serdült, Referendums in Switzerland, in M. Qvortrup (ed.), *Referendums around the world: the continued growth of direct democracy*, Palgrave Macmillan, Basingstoke, 2014, 65-121.

[36] U. Serdült, M. Germann, F. Mendez, A. Portenier and C. Wellig, *Fifteen Years of Internet Voting in Switzerland: History, Governance and Use*, in L. Terán and A. Meier (eds.), ICEDEG 2015: Second International Conference on eDemocracy & eGovernment, Quito, Ecuador, 8-10 April 2015, IEEE Xplore CFP1527Y-PRT, 149-156.

[37] U. Serdült and A.H. Trechsel, *Umfrage bei Stimmberechtigten der Zürcher Gemeinden Bertschikon, Bülach und Schlieren anlässlich des Pilotversuchs zum Vote électronique vom 27. November 2005*, Bundeskanzlei, Sektion Politische Rechte, Bern, 2006. http://www.edemocracycentre.ch/files/Umfrage%20bei%20Stimmberechtigten.pdf

[38] F.I. Solop, Digital Democracy Comes of Age: Internet Voting and the 2000 Arizona Democratic Primary Election, *Political Science and Politics* **34(2)** (2001), 289-93.

[39] A.H. Trechsel, G. Schwerdt, F. Breuer, M.R. Alvarez and T.E. Hall, *Internet Voting in the March 2007 Parliamentary Elections in Estonia*, Robert Schuman Centre for Advanced Studies, European University Institute, Florence, 2007.

[40] A.H. Trechsel and K.Vassil, *Internet Voting in Estonia. A Comparative Analysis of Five Elections since 2005*, Report for the Estonian National Electoral Committee, Robert Schuman Centre for Advanced Studies, European University Institute, Florence, 2011.

Instrument Categories
for Open Societal Innovation

Jörn von Lucke[a,1] and Katharina Große[b]
[ab]*Zeppelin University*

Abstract. In this paper, we introduce a new categorisation for innovation tools, which we employed to structure the Toolbox for Open Societal Innovation (TosiT) – a database of innovation software and services, which was developed in the international research project eSociety Bodensee 2020. The suggested categories are designed with the user in mind to make the toolbox easily accessible for interested actors. Open Societal Innovation (OSI) is a concept that combines the notions of open innovation and social innovation, in order to foster innovation from the society for the society.

Keywords. open innovation, open societal innovation, open government, participation, collaboration, co-operation

1. eSociety Bodensee 2020[2]

In the past years, nearly every aspect of our lives has undergone a radical change. New technologies transform the way we consume, travel, and interact. Our homes become connected, traffic and power grids are becoming responsive, and governments are opening to citizens, making their information transparent and engaging them in their problem solution. With the inter-connectedness of smart cities and smart regions, the problems we face will equally increase in complexity. In order to tackle this challenge, it will be crucial to establish a culture of open societal innovation in order to fully profit from the innovative potential of politics, administration, civil society, and businesses. How exactly can this collaboration be initiated and sustained? Which tools are suitable and available and how can they be introduced into every-day problem solving?

To answer these questions, the Internationale Bodensee Hochschule (IBH) funded the research project eSociety Bodensee 2020. Bridging the Lake Constance's region short distances between countries, scientists from Zeppelin University in Germany, the University of Liechtenstein and the Swiss University of Applied Sciences in St. Gallen examined how to establish and foster an innovation culture in the region.

The project developed the idea of open societal innovation (OSI), which we will introduce below. It also gathered best-practice examples of OSI in the Lake Con-

---

[1] joern.vonlucke@zu.de
[2] The contents of this paper were carefully adapted and enhanced from the published project handbooks, which so far are available in German only [20]

stance region and initiated pilot projects in the participating countries, which spanned from innovation in culture to social and political innovation.

Additionally, it created "The Open Societal Innovation Toolbox" (TosiT: http://www.tosit.org) – a database of innovation tools, which is meant to assist administrators, citizens, and project managers to find tools or services to support their OSI projects. The toolbox is freely available and can be expanded by the crowd.

On the subsequent pages, we will shortly introduce and explain OSI, in order to lay the basis for this paper's main focus: the instrument categories employed to structure TosiT. These categories were developed with the project and the toolbox's goal in mind: to foster open societal innovation. We hope TosiT's instrument categories will not only advance OSI around Lake Constance and enable other regions to follow suit. We also think they will add significantly to the body of innovation literature, especially in the context of state and society, as the suggested categories move away from a technology-centric towards a more user-centric point of view, developed to facilitate the use of innovation instruments for OSI actors.

2. Open Societal Innovation

"Open societal innovation (OSI) refers to the adaptation and subsequent sustainable use of appropriate open innovation approaches from business, adapted and utilized by state and society to solve societal challenges" [1]. On the subsequent pages, we will explain in more detail the concept of open societal innovation. In order to do so, we will firstly introduce the ideas of innovation, including social innovation, and open innovation, which all contribute to the concept of OSI.

2.1. Innovation and Social Innovation

Innovation has become an essential topic in business management and product development. Approaches for definitions vary from a resource-centric point [2] to a focus on degree of newness [3]. For this paper, the most important take away from the discussion about innovation in general is that it does not only concern products, but is equally including changing and re-structuring processes. While this might at times be disruptive, it can also be an incremental change or an ever-so-slight improvement.

However, there is also another, non-business side to innovation. Mulgan [4] describes „innovative activities and services that are motivated by the goal of meeting a social need and that are predominantly developed and diffused through organizations whose primary purposes are social" (p. 146). He calls this social innovation. While this definition implies that social innovation can only be brought about by social organization, mostly, the focus is less narrow and centres on the objective of the innovation: Social innovation can be interpreted as innovation with a focus on improving society, e.g. by tackling environmental problems, improving work situations and education or changing social systems to the better [5] [6] [7]. The World Economic Forum understands social innovation as improving the situation of under-privileged groups of society [8].

From this, we can take that social topics can be at the centre of innovation and a diverse array of actors may be involved, exceeding the corporate domain.

2.2. Open Innovation and Open Social Innovation

Companies have started years ago to open their innovation processes for external actors. Companies such as Audi, Adidas, BMW or Procter & Gamble [9] specifically invite customers and suppliers to contribute ideas to the improvement of products and processes. The discussion of open innovation has gained momentum since Henry W. Chesbrough coined the term in 2006 [10]. Additionally, many new possibilities of opening innovation processes arise due to the advances of ICT and the Internet. They enable asynchronous, location-independent communication and the involvement of a large number of people, or the crowd (crowdsourcing, [11]).

Independent of the specific technique that is used, open innovation aims at increasing the variety of ideas and solutions by increasing the number of people involved. This is nicely illustrated by a recent campaign in Vienna[3] that uses the slogan: Vienna has 1.7 million brains, let's use them!

Not only corporations are starting to understand the potential of opening innovation, but so do administrations. The problems that states face nowadays are increasingly complex [12]. They can greatly profit from the support of external actors [13].

Hilgers and Ihl [14] discuss this in the light of the phrase citizensourcing. They try applying open innovation concepts from the business perspective to the public sector. Similarly, Collm and Schedler [15] examine crowd innovation in public administration and Herzberg [16] discusses state modernisation through open innovation. Chesbrough and Di Minin [17] examine open social innovation, which they define as "the application of […] open innovation strategies, along with innovations in the associated business model of the organization, to social challenges" (p. 3).

As can be seen, opening innovation processes is relevant for both, industry and state. However, we are not satisfied by keeping apart the different actors and foci of innovation. We think society can profit greatly from a culture of innovation that combines all topics and actors. This is why we will now introduce our concept of open **societal** innovation.

2.3. Open Societal Innovation

The concept of social innovation illustrates that open innovation is not for corporations only. Social issues can be tackled equally and it might very well be employed by state actors or non-state third-sector organisations. However, both aficionados of open innovation and open social innovation tend to confine their efforts to their respective domain.

The British Ordinance Survey offers an admirable counter-example with its Geovation Challenge. In regular intervals, the administration issues challenges that tackle diverse social problems. It publishes the respective geodata, organises workshops and thereby invites innovators to contribute their suggestions. This is paired with an initial project financiation for the winners, which is meant to foster long-term entrepreneurial engagement.

It is the opportunity for this kind of holistic innovation strategy that leads us to suggest the concept open societal innovation (OSI). OSI "refers to the adaptation and

---

[3] http://digitalcity.wien/die-kreativitaet-und-das-wissen-der-17-mio-gehirne-nutzen

sustainable use of open innovation approaches from business, adapted and utilized by the state and society to solve societal challenges" [1].

The key element in OSI is that it does not only open the state's innovation processes for external actors (outside-in). It also intends to carry innovation impulses to the outside (inside-out). The two approaches can also be combined (coupled). The term 'societal' refers to the fact that not only does society participate in the innovation process (open innovation), or it is the beneficiary of innovation society (social innovation), but these aspects are combined into an open innovation from the society for the society [1].

While ICTs are not a necessity in open societal innovation, they can act as facilitator and significantly simplify and enhance its processes. In order to aid interested organisations and individuals to develop OSI projects, we deem it a necessity to develop an overview over available ICTs and techniques, to structure and make accessible the cornucopia of offers. This is why the project eSociety Bodensee 2020 decided to make one of its main outputs a database of OSI tools, openly and freely available: The Open Societal Innovation Toolbox (TosiT): http://www.tosit.org.

3. Instrument Categories for OSI

In the subsequent paragraphs, we will introduce one of the categorisations that we used to structure TosiT. Typically instrument classes for innovation focus on the type of tools or service offers: Is it a community, a challenge, a market place [18]? We use categories that differ significantly from these standard classes for innovation tools. We deem this necessary, because – in our experience – project initiators do not approach OSI projects with a specific idea about what kind of tool they need. They start with an idea about what they want to achieve with a tool, e.g. gather ideas. A toolbox for open innovation needs to be user-friendly and easily-accessible. Its categories must thus reflect the mind-set of its users. It might seem unsettling at first that our categories are not necessarily on the same level. They mix input-orientation (data), output-orientation (problem identification, ideas, problem solution, design), focus (future challenges), process-orientation (innovation management), and network (social media). Additionally, they are not necessarily mutually exclusive, as tools might very well fit to several categories.[4] However, these categories exactly reflect the concerns that administrators or projects managers have when approaching a database of ICTs and techniques to use in their innovation processes.

Users might consult TosiT with an output-orientation: We want to publish our geo-data, where can we do this and how (data)? We want to know where our street lamps are broken and there are holes in the road (problem identification). We need to collect ideas for reducing the noise of the new highway for the local residents (ideas). We need help keeping our senior citizens' driveways snow free in winter (problem solution). We need a new logo (design). We want to develop a strategy for our town – what do we need to consider and where do we want to go for the next 10 years (future challenges)? We have this new open innovation process planned with our citizens, but

---

[4] This is why we use the term category instead of class.

we somehow need to manage and organise it – are there tools to help us (innovation management)? We want to use facebook/twitter/youtube to engage our citizens and make them share their ideas (social media).

In the subsequent paragraphs, we will explain each category in turn and give examples.

3.1. Data

We start our description with tools that aid in the provision of data – probably the most basic form of OSI. Open Data and the deduced transparency and information are the basis for innovation. Government data portals might be named here. Data.gov in the USA is one of the more prominent examples. We also include repositories that store information such as images or texts or software. This also encompassed cloud services such as Dropbox or Google Drive. Jointly-assembled maps like Open Street Map or Wheelmap might also be considered.

It is important to realise that most repositories or open data portals require technical skills from the participants. Still, the potential they offer is considerable and their reach is theoretically unlimited. Quality of data and the maintenance of individual rights need to be considered.

3.2. Problem Identification

The second category summarises tools to identify, report and collect problems, challenges, and complaints. Famous examples are Fix My Street in the UK, See Click Fix in the USA or Maerker Brandenburg in Germany through which citizens can notify administration about potholes, broken streetlamps or un-emptied garbage containers. The complaints might be mapped or documented with a picture.

Crowdsourcing problem identification has a great potential for steering a city's maintenance resources to where they are needed. Also, if done correctly, it can increase the transparency of and citizens' understanding for the underlying processes. In order to achieve this, it is crucial that a problem reporting system is accurately integrated into an administration's hierarchy and processes. Also, problems must at least be acknowledged and requests must be responded to timely.

3.3. Ideas

Gathering ideas from citizens is probably one of the most straight forward types of OSI. Tools in this category will aid in collaborative brainstorming. Examples are MeetingSphere or IdeaScale. Also included is software to help with developing idea competitions and innovation contests.[5]

Opening up the idea generation to develop plans for improvements enables the influx of many new impulses and knowledge from outside the core administration/organisation. The heterogeneity of actors might lead to innovative solutions and great creativity. However, the more complex an issue, the more resources need to be

---

[5] For these contests, there is an overlap with the category problem solution. This stems from the fact that innovation competitions might call for both: ideas and fully fledged problem solutions.

invested to draw high-qualitative ideas from the crowd. Also, initiators must be prepared for a multitude of comments and ideas and must be willing to clearly define and explain the project to avoid confusion, centre the ideas, and motivate participants. It might be important in this case, to make sure that the generated ideas can be rightfully used by the initiator to avoid copyright infringements.

3.4. Problem Solution

This category unifies all tools that help organising the actual implementation of problem solutions. They might set-up virtual marketplaces or volunteer portals to distribute task, allocate volunteers or establish relationships. We also consider platforms such as Innocentive.com to be part of this category, because they requires users to deliver solutions that can be promptly implemented, as opposed to the joint gathering of ideas that might cover a small aspect of a problem.

The tools in this category can help to solve problems that otherwise would go unsolved. This might range from open source programming to solving scientific problems or securing a new childcare service. However, relying on opening problem solution to the crowd bares the risk of uncertainty. One cannot guarantee that a problem will actually be solved. Copyright, insurance and liability might be critical issues.

3.5. Design

While tools for design innovation could also be considered problem solving tools, opening up design processes places specific challenges on tools, and often is very distinct in the eye of the initiator, which is why we concede it a separate category. Instruments might help with finding a new logo or campaign picture. Initiators can describe their ideas and request – designers will submit their proposals. A well-known platform is 99designs.com.

While initiators might profit greatly from a variety of professional designs offered for their projects, they might consider it challenging to communicate clearly their vision and requirements. They might also find it difficult to choose as a multitude of offers might be available. Again, it is crucial to look into copyright regulations.

3.6. Future Challenges

Tools in this category focus on the identification of trends or possible developments and the creation of appropriate strategies. This calls for more standardised methods that generate reliable information. Well-known approaches include the Delphi-method and government foresight processes. Newer developments in this category are automated algorithms that can be used to make predictions like GoogleTrends and Google Correlate.

Employing these tools may enable administrations and organisations to move from an incremental "muddling through" [19] towards a more strategic approach to planning. Depending on the exact tool and method employed, the results must be regarded with caution and checked for intentional manipulation or bias due to non-representative user groups.

3.7. Innovation Management

We now introduce a more process-based category, which comprises tools that will support initiators throughout an entire life cycle of OSI or – if requested – manage specific stages. Examples are Hyve and Atizo360.

The support offered by the tools and services in this category holds the advantage that initiators can count on well-designed and tested processes that are readily available. They do not need to deploy and educate their own resources. However, these services are mostly cost-intensive.

3.8. Social Media

While social media can be used to identify problems, gather ideas, suggest solutions and maybe even designs – or at least to support and generate attention for OSI initiatives and processes – they deserve their own category, as many initiators will approach TosiT with the idea of engaging people via the well-known networks such as facebook or twitter. It is noteworthy that there is a great variety of networks and platforms as well as apps that are specifically designed to manage engagement. That means, they will facilitate surveys or allow idea competitions, and therefore add significantly to the features of the basic network.

Conducting OSI in social media certainly brings the advantage of "going where the people are" and well-known interfaces and thus might engage people that will be harder to reach on foreign platforms. Issues of data protection and privacy, which are and probably will remain unresolved with powerful social media corporations, need to be taken into consideration.

4. Conclusion

In this paper, we have introduced and argued for the concept of open societal innovation (OSI), which combines the notions of open and open social innovation. It does not only open innovation processes for society or pins society down as the beneficiary of innovation. It comprises both as open innovation from the society for the society. OSI includes corporations taking part in the process. It allows for and fosters new business-models. We further present The Open Societal Innovation Toolbox (TosiT), a database of tools for OSI. We finally advocate a new categorisation for the included tools, which is user-centric and therefore offers a more intuitive access for OSI initiators. TosiT uses eight categories of tools: data, problem identification, ideas, problem solution, design, future challenges, innovation management, and social media. We are convinced that these categories will aid OSI initiators to use TosiT and find the right support for their OSI projects.

References

[1] J. von Lucke, J. Herzberg, U. Kluge, J. vom Brocke & H.-D. Zimmermann, *Open Societal Innovation - The Alemannic Definition*, 2012. [Online]. Available: http://esocietybodensee2020.files.wordpress.com/2012/12/ibh121219-alemannic-definition-en-v1.pdf.

[2] J. Schumpeter, *Theorie der wirtschaftlichen Entwicklung*, Duncker & Humblot, Berlin, 2006.

[3] J. Hauschildt & S. Salomo, *Innovationsmanagement, Vahlen*, München, 2011.

[4] G. Mulgan, The Process of Social Innovation, *Innovations, Technology, Governance, Globalization* 1, 2 (2006), 145-162.

[5] E. Pol & S. Ville, Social Innovation: Buzz Word or Enduring Term?, *Journal of Socio-Economics* 38, 9 (2009), 878-885.

[6] J. Howaldt & M. Schwarz, *„Soziale Innovation" im Fokus: Skizze eines gesellschaftstheoretisch inspirierten Forschungskonzepts*, Transcript, Bielefeld, 2010.

[7] J. Howaldt & H. Jacobsen, *Soziale Innovation: auf dem Weg zu einem postindustriellen Innovationsparadigma*, VS Verlag für Sozialwissenschaften, Wiesbaden, 2010.

[8] World Economic Forum, *Global Agenda*, 2012. [Online]. Available: http://www3.weforum.org/docs/GAC12/IssueBrief/IB_SocialInnovation.

[9] R. Reichwald & F. T. Piller, *Interaktive Wertschöpfung: Open Innovation, Individualisierung und neue Formen der Arbeitsteilung*, Gabler, Wiesbaden, 2006.

[10] H. W. Chesbrough, *Open Innovation - The New Imperative for Creating And Profiting from Technology*, Harvard Business Review Press, Boston MA, 2006.

[11] J. Howe, *Crowdsourcing: Why the power of the crowd is driving the future of business*, Crown Business, New York NY, 2008.

[12] U. Schliesky, Über Begriff und Notwendigkeit eines staatlichen Innovationsmanagements, *Staatliches Innovationsmanagement*, Lorenz-von-Stein-Institut für Verwaltungswissenschaften, Staatliches Innovationsmanagement , 2010, 1-18.

[13] J. von Lucke, *Entdeckung, Erkundung und Entwicklung 2.0: Open Government, Open Government Data und Open Budget 2.0*, epubli GmbH, Berlin, 2012.

[14] D. Hilgers & C. Ihl, Citizensourcing: Applying the Concept of Open Innovation to the Public Sector, *International Journal of Public Participation* 4, 1 (2010), 67-88.

[15] A. Collm & K. Schedler, Managing Crowd Innovation in Public Administration, *International Public Management Review* 13, 2 (2102), 1-18.

[16] J. Herzberg, *Staatsmodernisierung durch Open Innovation: Problemlage, Theoriebildung, Handlungsempfehlungen*, epubli GmbH, Berlin, 2012.

[17] H. W. Chesbrough & A. Di Minin, Open Social Innovation, *New Frontiers in Open Innovation*, Oxford University Press, Oxford, 2014, 1-25.

[18] M. Amberg, F. Bodendorf & K. M. Möslein, *Wertschöpfungsorientierte Wirtschaftsinformatk*, Springer, Berlin Heidelberg , 2011.

[19] C. E. Lindblom, The Science of "Muddling" Through, *Public Administration Review* 19, 2 (1959), 79-88.

[20] C. Raffl, J. v. Lucke, O. Müller, H.-D. Zimmermann & J. v. Brocke, *Handbuch für offene gesellschaftliche Innovation*, epubli GmbH, Berlin, 2015.

This page intentionally left blank

Policy Modelling

This page intentionally left blank

Simulating the Core Dynamics of a Social Dilemma. Individual Choices, Time and Sanctions in the Tragedy of the Commons

Nicola LETTIERI [a,b 1], Margherita VESTOSO [b]
[a] *ISFOL, Rome, Italy*
[b] *University of Sannio, Dept. of Law, Economics, Management, Quantitative Methods,*
Benevento, Italy

Abstract. The understanding of the way in which collective phenomena emerge from the interaction between individual behaviors, environment and institutions, can play a crucial role in supporting the design of more contextualized policies. An apparently effective policy can easily fail if policy makers do not consider the interplay between individual decision making and social aggregate dynamics. This paper presents an ongoing research exploiting an agent-based simulation model to explore the core dynamics of the Tragedy of the Commons (ToC), a social dilemma known for being behind a series of societal problems spanning from pollution to resource depletion and climate change. The goal is twofold: capture the basic processes through which the ToC emerges and evolves; explore in an artificial society the effects of different strategies aiming to contrast the phenomenon. Our attention is focused on the interplay between different factors proven to be involved in the genesis of this dilemma: the selfish rationality of human beings, the temporal dimension of individual choices and the potential impact of sanctions.

Keywords. policy modeling, computational social science, Tragedy of the commons, agent-based models, evolutionary simulations, intertemporal choice

Introduction

The understanding of the way in which collective phenomena emerge from the interaction between individual behaviors, environment and institutions, can play a crucial role in supporting the design of more effective policies. A relevant limit of public policies is that they often fail in evaluating the effects that a given strategy could produce on the society. Part of this difficulty is due to fundamental features of social complexity: social systems are characterized by multiple ontological levels with multiple connections, proceeding not only from the micro to the macroscopic levels but also back from the macro to the micro levels [11], [26].

Recently, the complexity science perspective has fostered innovative approaches to policy making aiming at importing exact science methods and tools, advanced computing techniques and complexity mathematics into socio-economic policy (e.g. [2], [16]) and into rule making [36]. Interesting works have been published in this field [9]

---

[1] Corresponding author: Nicola Lettieri, ISFOL, C.so d'Italia 33, 00198, Rome, Italy; E-mail: n.lettieri@isfol.it.

showing how the intersection of complexity sciences with public policy and management is leading to new insights into very perplexing policy issues allowing the design of more successful policies "conceived not as something that takes place off-line, outside systems, but as a constitutive process interacting with self-organized system behavior" [50].

A fundamental contribution can be given, in this scenario, by computer simulations models as they provide new insights into social dynamics that can inform the design of effective policy solutions. Policy models are often unable to grasp and forecast complex social processes including the reaction of individuals to policy decisions, the aggregate effect of their interactions, and their consequences on large spatial–temporal scales [39], [50]. As frequently highlighted in recent policy informatics literature [14], [32], [33], simulation models can be in various way helpful in addressing this issue: they can not only offer new ways to predict how social systems co-evolve with the policy, but also a deeper understanding of the mechanics underlying the issues policy makers want to address. More in general, they promote a scientific habit of mind in policy making, pushing towards the adoption of explicit objectives expressed in quantitative and verifiable way, enabling testable predictions about the impact of the choices and of the actions undertaken. By revealing tradeoffs, uncertainties, and sensitivities, simulations can discipline the public dialogue about options making judgments more considered. Finally computer simulation models offer a chance to overcome the drawbacks of a disciplinary approach to policy design. Individual and collective phenomena that affect policy effectiveness are investigated by a number of distinct disciplines from sociology to psychology, law, or economics and simulations help us in understanding how the phenomena studied by these disciplines work together and influence each other.

According to the vision so far described, this paper presents an ongoing research exploiting an evolutionary agent-based simulation model to explore the core dynamics of the Tragedy of the commons (hereinafter ToC), the social dilemma theorized as being behind a series of societal problems spanning from pollution to resource depletion and climate change. The research goal is twofold: identify the core dynamics through which the ToC emerges and evolves and explore the effects of different strategies aiming to contrast the phenomenon. Our attention is focused, in particular, on the interplay between different factors proven to be involved in the genesis of this dilemma: the selfish rationality of human beings, the temporal dimension of individual choices and the potential impact of sanctions.

The paper is structured as follows: in Section 1 we present the ToC focusing attention on two of the most discussed factors involved in generating this social dilemma. Section 2 offers a general overview of the agent-based model we are developing and testing. Section 3 sketches preliminary results of the experiments so far conducted. Section 4 draws some conclusions and discusses future developments.

1. A puzzling dilemma for policy makers: the Tragedy of the Commons

The complexity of public policies is often connected with the need to cope with social dilemmas, situations that "*arise whenever a group of individuals must decide how to share a common resource while balancing short-term self-interests against long-term group interests*" [22]. In general terms, social dilemmas are characterized by two fundamental features: a) each individual of the community taken into account receives

a higher payoff for a socially defecting choice (e.g. using all the available energy, polluting neighbors) than for a socially cooperative choice, no matter what the other individuals in society do, but b) all individuals are better off if all cooperate than if all defect.

Social dilemmas are extremely relevant because they often create or lead to social issues, problems, or even disasters. Given the ubiquity and the global importance of some of them, it is essential for policy maker to learn how to deal with them: since social dilemmas become more complex there is an increasing need to understand how to design effective policies. That's why the scientific interest in social dilemmas - particularly those resulting from overpopulation, resource depletion, and pollution - has grown dramatically in the last ten years. The attention is shifting from pure laboratory research towards interdisciplinary approaches characterized by the cooperation between different research areas (spanning from computer science to anthropology, from biology, economics, neuroscience, to political science and psychology) aiming to develop together a unifying theoretical framework.

The ToC is a well-known social dilemma described for the first time by Garret Hardin in 1968 [23] and deriving by the fact that individuals acting autonomously and rationally according to their self-interest behave contrary to the interest of the whole group depleting common resources. Over the time, ToC has been considered as the prototype of a range of dilemmatic situations that occur in different social contexts from micro (local) to macro (global) level. The considerations made about the Tragedy can be therefore extended virtually to any instance in which society appeals to an individual exploiting a common resource to restrain himself for the general good by means of his conscience. ToC is often mentioned in connection with different issues from economic growth to environmental protection. It has been used in analyzing behavior in the fields of economics, evolutionary psychology, game theory, politics and sociology (extremely known the analysis conducted by Elinor Ostrom [40], [41]).

The evolution of the ToC has been traced back to many different social and individual causes. It is possible anyway to identify two basic factors involved in triggering the ToC, the two factors we focused on in designing of our simulation model.

a) Selfish rationality

In his paper, Hardin identifies in the individual tendency to increase well-being, the core mechanism generating the Tragedy. "A*s a rational being",* he states, *"each herdsman seeks to maximize his gain. Explicitly or implicitly, more or less consciously, he asks, "What is the utility to me of adding one more animal to my herd?" This utility has one negative and one positive component. 1) The positive component is a function of the increment of one animal. Since the herdsman receives all the proceeds from the sale of the additional animal, the positive utility is nearly +1. 2) The negative component is a function of the additional overgrazing created by one more animal. Since, however, the effects of overgrazing are shared by all the herdsmen, the negative utility for any particular decision-making herdsman is only a fraction of -1."* In this context, adding together the component partial utilities, a rational individual concludes that the best choice for him is to continue to add animals to his herd. But, as this is the conclusion reached by each rational herdsman sharing a commons, the tragedy becomes unavoidable.

b) Intertemporal choices and temporal discount

The selfish rationality is not the only ingredient of the ToC. As highlighted in psychological literature [12], the Tragedy is strictly connected, like many social dilemmas, with the "time lag": the fact that behaviors resulting in immediate reward

leads to long-term negative effects (see Platt's concept of "social traps" [43]). A recent study [30] focusing on climate change shows that a crucial role is played, in the genesis of the ToC, by the way in which human beings deal with time while making what are called "intertemporal choices". Intertemporal choices concern options that can be obtained at different points in time (e.g. buying expensive cars today or saving the money to ensure a sizable pension in the future) and that often oppose a smaller but sooner prize (e.g., a modest amount of food ready at hand) against a larger but delayed outcome (e.g., a more distant but also richer foraging opportunity). The research, published in 2013, aims to test how groups of individuals respond to the challenge of avoiding dangerous climate change in a setting that rewards defection immediately and rewards cooperation over three different time horizons to represent intra- and intergenerational discounting. According to the results of the public-good experiments presented in the paper, the difficulty of avoiding dangerous climate change (but the analysis can be extended to many other social dilemmas with the same structure and also to other instances of the Tragedy) arises not only from the tension between group and self-interest generated by rational selfish behaviors, but is also exacerbated by climate change's intergenerational nature. The dilemma lies in that *"present generation bears the costs of cooperation, whereas future generations accrue the benefits if present cooperation succeeds, or suffer if present cooperation fails"*.

The interplay between the two above mentioned factors makes it difficult to counteract the evolution of the Tragedy and suggests the introduction of incentives to cooperate, such as mutual coercion [23], punishment [17], [29] rewards [30] or even reputation [38] and shame [29]. A crucial issue for policy makers is therefore the comprehension of the way in which incentives can be administered in order to effectively fight the spread of the social dilemma.

2. An evolutionary model of the Tragedy of the Commons

As above highlighted, the research here presented aims to explore the core dynamics underlying the ToC and, at the same time, to investigate the way in which these dynamics can be altered by means of well tailored incentives to cooperate. To this end, drawing inspiration from a series of interesting research works on this topic [7], [13], [27], [31], [53]), we have implemented a simulation model using *NetLogo* [48]*,* an open source multi-agent agent-based programming language and integrated modeling environment widely used for simulation purposes in the social sciences. The model is available online at https://www.openabm.org/user/1294/models.

2.1. Theoretical framework

Taking cue from the considerations sketched in Section 1 about *selfish rationality* and *time discount*, we tried to build a comprehensive, multi-theoretical model in which facts are stylized according to different existing theories of human behavior. Without going into details, we can say that our choices in designing the model are mainly grounded in two theoretical frameworks:

- *Rational choice theory:* developed by the economist Gary Becker in 1976 [3] and still widely used to model human decision making in economics, sociology, and political science, rational choice theory revolves around the idea that people make

decisions about how they should act by comparing the costs and benefits of different courses of action.

    - *Intertemporal choice:* the scientific investigation of the intertemporal choices (as said above, the decisions with consequences that play out over time) has produced interesting findings in different research areas. After an initial emphasis on finding what mathematical model (typically a delay discounting function) would better fit the empirical data [34], [51], current research is focusing on exploring the cognitive mechanisms that produce the observed behavioral patterns, and tracing their evolutionary roots [4].

    Drawing upon insights offered by both rational choice theory and intertemporal choice research, we decided to explore the impact that the mechanism of sanction/punishment can have in preventing the evolution of the ToC. As widely acknowledged in a huge and heterogeneous literature, sanctions are essential in promoting social cooperation [17], [19] even in dilemmatic situations like the ToC [30]. A still challenging research goal, however, is to experimentally determine how different kinds of sanction can dynamically affect the evolution of social dilemmas taking into account the interplay between factors like selfish rationality and intertemporal choices and computer simulations have proven to be a suitable tool in this regard (see, for example, [35]).

    Our model is extremely simplified with respect to the phenomena under investigation while, as authoritatively claimed [10], policy modeling will more and more often need complex simulations to yield conclusions concretely applicable to real world problems. Even very abstract models, anyway, can be useful when they are able to grasp the core mechanics of the socio-economic dynamics impacted by policy measures. In this perspective, the simulation here presented is expected to help to figure out what can be done to increase cooperation in dilemmatic situations (how, to what extent and when apply sanctions) offering insights that can integrate more traditional regulatory impact assessment tools [52]. Policy solutions inspired by simulation experiments can account for both individual decision making and the counterintuitive effects generated by the interactions between the individuals, the environment and the policy measures. Obviously, when needed, it will be possible to increase the level of realism of the model developing scenarios more grounded in real data in which the actors, their behaviors and the environment are described with a higher level of detail. This will offer policy makers results with a higher predictive power and, therefore, more immediately exploitable in real settings.

2.2. Modeling approach

Every simulation is based on a more or less abstract model that simplify the representation of a target phenomenon. The design of the model and the way in which facts are stylized inside the model, depends on the research objectives, the scientific approach and the researcher's point of view about the causes of the phenomenon under investigation.

    Different approaches to the simulation of social phenomena have emerged so far from system dynamics [18], [44], to microsimulations [49] and agent-based simulations [11], [15], [20]. In particular, the agent-based paradigm (ABM) is a specific kind of social simulation that can be defined as a ''*computational method that enables a researcher to create, analyze, and experiment with models composed of agents that*

interact within an environment''[20]. Based on the identification of the scientific explanation with the reproduction ''*in silico*'' (i.e. in a computer simulation) of the social processes being investigated, ABM has contributed to promote a generative [15] approach to social science research: social macro dynamics and structures are interpreted, described, reproduced and explained as the result of micro-interactions between computational entities (agents) simulating the behavior of real individuals. In this perspective, modeling the structural properties of social systems and exploring their spatio-temporal development via computer simulation are crucial steps to provide explanations of complex social outcomes.

Over time, ABM research has generated different approaches to modeling phenomena that are very close to those discussed in this paper such as the effects of punishment on cooperation. These approaches, roughly speaking, can be classified into two macro-categories. On the one hand, there are simulation models whose main objective is the analysis of the socio-cognitive underpinnings of human behaviors. This category of models, in which agents are endowed with complex cognitive architectures reproducing mental processes - *Beliefs, Desires, Intentions* (BDI) [45] and *Beliefs, Desires, Obligations, Intentions* (BOID) [6]. On the other hand, there are *evolutionary models* which are somehow bio-inspired and are not so much interested in the internal dynamics of the agent but focus on the effects of mutual influences between individuals and the social environment and on analyzing the condition under which pre-programmed strategies can become stable patterns of behavior. So far, evolutionary simulation models have been frequently used to understand how social outcomes spanning from cooperation [1], [21], [25], to social learning [37], can be explained as the result of adaptation strategies. The study of punishment, in particular, has already exploited evolutionary models [5], [24].

From a technical point of view, the core of the evolutionary approach is represented by modeling and programming techniques trying to mimic natural processes of adaptation. One of the most relevant of these techniques is the genetic algorithm (GA) that imitates the evolutionary process of learning based on research and exploration [28]. Often used in social simulation research [8], [41], GA allows to model populations of adaptive agents that are not fully rational in the sense that they are only capable of refining the strategies adopted by trials and errors. Using the selective reproduction of agents and the constant addition of random mutations, most effective strategies can emerge thanks to the research conducted by a succession of generations of agents. According to this second approach, we used GA to simulate learning, where learning occurs across generations of agents rather than during an agent's life. Obviously, we interpret our GA not in biological but in cultural terms [46].

2.3. Model overview

Model's structure has been designed aiming to stylize and reproduce somehow the basic interaction structure leading to the ToC. 100 agents move in an environment which is a grid of 41x41=1681 patches that contains a given number of randomly distributed tokens. When an agent reaches a token, the agent takes possession of the token but another token appears in another position of the environment so that the total number of tokens remains always the same. All agents are equally able to reach the tokens but the speed with which the agents move to reach the tokens varies from agent to agent and this speed is encoded in the agents' genes. The simulation is a succession of generations each composed of 100 agents.

The agents of the first generation have random genes and this means that each agent moves at a different speed. The 10 agents that happen to have better genes and therefore run faster and reach more tokens (fitness) generate 10 offspring each and the offspring inherit the genes of their (single) parent with the addition of some random variations (genetic mutations) that can make some offspring to have better genes and to run faster than their parent. The 10x10=100 offspring are the second generation of agents and the simulation goes on for 5 generations. While moving in the environment to reach the tokens, agents pollute the environment and the quantity of pollution depends on the speed with which they move. An agent which moves faster pollutes the environment more than an agent which moves more slowly. Pollution reduces the agents' fitness. The fitness of an agent and, therefore, the probability that the agent will have offspring depends not only on the number of tokens that the agent is able to reach, but also on the level of pollution of the environment. Living in a polluted environment implies a reduction of fitness which is proportional to the level of pollution.

In this way we have stylized the basic dynamic of the ToC, leading the agents to face with the dilemmatic issue: either they move more slowly and eat fewer tokens and in this way they do not contribute to the pollution of the environment, or they move faster and eat more tokens but they pollute the environment. The problem is that all the agents contribute to the pollution of the environment. If an agent because of its genes "decides" to move more slowly in order not to pollute the environment, this does not mean that the environment will not be polluted because other agents may "decide" to move faster and, therefore, the environment will equally be polluted. An agent has only disadvantages if it moves slowly and, therefore, to have more fitness all agents will move fast and the environment will become progressively more polluted - with an increasing damage for all the agents.

3. Experiments and preliminary results

Starting from the basic interaction structure so far sketched and drawing upon the theoretical framework above outlined to model agents' behavior, we are conducting a series of experiments manipulating different relevant variables of the simulation model. We have introduced in the model the possibility to apply a "sanction" to agents that move too fast and pollute the environment producing therefore a higher level of pollution. The sanction produces a reduction of the fitness of the agent that exceeds a given speed limit and its amount can be varied according to researcher's aim. Our main goal is to explore interplay between selfish rationality, punishment and intertemporal choice on the emergence and development of cooperative behavior exploiting the genetic algorithm as a device to simulate learning processes.

The parameters of the simulation are numerous (token density, agents' life length, speed limit etc.). Each of them can affect the result of the experiment and each of them has a specific semantic value. Currently, we are investigating how the combination of different kinds of sanction with different temporal scenarios affects the evolution of ToC dynamics. The simulation model allows different sanction "regimes":

- *fixed*: the sanction has the same fixed value for each agent that exceeds the limit;
- *speed proportional:* the sanction value is given by difference between real speed of the agent and speed limit;
- *fitness proportional*: the sanction has a value which is proportional (10%) to the fitness of the agent that exceed the speed limit.

   The sanction regimes can be combined with different temporal scenarios in which the effects of pollution and the effects of sanctions are more or less delayed in time. Results so far obtained (Table 1) are showing interesting correlations between the delay of sanction and cooperative behavior: the effectiveness of sanction depends not only on the amount and type of sanction, but also on the time of its application. In particular, we found that a large delayed sanction reduce the effectiveness of the sanction itself, so the agents keep on exceeding the speed limit. Moreover, we also noted that when the sanction is proportional to the agents' fitness it seems to be anyway effective, even if it is largely delayed.

Table 1. Simulation results. The speed limit is set on the same value (2 in a range 1 to 5) during all the experiments. "+" indicates that the sanction is "effective", (the sanction allows to contain the speed of the whole population) under the speed limit; while "x" indicates an "ineffective" sanction.

		Sanction delay		
		None	*Medium (10 ticks)*	*High (30 ticks)*
	Fixed (low: -1 fitness)	x	x	x
	Fixed (medium: -2.5 fitness)	+	+	x
Sanction regime	*Fixed (high: -5 fitness)*	+	+	+
	Speed proportional	+	x	x
	Fitness proportional	+	+	+

4.   Conclusions

Even if very preliminary, results so far obtained allow us to draw some conclusions about future experiments and the developments of the model. As to the first point, we are planning to conduct other experiments sweeping parameters over a range of scenarios to identify uncertainties and important thresholds. Our attention will be focused, in particular, on the interplay between the delay of pollution effects, the sanctions and the temporal dimension of individual choices. More in general, the analysis of the experiments have highlighted the need of a more semantically rich solutions to stylize the way individuals deal with the interaction structure that characterizes the ToC. The hyperbolic discount function so far adopted does not account for cognitive dynamics that play a significant role in determining the evolution of intertemporal choice.

   In real settings, the individual propensity to cooperate with other individuals is conditioned by factors that go beyond the simple (even if temporally discounted) assessment of the costs and the benefits deriving from selfish behavior and sanctions. The choice to cooperate in dilemmatic scenarios (in our case by refraining from polluting or depleting common resources) is also conditioned by the prescriptive power of social norms and, therefore, by all the mechanisms that supports their spreading and stabilization. That's why, taking also cue from a recent simulation work on the cognitive implications of the ToC [53] we are planning to endow agents with a cognitive architecture (a software model of reasoning used for programming intelligent agents) accounting for the process of norm internalization.

Acknowledgements

References

[1] R. Axelrod, T*he evolution of cooperation*. Basic Books, New York,1984.
[2] P. Ball, *Why Society is a Complex Matter. Meeting Twenty-first Century Challenges with a New Kind of Science*, Springer-Verlag Berlin Heidelberg, 2012.
[3] G.S. Becker, *The Economic Approach to Human Behavior*, University of Chicago press, Chicago, 1976.
[4] G.S. Berns, D. Laibson, G. Loewenstein, Intertemporal choice–toward an integrative framework, *Trends in cognitive sciences* 11 (2007), 482-488.
[5] R. Boyd, H. Gintis, S. Bowles, P. Richerson, The evolution of altruistic punishment, *Proceedings of the National Academy of Sciences* 100.6 (2003), 3531-3535.
[6] J. Broersen et al. The BOID architecture: conflicts between beliefs, obligations, intentions and desires. In: *Proceedings of the fifth international conference on autonomous agents*, ACM, New York, 9-16.
[7] M.A. Chakra, A. Traulsen, Evolutionary dynamics of strategic behavior in a collective-risk dilemma, *PLoS computational biology,* **8 8** (2012), e1002652.
[8] T. Chmura, J. Kaiser, T. Pitz, Simulating complex social behaviour with the genetic action tree kernel, *Comput Math Organ Theory Arch* **13 4** (2007), 355–377.
[9] D. Colander, R. Kupers, *Complexity and the Art of Public Policy: Solving Society's Problems from the Bottom Up*, Princeton University Press, Princeton, 2014.
[10] R. Conte, N. Gilbert, G. Bonelli, C. Cioffi-Revilla, G. Deffuant, J. Kertesz, D. Helbing, Manifesto of computational social science, *The European Physical Journal Special Topics* **214 1**(2012), 325-346.
[11] R. Conte, R. Hegselmann, P. Terna, *Simulating social phenomena*, Springer, Berlin, 1997.
[12] R.M. Dawes, Social dilemmas, *Annual review of psychology* **31 1** (1980), 169-193.
[13] P.J. Deadman, E. Schlager, R. Gimblett, R., Simulating common pool resource management experiments with adaptive agents employing alternate communication routines*, Journal of Artificial Societies and Social Simulation*, **3 2** (2000), 2.
[14] A. Desai (ed), *Simulation for policy inquiry*. Springer Science & Business Media, 2012.
[15] J. Epstein, *Generative Social Science: Studies in Agent-Based Computational Modeling*, Princeton University Press. Princeton, 2007.
[16] D.J. Farmer, D. Foley, The economy needs agent-based modelling, *Nature* **460 7256** (2009), 685-686.
[17] E. Fehr, S. Gächter, Altruistic punishment in humans, *Nature* **415 6868** (2002), 137-140.
[18] J.W. Forrester, *Principles of systems*, MIT Press, Cambridge, 1968.
[19] S. Gächter, E. Renner, M. Sefton, The long-run benefits of punishment, *Science*, **322** (2008), 1510.
[20] N. Gilbert, *Agent-based models*, Sage Publications, London, 2008.
[21] H. Gintis, S. Bowles, R. Boyd, E. Fehr, Explaining altruistic behavior in humans, *Evol Hum Behav*, **24** (2003), 153–172.
[22] G. Greenwood, Evolution of strategies for the collective-risk social dilemma relating to climate change, *EPL (Europhysics Letters)* **95 4** (2011), 40006.
[23] G. Hardin, The tragedy of the commons, *Science* **162**(3859) (1968), 1243-1248.
[24] C. Hauert, A. Traulsen, H. Brandt, M.A. Nowak, K. Sigmund, K., Via freedom to coercion: the emergence of costly punishment, *Science* **316**(5833) (2007), 1905–1907.
[25] J. Heinrich, N. Heinrich, *Why humans cooperate. A cultural and evolutionary explanation*. Oxford University Press, Oxford, 2007.
[26] D. Helbing, J. Jost, D. Lane, Social systems and complexity, *Advances in complex systems,* **11 4** (2008), 485-486.
[27] C. Hilbe, M.A. Chakra, P.M. Altrock, A. Traulsen, The evolution of strategic timing in collective-risk dilemmas, *PLoS one*, **8 6** (2013), e66490.
[28] J.H. Holland, Genetic algorithms, *Sci Am* **267 1** (1992), 66–72.

[29] J. Jacquet, C. Hauert, A. Traulsen, M. Milinski, Shame and honour drive cooperation, *Biol. Lett.* **7** (2011), 899–901.

[30] J. Jacquet et al., Intra-and intergenerational discounting in the climate game. *Nature climate change* **3 12** (2013): 1025-1028.

[31] W. Jager, M.A. Janssen, Using artificial agents to understand laboratory experiments of common-pool resources with real agents, *Complexity and ecosystem management: The theory and practice of multi-agent systems* (2002), 75-102.

[32] E.W. Johnston, Kevin C. Desouza (eds). Governance in the Information Era: Theory and Practice of Policy Informatics. Routledge, New York, 2015.

[33] C. Koliba, A. Zia. Governance Informatics: Using Computer Simulation Models to Deepen Situational Awareness and Governance Design Considerations, in E.W. Johnston, K.C. Desouza (eds). *Governance in the Information Era: Theory and Practice of Policy Informatics*. Routledge, New York ,2015.

[34] D. Laibson, Golden eggs and hyperbolic discounting, *Quarterly Journal of Economics,* **112** (1997), 443–477.

[35] N. Lettieri, D. Parisi, Neminem laedere. An evolutionary agent-based model of the interplay between punishment and damaging behaviours, *Artificial intelligence and law* **21 4** (2013), 425-453.

[36] N. Lettieri, S. Faro, Computational social science and its potential impact upon law, *European Journal of Law and Technology,* **3 3**, 2012.

[37] M. Macy, Natural selection and social learning in prisoner's dilemma: co-adaptation with genetic algorithms and artificial neural networks, in W.B.G. Liebrand, M. Messick, *Frontiers in social dilemmas research,* Springer, Berlin, 1996, 235–265.

[38] M. Milinski, D. Semmann, H.J. Krambeck, Reputation helps solve the 'tragedy of the commons'*, Nature*, **415 6870** (2002), 424-426.

[39] S. Moss, Policy analysis from first principles, *Proc Natl Acad Sci* **99 3** (2002), 7267–7274.

[40] E. Ostrom, *Governing the Commons: The Evolution of Institutions for Collective Action*, Cambridge University Press, Cambridge, 1990.

[41] E. Ostrom, J. Burger, C.B. Field, R.B. Norgaard, D. Policansky, Revisiting the commons: local lessons, global challenges, *Science* **284 5412** (1999), 278-282.

[42] W. Pinata, Genetic algorithm and social simulation, in M. Ishizuka, A. Sattar (eds) *PRICAI 2002: trends in artificial intelligence. Lecture notes in computer science*, **2417** (2002), 148–157.

[43] J. Platt, "Social Traps". *American Psychologist* **28 8** (1973), 641- 651.

[44] J. Randers, *Elements of the system dynamics method*, Productivity Press, Portland, 1980.

[45] A.S. Rao, M.P. Georgeff MP BDI-agents: from theory to practice. In P*roceedings of the first international conference on multiagent systems* (ICMAS'95), Victor Lesser, San Francisco, (1995)

[46] R. Reynolds*,* An introduction to cultural algorithms, in: *Proceedings of the 3rd annual conference on evolutionary programming*, World Scientific Publishing, Singapore, 1994.

[47] J. Schindler, Rethinking the tragedy of the commons: The integration of socio-psychological dispositions, *Journal of Artificial Societies and Social Simulation*, **15 1** (2012), 4.

[48] E. Sklar, NetLogo, a multi-agent simulation environment. *Artificial Life* **13 3** (2007), 303–311.

[49] M. Spielauer, What is social science microsimulation?, *Soc Sci Comput Rev* **29 1** (2011), 9-20.

[50] F. Squazzoni, A social science-inspired complexity policy: Beyond the mantra of incentivization, *Complexity*, **19 6** (2014), 5-13.

[51] R. Strotz, Myopia and inconsistency in dynamic utility maximization, *Review of Economic Studies* **23** (1956), 165–180.

[52] K. Troitzsch, Legislation, regulatory impact assessment and simulation, in S. Faro, N. Lettieri, *Law and Computational Social Science,* ESI, Naples, 2013, 57-74.

[53] D. Villatoro, G. Andrighetto, R. Conte, J. Sabater-Mir, Self-Policing Through Norm Internalization: A Cognitive Solution to the Tragedy of the Digital Commons in Social Networks, *Journal of Artificial Societies and Social Simulation*, **18 2** (2015), 2.

Aggregating Expectations to Predict Policy Indices with Information Markets

Niki Nikolakakou[1], Efthimios Bothos[1] and Gregoris Mentzas[1]
[1]School of Electrical and Computer Engineering,
National Technical University of Athens, Greece

Abstract. As our world is becoming increasingly interconnected and globalized, policy modelers and policy makers are faced with complex decisions for challenges like climate change and financial crises. To address such challenges in an efficient and effective way, policy officials require timely and reliable information. Accordingly, the availability of appropriate methods and tools to aggregate scattered pieces of information and stakeholders' expectations is becoming increasingly important. In this paper we focus on the problem of aggregating and interpreting heterogeneous information dispersed among individuals in a timely manner to support public policy making. We propose an approach that involves the creation of belief (Bayesian) networks to model conditional and probabilistic dependencies between policy indices and the use of Information Markets (IMs) for aggregating stakeholder expectations on uncertain policy indices. An example case focusing on predicting Greenhouse Gas emissions shows how our approach can support real life policies and decisions.

Keywords. ICT Solutions for Policy Modelling, Bayesian Networks, Information Markets, Decision support, Uncertainty modeling.

1. Introduction

In our increasingly interconnected and globalized world policy modelers and policy makers are faced with complex decisions for challenges like the climate change and the financial crisis that are shared worldwide and tightly connected with policies across different sectors. An issue of major concern is that solutions for addressing such highly interconnected challenges in a 'system of systems' world, tend to address only subsystems and so fail to anticipate risks, achieve systemic change and avoid unintended consequences of public action. For example, the widespread unforeseen turmoil that hit the global economy as well as the abrupt climate changes point out how unstable and uncertain the future is, and highlight the need for proper tools able to efficiently and effectively inform decisions. Commonly, the process of defining, analyzing and modeling policies involves acquiring and organizing bits and pieces of dispersed information in a way that will more clearly inform a decision that has to be made [1].

We specifically focus on the problem of aggregating and interpreting heterogeneous information dispersed among individuals in a timely manner. We aim to address policy officials' need for concise and reliable information, which means to generate reliable and real-time indicators on the activities of economic and production agents, including their expectations, preference, opinions and related interdependencies.

Using such information, policy makers can understand what the future holds so that current policies can be formed appropriately. For example monetary policies affect economic activity with a lag of six months or more and affect inflation with an even longer lag, so the European Central Bank must anticipate the state of the economy some months in the future in order to make the proper monetary policy decision today [2].

Our approach involves the creation of belief (Bayesian) networks to model conditional and probabilistic dependencies between policy indices and the use of Information Markets (IMs) for aggregating stakeholder expectations on uncertain policy indices. The rest of the paper proceeds as follows. In Section 2 we briefly introduce background information on Bayesian networks and IMs together with the related work. Section 3 provides the details of our approach and an IM tool we have developed for information aggregation. Section 4 presents an example case of our approach. Section 5 concludes the paper with our final remarks and outlook for future work.

2. Background and Related Work

Nowadays several institutions in industrialized countries collect data on sentiments of investors and consumers, and periodically (monthly, quarterly etc.) release aggregated indices. These indices are typically based on surveys among financial and climate experts, households or consumers, in the order of several hundreds to few thousands (see e.g. the Eurobarometer [1] established by the European Commission). A large number of indices exist world-wide and are released by governmental institutions, universities and private corporations. Related literature (see e.g. [2]) shows that these indices closely follow the economic situation as described by different variables, like GDP, inflation or interest rates, and even anticipate turning points in the economic cycle. While the usefulness of monitoring expectations, opinions and sentiments of economic agents is undoubted, one can argue that all these indices are based on a selected number of surveys and a low monitoring frequency resulting in a partial view of more complex dynamics. Moreover survey based methodologies are often not incentive compatible and therefore can be of low precision in revealing the 'real' preferences and expectations of experts and the population at large [3]. In order to address these shortcomings, in the present paper we use Information Markets to aggregate expectations on policy indices and adopt Bayesian networks as a mechanism which allows modelling dependencies between indices.

2.1. Bayesian networks

A Bayesian network (BN) is a graphical structure that allows us to represent and reason about an uncertain domain. More formally, Bayesian networks are directed acyclic graphs whose nodes represent random variables $X = X1; ::Xi; :::Xn$, from the domain of interest, in the Bayesian sense [4]. These can be observable quantities, latent variables, unknown parameters or hypotheses. The edges, E_{ij}, represent conditional dependencies between variables. Each node is associated with a probability function

---

[1] http://ec.europa.eu/public_opinion/index_en.htm

that takes as input a particular set of values for the node's parent variables and gives the probability of the variable represented by the node. The construction process of a Bayesian network commonly follows a knowledge engineering approach [5]. It involves the translation of knowledge that is available in the problem domain to an appropriate graphical structure, the choice of the type of the associated local probability distributions and the setting of the prior probabilities before any evidence is obtained. Once evidence about a current situation becomes known, the beliefs in the network are updated and predictive or diagnostic reasoning can be performed.

Bayesian networks have been extensively used for modeling uncertain decisions and outcomes in various domains, including financial risks and indicators, medical conditions diagnoses, success or failure of products, etc.; see [6] for an analysis of related applications.

2.2. Information Markets

Markets are institutional settings able to efficiently allocate scarce resources and accommodate hedging against various types of risks. Moreover, markets have been known to aggregate and disseminate information into prices [7]. Information Markets (IMs) rely on this property and are designed and run for the primary purpose of mining and aggregating information scattered among participants and subsequently using this information in the form of market values in order to make predictions about specific future events. IMs make use of specifically designed contracts that yield payments based on the outcome of future events. IMs have provided accurate predictions of future outcomes in a wide range of domains, e.g., political elections, economic indices, or sport events etc. [8]. Prices of contracts can be interpreted as a measure of the probability of the event, and the metric depends on contract specifications and market design.

IMs efficiency to aggregate information, relies on sound cohorts. Surowiecki [9] has provided a qualitative analysis of participant characteristics necessary for the market to be trustworthy: diversity of opinion, independence of thought and decentralization of knowledge. Wolfers and Zitchevitz [8] established a theoretical model and provided an account of sufficient conditions under which IM prices aggregate private information held amongst participants. They concluded that, when participants are typically well-informed, IM prices will aggregate information into useful information. In an IM, participants who hold the correct information are rewarded and their wealth increases. On the contrary the assets of those who provide erroneous input decrease. The resulting effect is a higher long term accuracy as 'good predictors' obtain greater influence on the market results due to their increased wealth.

IMs offer incentive compatible mechanisms such as anonymity in the transactions and rewards (usually non-monetary e.g. in the form of prizes). Such features encourage people to provide their true forecasts without fearing the consequences of making errors. For example, participants can feel comfortable in investing on outcomes that they really believe will happen, even if this contradicts their organizations' official position. Also, participants are more likely to take an extreme position that they believe in, without feeling pressure to conform to the consensus forecast.

The combination of Bayesian networks with IMs has been proposed in [10] and [11]. The authors provide a generalized model that includes the decomposition of a target question into a set of related variables and subsequently the update of the variables with the use of market estimates. The model has been evaluated to forecast

the risk of Greece leaving the European Union in [10] and state failed indices in [12]. In this paper we extend the work of Berea et al., [10] and provide an approach to predict policy indices of interest.

3. Proposed Approach

We aim to support the aggregation of expectations on policy indices through a four step process as shown in Figure 1.

| 1. Define Model Structure and Initialize Probabilities | → | 2. Select Uncertain Questions and Initialize IM | → | 3. Invite Stakeholders in IM and aggregate expectations | → | 4. Update Model using aggregated expectations |
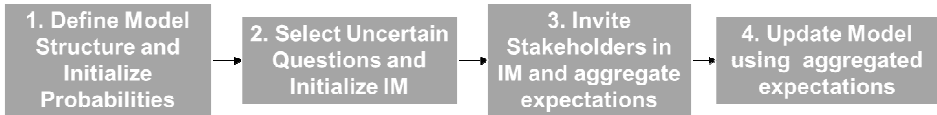
Figure 1. Overview of our approach for aggregating expectations on policy indices.

The first step refers to Bayesian decomposition and parallels Bayesian networks modeling and analysis projects. Policy makers and policy modelers define the model structure of a Bayesian network that decomposes complex policy indices. Simpler indices are used in order to compute the expectations of more complex. In practice, the experts define the policy indices of interest and their interrelations in a graphical model. In order to complete the Bayesian network, policy experts set the prior probabilities according to their expectations. The second step involves the selection of a subset of indices that are uncertain and for which interested stakeholders can provide expectations to forecast their future values. These indices are introduced in the IM in the form of questions. The third step is to run the IM. The stakeholders (who could be knowledgeable experts from public or private organizations) are invited and they start investing on the outcomes of the questions which they believe are going to be correct. The final step is to update the Bayesian network using the aggregated expectations from the IM. The network can be updated at regular intervals with updated expectations from the IM.

For the purposes of our approach we have implemented a play money IM called PolicyOracle designed to aggregate expectations on economic, environmental and energy indicators. Policy modelers and policy makers define and publish questions on policy issues they need information along with a set of mutually exclusive and collectively exhaustive answers. The answers comprise the contracts of the IM. Participants can buy or sell them at a price that varies between 0 and 100 points.

Participants are interested stakeholders and experts. They register in the platform and receive an initial amount of 20000 points which they can invest on the answer(s) they believe will be correct by buying corresponding contracts. If they change their opinion due to e.g. access to new information and expect that a different answer will be the correct, they can modify their portfolio to express their updated opinion immediately in the market. For example they can sell the contracts of the answer they now see as not correct and invest on the answer they believe will be correct. The trading interface is displayed in Figure 2. Participants select the answer they want to invest and the amount of shares they want to buy or sell. The software provides information related to the cost of the transaction and the new contract prices. If the user agrees, she/he confirms the transaction and her/his portfolio is updated. In addition, the software offers functionalities for users to provide comments and participate in

discussions, enabling information and opinion exchange. Finally, a background information section provides a detailed description of the question as well as links to external information sources while price charts show fluctuations in the prices of contracts over time (see Figure 3). The application is available at http://experts.policyoracle.org/, while the current set of questions focuses on the following policy indices: Inflation, Consumer Price Index, Unemployment, GDP, Brent Oil price and Emissions Trading System (ETS) price of carbon.
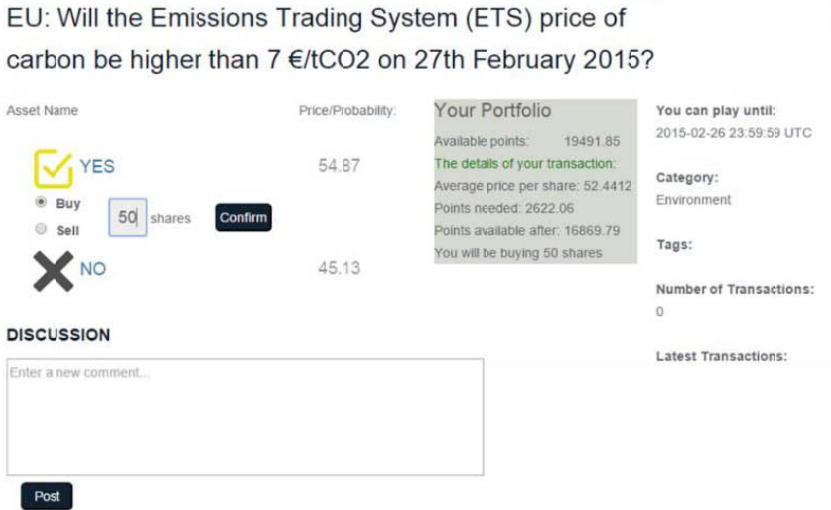
Figure 2. View of the PolicyOracle trading interface. Users can buy or sell shares of contracts that represent different answers to questions related to trends of policy indices.
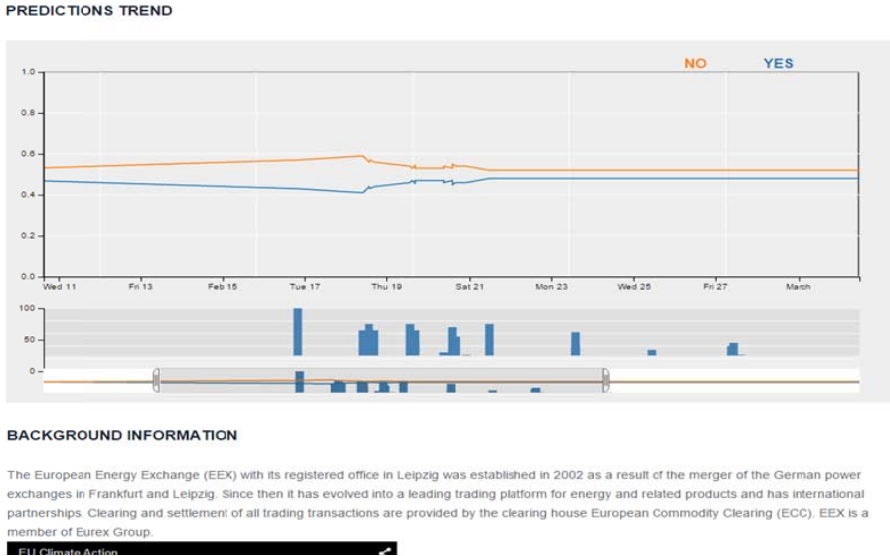
Figure 3. Price chart and Background Information section. PolicyOracle is a virtual stock market.

Buying and selling transactions in PolicyOracle affect the prices of the contracts which are set by an automated market maker (AMM). The AMM can be conceptualized as an agent who is always available to buy or sell contracts based on a price function that determines the cost of each transaction. Our implementation is based on the concept of logarithmic scoring rules introduced by [13]. When the outcome of a question becomes known the final contract values are set. Contracts of the correct answer worth 100 points each while contracts of the false answer worth 0 points. This means that participants who invest on the right answer earn points otherwise they lose.

4. Predicting Greenhouse Gas Emissions

In this section we demonstrate our approach through a trial case focusing on the topic of Greenhouse Gas (GHG) emissions. Preventing dangerous climate change is a strategic priority for the European Union and since GHG emissions consist a major factor for climate change, Europe is working hard to cut its emissions substantially while encouraging other nations and regions to do likewise. This policy agenda is inline with the "Copenhagen Accord" [14] which underlines that climate change is one of the greatest challenges of our time and accepts that in order to prevent the most severe impacts of climate change, global warming should be kept below 2ºC compared to the temperature in pre-industrial times. That means a temperature increase of no more than around 1.2°C above today's level. In practice, the world must stop the growth in global greenhouse gas emissions by 2020 at the latest, reduce them by at least half of 1990 levels by 2050 and continue cutting them thereafter. The EU has made a unilateral commitment to reduce overall GHG emissions from its 28 Member States even higher for 2020, by 20% compared to 1990 levels, which is one of the headline targets of the EU 2020 strategy. According to latest estimates, total EU GHG emissions in 2013 fell by 1.8% compared to 2012, which brings them at around 19% below 1990 levels [15]. This keeps the EU on track to meet its 20% target by 2020.

Our aim is to construct a network of factors influencing GHG emissions in order to predict whether emissions will increase or decrease. We define our hypothesis node as the probability to the question: "Will the GHG emissions in EU at the end of 2015 be higher compared to 2014?". The next step is to decompose the hypothesis into parent nodes (these are the causal or evidential nodes). Starting from a literature review we identified that GHG emissions depend on population size and economic output as expressed in the Gross Domestic Product (GDP) (see e.g. [16]). Moreover we consulted policy modelers from the organizations German Climate Forum and Germanwatch with expertise in environmental sustainability who advised us to incorporate dependencies on the share of renewables in energy consumption, the price of ETS carbon emissions and the price of crude oil. The Bayesian network we constructed is presented in Figure 4. We now have a short model and we initialize its conditional probabilities. The elicitation of prior probabilities were based on expert judgment and are shown in Figure 4.
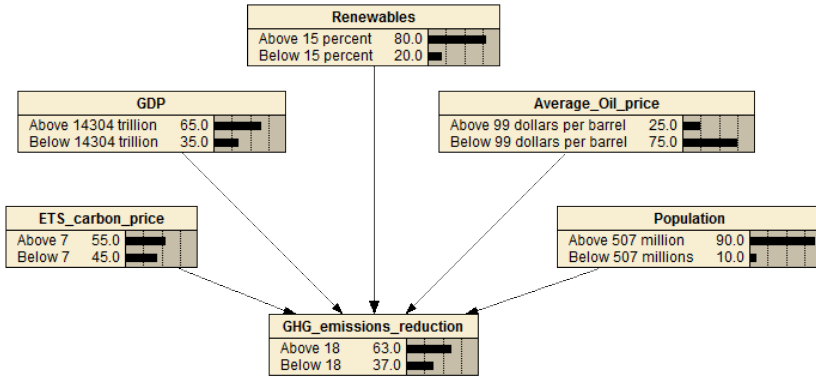
Figure 4. Basic influence diagram for the Bayesian network on GHG emissions.

The probabilities of the parent nodes can be inferred from experts participating in our IM. For this, we have to present them in the form of questions. Considering the network presented in Figure 4, we formulate the following set of questions with a binary (yes /no) outcome, each corresponding to one of the parent nodes:

- Will the population in EU be higher than 507 millions in December 2015?[2]
- Will the GDP in EU be higher than €14 trillion in Q4 2015?[3]
- Will the share of renewable energy in energy consumption in EU be higher that 15% in December 2015?[4]
- Will the average of closing Brent oil price be higher than $99/barrel in 2015?[5]
- Will the Emissions Trading System (ETS) price of carbon be higher than 7 €/tCO2 on 30th December 2015?[6]

While the IM is running, participants express their expectations and modify the probabilities of the answers to the above questions. The Bayesian network can be regularly updated with the new evidence coming from the IM leading to changes of the estimate of the hypothesis node. Depending on the results, policy makers may validate or modify their actions. For example, if the experts' expectations reveal a significant drop of the share of renewable energy in final energy consumption and this causes increase in GHG emissions, they can design new policies to reverse the situation. On the other hand if their policies are performing well and the network shows a decrease of GHG emissions, they can feel confident that their selected policies are effective.

---

[2] Using as a baseline the EU population at the end of 2014 which was 507 millions: http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode=tps00002&plugin=1

[3] Using as a baseline the EU GDP at the end of 2014 which was 14 trillions: http://sdw.ecb.europa.eu/quickview.do?SERIES_KEY=MNA.Q.Y.I7.W2.S1.S1.B.B1GQ._Z._Z._Z.EUR.V. N

[4] Using as a baseline the share of renewables in EU at the end of 2014 which was 15%: http://ec.europa.eu/eurostat/tgm/table.do?tab=table&init=1&language=en&pcode=t2020_31&plugin=1

[5] Using as a baseline the Brent oil price for 2014 which was $99.57/barrel: http://www.investing.com/commodities/brent-oil-historical-data

[6] Based on reports of ETS price of carbon: http://www.eex.com/en/

The questions will be launched within spring of 2015, upon the publication of the EU GHG inventory report for the 2014 year by the European Environment Agency and will be active until the end of 2015. Our plan is to evaluate the usefulness and accuracy of our approach with the help of experts to whom we will regularly present the outcome of the Bayesian network and the expectations provided by our IM application.

5. Conclusions & Future Work

We presented our approach to predict policy indices using Bayesian networks to model interdependencies of events and IMs to aggregate expectations from interested stakeholders. Currently, we are in the process of running the trial case presented in Section 4 and we expect to have the evaluation results available until September 2015. We will be collaborating with experts and policy modelers as part of the project SYMPHONY (http://projectsymphony.eu/).

As the work reported in this paper is ongoing research, a number of questions require further investigations. We have described the design of a case focusing on GHG emissions, however there is a large number of uncertain events in policy design and analysis which could be supported by our approach. We are currently looking into modeling and predicting complex economic indicators, including the change in consumer spending (propensity to consume), citizens' willingness to switch job and citizens' willingness to participate to labor market. Furthermore we are looking into combinatorial designs of our IM, following the work described in [11]. In such a setting participants will be able to express their expectations by investing on combinations of events; for example, a participant could invest on the event that the propensity to consume is going to increase, conditional to an increase in GDP and a decrease of unemployment. With such formulations the IM will be able to capture expectations on interrelated and dependent variables modelled in the Bayesian network and lead to more accurate predictions.

Acknowledgements

References

[1] J. O. Ledyard, Designing information markets for policy analysis. Information Markets: A New Way of Making Decisions, RW Hahn and PC Tetlock (editors), AEI'Brookings Joint Center for Regulatory Studies (2006).
[2] C. D. Carroll, Macroeconomic expectations of households and professional forecasters., the Quarterly Journal of economics (2003), 269-298.
[3] T. T. Kranz, F. Teschner, C. Weinhardt. Combining Prediction Markets and Surveys: an Experimental Study. Proceedings of the 22nd European Conference on Information Systems, (2014).
[4] D. Heckerman. A tutorial on learning with Bayesian networks. Springer Netherlands, 1998.
[5] A. Zagorecki, M. J. Druzdzel, Knowledge engineering for Bayesian networks: How common are noisy-MAX distributions in practice?. Systems, Man, and Cybernetics: Systems, IEEE Transactions on, 43(1) (2013), 186-195.

[6] N. Fenton, M., Neil, Risk assessment and decision analysis with Bayesian networks. CRC Press (2012).

[7] V., L., Smith, Markets as economizers of information: experimental examination of the "Hayek hypothesis". Economic Inquiry 20(2) (1982), 165-179.

[8] J. Wolfers and E. Zitzewitz. Prediction markets. The Journal of Economic Perspectives, 18 (2), 107-126, 2004.

[9] J. Surowiecki. The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations. Doubleday Books, 2004.

[10] A., Berea, D., Maxwell, C., Twardy, Improving Forecasting Accuracy Using Bayesian Network Decomposition in Prediction Markets., 2012 AAAI Fall Symposium Series, (2012).

[11] W., Sun, R., Hanson, K. B., Laskey, C., Twardy, Probability and asset updating using bayesian networks for combinatorial prediction markets. arXiv preprint arXiv:1210.4900, (2012).

[12] A., Berea, C., Twardy, D. T., Maxwell, Forecasting the Failed States Index with an Automated Trader in a Combinatorial Market, Journal of Strategic Security 6(5) (2013), 6.

[13] R. Hanson. Combinatorial information market design. Information Systems Frontiers, 5(1):107{119, 2003.

[14] Accord, Copenhagen. Decision 2/CP. 15 Copenhagen Accord. FCCC/CP/2009/11/Add. 1, 2009.

[15] European Commission report, 'Progress towards achieving the Kyoto and EU 2020 objectives', COM (2014) 689.

[16] V., Detlef, P., Bert de Vries, A., Beusen, P., Heuberger, Conditional probabilistic estimates of 21st century greenhouse gas emissions based on the storylines of the IPCC-SRES scenarios., Global Environmental Change 18(4) (2008), 635-654.

Constructing a Methodology Toward Policy Analysts for Understanding Online Public Opinions: A Probabilistic Topic Modeling Approach

Nan ZHANG[a] and Baojun MA[b]

[a] *School of Public Policy and Management, Tsinghua University, Beijing 100084, P. R. China*
[b] *School of Economics and Management, Beijing University of Posts and Telecommunications, Beijing 100876, P. R. China*

Abstract. Public opinion always has an important influence on the policy process. The development of social networking sites and applications has given the public more opportunities to express their views about the related policies. In cases where the coverage of the traditional hearing system challenged the policy process, how to measure accurately the public concern and attitudes regarding policies based on online public generated content using a data mining method will be very important issue in policy informatics. Our paper provides a probabilistic topic modeling approach, mainly based on Latent Dirichlet Allocation (LDA) model, to transform the complex semanteme of online public opinions into the values could be measured. A simple case could show the usefulness of the too toward policy analysts also be provided and discussed briefly.

Keywords. Online Public Opinions, Public generated content, Probabilistic Topic Modeling, Latent Dirichlet Allocation (LDA), Policy Informatics.

1. Introduction

The media, public and decision makers are considered the three primary factors in the traditional policy agenda settings [1]. The prior research believes that these three factors are independent of each other and mutually affect each other [2]. In recent years, the rapid development of Web 2.0 applications, such as Facebook, Twitter and Wikipedia, has provided citizens with more approaches to participate in public policy discussions more easily. Ordinary people have the opportunity to express their opinions and exert an unprecedented influence as long as their viewpoints are typical and comprehensive. It seems likely that ordinary people, who lack discourse power in real society, will become the opinion leaders on the new media overnight. Hence, the age in which we live now is called the 'we-media era' by scholars in the communications field [3], which means that the roles of the public and media are deeply integrated in policy agenda settings. Therefore, the capability of the professional media, such as newspapers and TV, to initiate topics for discussion and lead public opinion may be further weakened while, in contrast, public expression will play an increasingly important role in the process of making and revising policies.

These new characteristics of the age mentioned above have undoubtedly led to more challenges associated with obtaining public attitudes and feedback as was the case during the traditional policy process. The hearing system that was generally adopted before the policy promulgation is being increasingly questioned [4]. It seems that, no matter how the participants of the hearing are selected, it is impossible to satisfy everyone, which constitutes an unavoidable hindrance to the representative system in such an increasingly flat world [5]. Moreover, even online opinion surveys, that emerged with the popularity of the network, face challenges: those stakeholders directly benefiting from certain policies or the group or individuals who pay more attention to certain issues usually perform in the manner of voting several times in support of their own views, or inviting friends to vote in support of those views to exaggerate the scale of their own side which, to a certain degree, is misleading to policy makers. However, as the social network has become a crucial way for the public to express their attitudes and exchange ideas, it is possible to find a reliable method to determine public attitudes and examine the mass view in order to create policy from the fragmented expression of opinions that are to be found via various social network applications?

The direct approach that has been applied is public opinion analysis, based on keyword searches and filtering [6]. Obviously, this type of method may only describe the problems roughly, leading to a boom in related studies in the field of communication, and it seems that the method cannot really support the policy process at present. The biggest challenge originates from the definition of "data" and the expansion of the processing method, and a deeper analysis of public generated content on social media must surmount the boundary of Codd (1982)'s relational database [7]. The progressive data mining methods, including the semantic analytical methods, have made it possible to analyze these unconventional data. The related analysis contributes not only to investigating the degree of public support and practical effects made during the phase of implementing policies, but also the collecting of wider public opinion during the policy-making stage to engender more extensive public participation in the policy-making process. The decision makers need to make political responses rapidly by employing semantic analysis and other methods to learn about the public attitudes towards various policy agendas on the Internet quickly and accurately, which should also be one of the issues considered and addressed by the new research branch, policy informatics.

Under such a context, this research attempts to construct, based on Latent Dirichlet Allocation (LDA) [8], an effective approach to measuring the popularity and polarity of the discussion by means of refining related subjects from a large amount of web text. The method we present is explored by taking, as an example, public attitudes drawn from an online community regarding the car license lottery policy promulgated at the end of 2010 in Beijing, China.

2. Framework of the Methodology

2.1. Briefly Review of Topic Modeling and Latent Dirichlet allocation

Topic modeling algorithms are statistical methods that analyze the words of an original text to discover the latent themes or topics that run through it, originally proposed and applied in the field of information retrieval [9].

Without directly dealing with the topics inherent in the documents, the TF-IDF schema and vector space model (VSM) provided a rough solution to describing and modeling documents and their content or topic similarities [10, 11], with the disadvantage that documents with a similar context but different vocabulary will not be associated with each other [12]. To deal with this problem, latent semantic analysis (LSA) has been proposed to convert the high-dimensionality word-space representations of the documents into low-dimensionality vectors of the topics [13], in which topics can be obtained using singular-value decomposition (SVD). Using such a technique yielded some improvement over a TF-IDF baseline.

A closely related technique−pLSI [14] −tries to create a set of topics in a probabilistic framework [15]. The topics in pLSA are probabilistic instead of the heuristic geometric distances in LSA. pLSA has been successfully used with large collections regarding information retrieval, because its does not need to run the expensive SVD operation.

Latent Dirichlet allocation (LDA) is a popular topic modeling tool, designed to learn a set of topics (word distributions) and infer mixtures of these topics to build low-dimensionality representations of documents [9], which further refines the pLSI model within a Bayesian framework [14]. Among all of these topic models, LDA appears to be the most effective [8, 16]. A simple introduction to the LDA model will be presented in more detail in the following section.

The intuition behind the LDA model is that documents exhibit multiple topics [9]. LDA generates ''topics'' as lists of words drawn from the vocabulary used in the text corpus; the topic is based on the distributions of those words over the vocabulary. The topics are generated inductively by the model based on the likelihood of words to co-occur within documents. LDA produces the topics through a probabilistic approximation of Bayesian inference. Starting with a set of seed topics (often randomly generated), the algorithm iteratively alters these topics to best match the set of data being learned. LDA also generates proportions for each document for each topic, so that each document can be described as being proportionally composed of (or, interpretively speaking, "about'') a number of topics that are expressed by the words used in that document. For details about the algorithmic and computational aspects of LDA, see [8].

Topic modeling is a good match for a data source like public comments. Because the procedure is automated, it enables analysis of much larger text corpora than would be feasible by hand. Since topics are generated inductively by the model and not predefined by the researcher, the technique protects against implicit coding bias caused by the constraints of researcher knowledge. And because the method assumes that a single document can contain multiple topics (in contrast to some other document clustering methods; see [17]), it enables researchers to draw insights from interrelations among themes, both within documents and within the dataset as a whole.

As described, public comments are rich data because commenters express all manner of concerns in unstructured ways. Comments run the gamut from technical specifications to personal stories and from thoughtful reflection to threats and name-calling. Many comments defy easy categorization as being for or against the proposal at issue. Thus, hand-coding such documents can be a particularly difficult task; topic modeling appeals because it can uncover hidden patterns in even a varied set of documents.

Moreover, it is noted that the LDA model does not require any prior annotations or labeling of the documents and the topics emerge from the analysis of the original texts.

The LDA model enables us to organize and summarize electronic archives on a scale that would be impossible using human annotation [9].

2.2. Overall Consideration and Detailed Procedures

Given the increasing amount of content about public policy feedback being available online in the Web 2.0 era, especially via blogs, microblogs, Wikipedia and social networks, the policy makers or government do not have the time to read and study all of them to learn about the detailed contents and viewpoints of the public. In addition, due to the characteristic of being an online public communication, the topics or themes of public discussion tend to be diverse and may evolve, led by someone with time, even about a certain event or general topic. Thus, the later public discussion may deviate from the original topic, and the total content may contain many issues that are irrelevant to the policy makers or government.

To this end, we have designed a methodology framework for public policy feedback monitoring by applying a probabilistic topic modeling algorithm, specifically Latent Dirichlet Allocation (LDA) [8], which will be introduced in detail in the next subsection. In the framework, the LDA model has been utilized to discover and annotate large archives of public-generated documents with thematic information to determine how these themes are connected to each other and how they change over time. Thus, we could provide the policy makers or government with filtered, accurate content that they are actually concerned about.

We will introduce the detailed procedures of public policy feedback monitoring by applying the LDA modeling. The steps involved in probabilistic topic modeling and concerned content filtering are as follows:

- Step 1: Data collection and pre-processing.

- Step 2: Probabilistic topic modeling using the LDA model.

- Step 3: Choosing interested or concerned public policy topics.

- Step 4: Document topic assignment and concerned content filtering.

- Step 5: Concerned document hotness and relevance calculation.

Firstly, the original relevant data need to be collected from the online platform based on certain events, topics, persons or contents in which the policy makers are interested, which are normally conducted by web crawlers or programs [12, 18, 19]. Then, in order to explore the text content information further, several necessary pre-processing operations must be performed, such as word segmentation [20, 21], stemming [22, 23] and stop-word removal [18, 19]. Thereafter, each text's content is represented by a vector of words.

In Step 2, the core task lies in using the LDA to conduct probabilistic topic modeling. Specifically, LDA modeling applies training and inference to all of the text vectors and can discover any latent topics or themes inherent in these data [9], as will be discussed in the next subsection in greater detail. After the topic modeling, we obtain the following useful results related to the words, documents and topics in these documents:

- Latent topics with the most likely words in each topic;

- A Topic-document distribution matrix.

Theoretically, for any given topic, there is a corresponding distribution across all of the words in the vocabulary, and the LDA modeling process could provide the most likely words with the highest probability with respect to each topic, whereby the number of the most likely words can be given in advance. After the LDA modeling, the thematic keywords based on the given topic number become available to policy makers, and we could invite them to select topics of interest to them. Thus, this selection step is intuitive, easy to implement, and would not impose much time and cost on the policy makers. More conservatively, the topics or themes of concern to the policy makers or government could be easily predicted or learnt in advance.

In Step 4, we first utilize the topic-document distribution matrix to conduct topic assignment for each text document. In detail, the topic or topics with the highest document-topic probability will be assigned to each document. More flexibly, the topics with the top t highest document-topic probabilities could be simultaneously assigned to each document, with $1 < t < K$ based on either the preference of the policy makers or the document number to be filtered. By combining the topics of concern selected in Step 3 and the topic assignment results together, documents about irrelevant topics could be easily filtered out.

Finally, after we obtain all of the contents of interest regarding public policy feedback, some related statistical information, such as the documents' hotness and cumulative relevance ratio per day, could be calculated and provided for the policy makers. Specifically, to show the discussion hotness of these topics each day, the cumulative document number concerned can be easily obtained. In the meantime, in order to evaluate the overall relevance of the documents concerned with respect to the topics concerned, the daily cumulative relevance and cumulative relevance ratio can be calculated.

3. Case and Discussion

3.1. Case Background and Analysis Results

The government of Beijing promulgated the policy of a car license plate lottery at the end of 2010. In our research, we selected the forum named "AutoWorld" in Shuimu Community as the targeted concern. AutoWorld forum is one of the most active forums in Shuimu Community, especially around and after the time when the lottery policy for buying vehicles in Beijing City was announced on December 23th, 2010. Since data in the recycling box can only be kept for around four or five months, in order to obtain all of the data from the AutoWorld forum about this event, the time window was chosen of around three months, which was specifically from December 15th, 2010 (i.e., around a week before the policy announcement) to April 5th, 2011. After removing the duplicate posts that appeared in both the forum and the recycling box, such as the collection post information, we finally obtained a total of 359,715 unique posts.

We conducted probabilistic topic modeling for all 359,715 post texts using the LDA model with the topic number 25, 50, 75, 100, 125, 150, 175 and 200 respectively. As we discussed above, for any given topic, there is a corresponding distribution across all of the words in the vocabulary, and the LDA modeling process could provide the most likely words with the highest probabilities with respect to each topic, in which the

number of the most likely words can be given in advance. In our case, we set this number at 100. By manually looking through all 100 thematic keywords for different topic numbers, it was found that the semantic results for the thematic keywords for each topic appear best when the topic number is 50.

In this case, we assume that the policy makers and government are notably interested in public feedback on the lottery policy of car license plates in Beijing City, as introduced in the background section. Thus, public discussion of topic 10 (i.e., the public countermeasure to the lottery policy) and topic 27 (i.e., the lottery policy) were chosen as the topics of concern or interest in our case. Then, through the procedure of document topic assignment (the highest probability strategy) and concerned content filtering, we finally obtain over 23,074 related posts, with 9,543 about the lottery policy itself and 14,146 about the public countermeasure. Figure 1 represents the daily discussion hotness of the policy and public countermeasure based on our proposed LDA-based framework, showing the important policy time points.
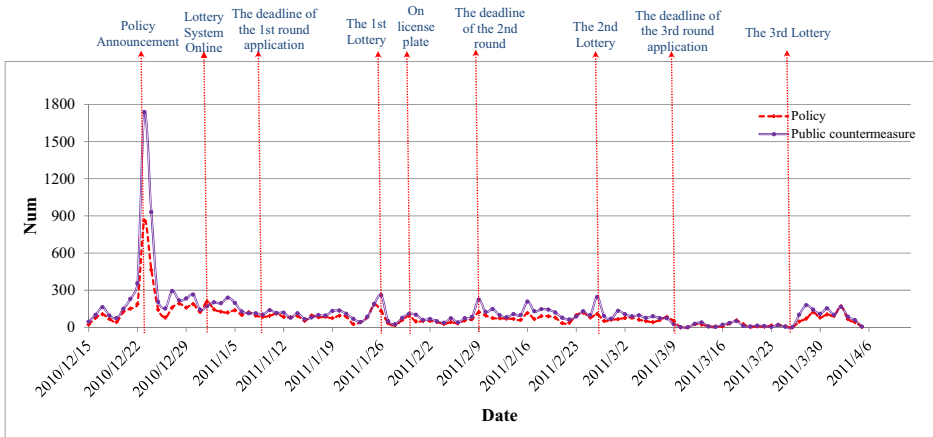
Figure 1. Comparison of daily discussion hotness of policy and public countermeasure.

Moreover, we also compared our results based on LDA modeling with those from the keywords filtering strategy, in which only containing corresponding policy terms, such as "摇号" (i.e., lottery) or "限购" (i.e., quota) in Chinese, would be retained. In detail, if any post contains at least one of the above policy terms, its original post and all of its reply posts would be regarded as potential content of interest, which resulted in 31,628 posts in total. Figure 2 shows the results of the comparison between our method and the keywords filtering method in relation to important policy time points. It is obvious that, compared with our LDA-based method, the keywords filtering strategy usually suffers from the disadvantage of containing too much irrelevant information, ignoring topic transfer as well as containing words that do not discuss the real topic, as explained above.
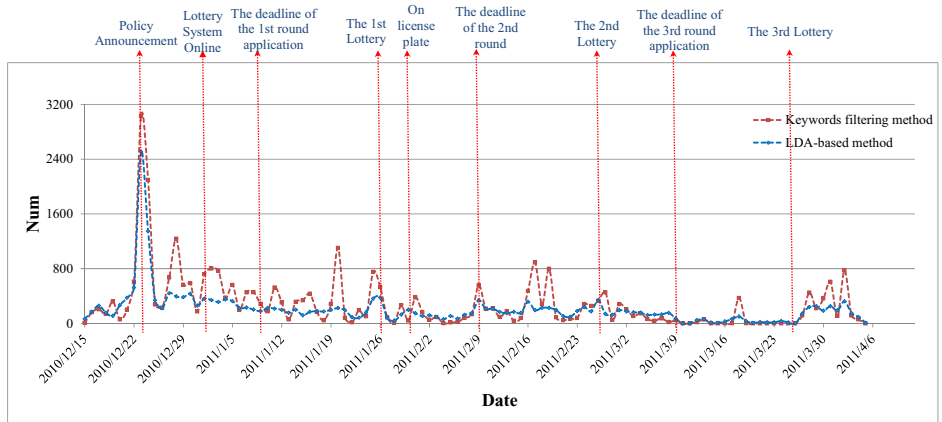
Figure 2. Comparison of daily discussion hotness by LDA-based method and keywords filtering method.

3.2. Concluding Remarks

The results of the case indicate that it is possible to obtain more accurately contents of interest from the mass of information by using the method based on LDA and be in a position to be better able to describe public concerns, and their types and polarity, related to the policy process via the related analysis. The core idea of the proposed methodology is transform the complex semanteme a relation matrix between texts and topics. After the transformation, we could construct many functions based on matrix values from different perspectives. This case only discusses one of the simplest functions and describes the trend of the hot topic. However, we could understand deeper policy related issues based on function construction.

In the age of large-scale data, a large amount of web text information that rapidly develops every day should explicitly serve as a treasure trove to be mined by policy analysis science, which will necessarily require more data mining methods and applications for policy analysis. The process undoubtedly requires that scholars of data mining and policy analysis should cooperate with each other, as in the case of the paper. It is expected that the method explored in this paper and its implications will facilitate both academia and practitioners to gain a better understanding of the opportunities and challenges existing in this rising research field.

Acknowledgements

References

[1] Lazarsfeld, P. F., & Merton, R. K. (1971). Mass communication, popular taste and organized social action. In Media Studies: A Reader (2nd ed., pp. 18-30). Edinburgh: Edinburgh University Press.

[2]   Wanta, W., & Hu, Y.-W. (1993). The agenda-setting effects of international news coverage: An examination of differing news frames. International Journal of Public Opinion Research, 5(3), 250-264.

[3]   Cottle, S. (2006). Mediatized rituals: Beyond manufacturing consent. Media, Culture & Society, 28(3), 411-432.

[4]   Leighninger, M. (2012). Public Hearings ≠ Public Values. Public Administration Review, 72(5), 708-709.

[5]   Freidman, T. (2005). The world is flat. New York: Farrar, Straus and Giroux.

[6]   Wojcieszak, M. (2011). Pulling Toward or Pulling Away: Deliberation, Disagreement, and Opinion Extremity in Political Participation*. Social Science Quarterly, 92(1), 207-225.

[7]   Codd, E. F. (1982). Relational database: a practical foundation for productivity. Communications of the ACM, 25(2), 109-117.

[8]   Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. The Journal of Machine Learning Research, 3, 993-1022.

[9]   Blei, D. M. (2012). Probabilistic topic models. Communications of the ACM, 55(4), 77-84.

[10] Jones, K. S. (1972). A statistical interpretation of term specificity and its application in retrieval. Journal of Documentation, 28(1), 11-21.

[11] Salton, G., Wong, A., & Yang, C. S. (1975). A vector space model for automatic indexing. Communication of the ACM, 18(11), 613-620.

[12] Liu, B. (2007). Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data: Springer Berlin Heidelberg.

[13] Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. Journal of the American Society for Information Science, 41(6), 391-407.

[14] Hofmann, T. (1999b). Probabilistic latent semantic indexing, Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval (pp. 50-57). Berkeley, California, USA: ACM.

[15] Papadimitriou, C. H., Tamaki, H., Raghavan, P., & Vempala, S. (1998). Latent semantic indexing: a probabilistic analysis, Proceedings of the seventeenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems (pp. 159-168). Seattle, Washington, USA: ACM.

[16] Blei, D. M., & Lafferty, J. D. (2009). Topic models. In A. Srivastava & M. Sahami (Eds.), Text Mining: Classification, Clustering, and Applications (pp. 71-93). Boca Raton, FL: Taylor & Francis Group.

[17] Carterette, B., & Chandar, P. (2009). Probabilistic models of ranking novel documents for faceted topic retrieval, Proceeding of the 18th ACM conference on Information and knowledge management (pp. 1287-1296). Hong Kong, China: ACM.

[18] Baeza-Yates, R., & Ribeiro-Neto, B. (1999). Modern information retrieval: Addison-Wesley New York.

[19] Manning, C. D., Raghavan, P., & Schütze, H. (2008). Introduction to information retrieval. Cambridge, England: Cambridge University Press.

[20] Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-Driven Computations in Speech Processing. Science, 298(5593), 604-607.

[21] Yang, C. D. (2002). Knowledge and learning in natural language. New York: Oxford University Press.

[22] Hull, D. A. (1996). Stemming algorithms: A case study for detailed evaluation. Journal of the American Society for Information Science, 47(1), 70-84.

[23] Porter, M. F. (1980). An algorithm for suffix stripping. Program, 14(3), 130-137.

Policy modelling and maintenance in social housing: motivating sustainability and managing change

Bojan CESTNIK[a,1] and Alenka KERN[b]

[a] *Temida d.o.o. and Jozef Stefan Institute, Dunajska cesta 51, 1000 Ljubljana, Slovenia,*
bojan.cestnik@temida.si
[b] *The Housing Fund of The Republic of Slovenia, Poljanska cesta 31, 1000 Ljubljana,*
Slovenia, alenka.kern@ssrs.si

Abstract. Public policies are collections of principles that guide state government actions in a given domain. The aim of public policy used by the state Housing Fund for distributing social housing facilities in Slovenia is to ensure transparent and equally fair treatment of all beneficiaries. During the last decade the Fund has established a sustainable structure for policy models that consist of three parts: written requirements for applications, a detailed process description, and a comprehensive decision support model containing rules for ranking the applications according to their priority. In this paper we analyze the main causes of changes in the policy in the last decade and investigate how they are reflected in the underlying conceptual and decision support models. The results of our analysis indicate that the key critical success factor was to maintain a delicate balance between the sustainability of the models and changes in the policy. As a result, the state housing Fund was able to leverage the supply on the Slovenian housing market by efficiently conducting twelve consecutive tenders for offering apartments under favorable terms to citizens.

Keywords. e-Government, policy modelling, policy governance, social housing, change management

1. Introduction

The Housing Fund of the Republic of Slovenia, public fund (hereinafter the Fund) has since its foundation been actively involved in a number of different tasks and activities. Through these actions the Fund wanted to improve conditions and opportunities in the housing area for the citizens of the Republic of Slovenia. Since the establishment the Fund has undergone numerous changes and adjustments of business strategy and activities, mostly because of the external factors such as changes in the legislation, the progress in the information technology, the web development and the economic crisis. The Fund was established in 1991 with the Housing Act as a central national institution for financing or implementation of the national housing program or the promotion of housing construction, renovation and maintenance of apartments and houses. In the first year after the establishment, the Fund has published its first tender for favorable

---

[1] Corresponding Author: Bojan Cestnik, Temida d.o.o. and Jozef Stefan Institute, Dunajska cesta 51, 1000 Ljubljana, Slovenia, bojan.cestnik@temida.si.

long-term loans for refinancing unfavorable bank loans, which was followed by a tender of housing loans for the first purchase, construction or renovation of the apartments. Until 2005, the Fund has with its favorable loans helped to solve the housing problems of over 30.000 Slovenian citizens. In 1999 the Fund published a public appeal for saving in the National Housing Saving Schema [3, 8], which has a goal to insure savings for real estate, especially apartments and consequently after the end of saving, which was usually 5 years, increased the demand for apartments.

Public policy studies are used to analytically evaluate the past and innovatively explore the future [5]. Policy studies may refer to the policy process, to the substance of policies themselves, or to their effects [13]. The policy studies are typically concerned with the process and substance, since the outcomes require a longer perspective and different approach [7]. This study aims to fill the gap by describing the analysis of the policy outcomes and, based on the findings and observations, as well as other contributing factors like management strategy, proposing methods to sustainably maintain the underlying policy, preserving the policy's transparency, openness, and collaborativeness. In general, policies should be designed to be stable and enduring [11]. Therefore, eventual changes should be taken into account by encapsulating them in a relatively isolated and localized frame within the policy. Note, however, that overemphasizing the importance of policies and focus on policies alone might in some cases lead to the so called "bureaucracy myopia" [6].

The problem addressed in this paper is how the analysis of the policy outcomes observed in a given time period influences changes in the policy process and/or substance. The underlying goal is how to design and maintain a policy governing a particular strategic task so that it is stable (i.e. it does not change often and much in process), but it at the same time allows for transparent introduction and reflection of subtle changes in substance required by the management and state regulations. The addressed problem is described in the context of distributing housing facilities, which is itself a strategic and delicate task. When dealing with delicate and problems concerning general public interest, the responsible government officials are expected to, apart from general knowledge of the problem domain and other related approaches, exhibit skills for efficiently handling communications with the concerned citizens.

This paper is organized as follows. First, we give an overview of the methodology used in our study. In section 3 we describe the results of the analysis of the policy changes during the twelve consecutive tenders for selling housing facilities to citizens by the Fund. Lessons learned from the presented approach are discussed in section 4. The most important findings are summarized in the conclusions.

2. Methodology

One of the important governmental actions to stimulate private and business endeavors is to give out incentives through the programmes of distributing funds and grants [2]. Typical process phases for such actions include (1) publishing a call that defines eligibility criteria and application requirements, (2) application gathering, validation and processing, and (3) informing beneficiaries about the decisions. The last step might include also complaints gathering phase and related procedures. The goal of such governmental actions is first to inform the potential target population and second to distribute the earmarked funds or grants objectively according to the policy, principles and criteria published in the call.

The purpose of the work described in this paper is analyze the main causes of changes in the policy used in the tenders for distributing apartments by the Fund in the last decade and investigate how they are reflected in the underlying conceptual and decision support models. The context of our study is a decision support system for selling apartments by The Housing Fund of the Republic of Slovenia to citizens [4]. One of the Fund's missions is to balance the demand and supply in the real estate market by offering apartments at favorable prices. Since the Fund's offer represents only a small fraction for the total supply of apartments, a priority ranking is used to establish the preferential order of applicants. In case when two or more applicants fall into the same priority rank, a single buyer is selected by a random choice [12]. For this reason, the whole business process is popularly called the Housing Lottery [4]. Since its first use, the described system received considerable media attention. In the past 13 years the system has been used to support twelve consecutive tenders that are described in Table 1. It implements a complex decision policy that was adopted by the Fund's management. Besides, its main characteristics are robustness, flexibility and transparency. The system provides the transparent and understandable insight in all the steps of the process and includes more controlling mechanisms which assure integrity and the confidentiality of the processed data.

Table 1. Twelve public tenders described by the number of available apartments, submitted and approved applications, and number of categories for ranking the applications

Label	Year	Apartments	Applications	Approved	Categories
SA	2002	78*	36	30	4
SB	2002	73	135	73	4
SC	2002	32	311	32	10
SD	2003	429	1.831	450	38
SE	2004	852	2.057	872	220
SF	2006	258	424	262	180
SG	2006	116	3.427	118	308
SH	2007	387	285	238	352
SI	2008	140	24	24	352
SJ	2010	24	76	49	1.134
SL**	2012	504	493	393	80
SM	2014	342	403	239	108

The business process of selling the market apartments consists of six main steps. In the first step, the Fund prepares the offer or public tender for sale of apartments and publishes it on its website and in the media, normally daily newspapers and on the radio [4]. The first requirement that comes from the business process is the suitable presentation of offered apartments, which provides interested customers with precise insight in the buildings and as such eases their decision. Interested applicants must then fill in the prescribed web form – an application for the purchase of the apartment, which is a part of the tender's documentation. In the third step, the Fund's employees identify and validate the received applications and assure that the gathered applications are correctly and accurately stored in the database that is then used by the special computer programme, which was specially designed for this purpose. In the next step all the completed applications are ranked according the priority category and are allowed to participate in the allocation process. When there are two or more applicants ranked in the same priority category, the fifth step carries out the procedure of random selection, so there is just one buyer of the apartment selected. The main goal of this

business process is to ensure the correctness of the information and consequently the allocation of the apartments to the applicants from the highest priority rank. In the sixth step all the applicants are informed of the outcome of the sub process of allocating the apartments and the selected buyers are thereafter invited to sign the contract of sale.

The quality of services that are carried out by an organization depends heavily on organization's ability to understand, model and control business processes [9. 10]. Due to relatively complex nature of the Fund's operations, carefully designed and transparently presented policies are required. The concrete detailed policies, like the policy for distributing apartments, are aligned with the state housing policy governing general processes and priorities in the housing field. Besides, they facilitate the introduction of various controlling mechanisms, thereby increasing the reliability of the Fund's operations.

The Fund's policy for selling apartments consist of three parts: written requirements for applications, a detailed process description, and a comprehensive decision support model containing rules for ranking the applications according to their priority. The first two parts are represented as written documents prepared by the Fund's legal department and are for every tender published on the Fund's web page. The third part representing the decision model is represented in the form of a table containing decision rules that unambiguously determine the actual rank of each applicant. Typically, it is prepared in a collaborative meeting including all important stakeholders. After reaching a consensus, a set of decision-support rules that encapsulate the views and priorities to be used in the current tender is formulated. The main purpose of the decision model, which is also published on the web site, is to clarify eventual ambiguities that might arise from the textual documents alone. It is the responsibility of the Fund's management to confirm the correctness and validity of all the documents that constitute the policy for each particular tender.

The general, the requirements for each tender are not supposed to change much during the years. However, the variables that determine the preferential ranking order for the applications are subject to change both from the Fund's management and other stakeholders like the Ministry and general public. The changes are of four types: (1) adding new attributes, (2) changing the order of importance of the attributes, (3) changing the domain values of the attributes, and (4) changing the interpretation of the attribute values. All the four types of the changes must be reflected both in the textual documents and, even more importantly, in the decision model.

In Table 1 twelve past tenders for selling apartments are described with the label, year, number of available apartments, number of submitted applications, number of approved applications and number of categories for ranking the applications. All the tenders were used to support selling apartments to citizens, except for the tender SL (labeled with **) which is used to support renting apartments to citizens. In the first three tenders in 2002 the grounds were set for the following tenders. In all the three tenders there were 135 apartments available; if they were not sold in SA (labeled with *) tender, they were moved to the subsequent SB and SC tender. At the beginning, in tenders SA and SB, there were only 4 categories for ranking. Then, in the next tenders, the number had risen to several hundredths and even 1.134 in tender SJ. The reasons for this behaviour are explained in the next section; they can be mostly accounted to the introduction of new attributes, as well as the general tendency of the Fund's management trying to minimize the number of random selections within the process.

The process of selecting the attributes for the decision making is described in the next section. Note that the actual variable selection and preferential order of columns depends on the guidelines obtained from the state government housing policy, the Fund's management, and the observed opinions of the concerned general public. The final decision table is carefully produced on the joint meetings to incorporate directives acquired from all three sources. The knowledge and experience from the past Fund's projects [1] were very beneficial for handling and supporting group consensus meetings.

The trend of growing the number of categories for ranking can be observed in the first ten tenders (SA to SJ). While the aim of introducing more ranking categories was to reduce the unpopular number of random selections (i.e. housing lottery) between equal ranks, the implications of such solution were twofold. First, random selections were just smaller in number and could not be completely avoided. And second, the resulting decision table got huge and more difficult to understand and interpret without the computer technology. Therefore, observe the "normalization" of the number of categories for ranking in the last two tenders SL and SM.

3. Analysis and results

In this section we present our analysis of the outcomes of twelve tenders for selling apartments by the Fund. We describe the factors that influence the outcomes of a particular tender and identify the main causes for the decision policy changes in the next tenders.

In year 2002 the Fund published its first tenders (SA, SB) for selling apartments at prices that were lower than on the housing market. The applicants who were granted by a loan from the Fund and the savers in the National Saving Scheme had the advantage of buying these apartments and among them especially young families under the criteria of age of the parents. The next tender (SC) defined as a main priority category for selecting buyers of apartments saving in the National Saving Scheme and among them additional priority for those that already concluded their saving. Young families had once again the priority inside the category of the savers.

In 2003, the Fund within the offer (SD) also introduced additional priority for disability apartments; there was an absolute priority for physically disabled persons at buying these apartments. In 2004, the first savers, this means 21.000 people, concluded their saving. In this year the Fund offered the biggest number of apartments, altogether 852 on the public tender (SE). In spite of the stated the expected interest was bigger than the offer. This is why it was necessary to create the decisive criteria for priorities at the purchase. The main target group were the savers that completed their saving and among them additional priority for young families, young people and families. For fulfilment of the status of the young family there were still the criteria of parents' age; the young people were people under 30 years. Among young families and families there was an extra priority for families with more children.

In 2006 the supplementation to the legislation from the field of saving in the National Saving Scheme was accepted. Among other things it redefined the status of a young family, which wasn't linked anymore to the age of the parents but to the age of their children. According to the change of the legislation and to the other changes, as well on the real estate market, the Fund has within the framework of public activity carefully created the priority categories (tenders SF and SG) for its tenders for apartments' sale and this has turned up to be the right decision. During the preparation

stage of the decisive criteria, the expert services passed numerous studies and analyses were the basis for the decisions.

Decision-making is one of the central activities of management and is a huge part of any process of implementation, especially because of the fact that this time there were just 116 apartments to offer on the market in the capital city where the needs are the biggest (tender SG). In order to ensure the successful implementation of the public tender it was necessary to carefully and transparently prepare complete tender's documentation and the priority classes are a part of it. The decisive criteria were adopted as it follows: young family, family, physically disabled persons and young couples. Among young families and young couples there was an additional priority for those who were solving their housing problem for the first time. The main novelty of the tender was a five-year preemptive right to limit potential speculative practices.

The decision-making criteria were in force until the 2014 (for tenders SH, SI, and SJ), when there were additional defining decisions over some of the priority categories in the spirit of the business policy adopted in 2012. The young families with at least one pre-school child stay as the main priority category. They are followed by families with schoolchildren, families with dependent children, physically disabled persons and young persons. According to the carried out analyses and set goals the decision was made that the young person is a person younger than 27 years (tenders SL and SM).

In Fig. 1 the number of approved and rejected applications according to priority ranks is shown for the twelve conducted tenders of the housing Fund. For easier comparison, all the categories are normalized to the interval 1 to 20. The graph shows approved and rejected applications from the lowest category (the highest rank) to the highest one (the lowest rank) in a cumulative fashion. Observe the general pattern that in the first categories the ratio between approved and rejected applications is the highest. This observation can be used to justify the use of priority ranking as a mechanism for selecting suitable candidates. Note also that the actual number on the y axis (number of applications) may differ from the tender to tender.

Let us compare the resulting graphs for different tenders depicted in Figure 1. As stated earlier, in the first three tenders there were only 4 categories (10 in SC). At the beginning, the used categories were well accepted by the general public; however, the omission of the social categories like number of children, family's financial status, etc., was soon noted in the communication. The number of categories grew from tender to tender, culminating in the number above 1.000 for the tender SJ. In the last two tenders (SL and SM) a simplification of the decision table was suggested; they include 80 and 108 categories respectively. Note also that the actual number of submitted applications to the tender does not depend on the number of categories; it most heavily depends on the attractiveness of the location of apartments sold in a particular tender.

In Figure 1 the numbers of approved and rejected applications for three distinguished tenders (SE, SG, and SM) were depicted. Tender SE can be viewed as a typical tender; there were 852 apartments available, 2.057 applications were submitted and 872 granted. Note that the last number can be higher than the first, because 20 applicants decided to withdraw after the approval. There were 220 categories for ranking. The apartments were mostly allocated to applicants with the higher ranking order, so there were only a few complaints after the completed process. In tender SG, on the contrary, the number of applications (3.427) was more than 20 times higher than the number of available apartments (116). Even though there were 308 categories for ranking, the graph depicted in Fig. 1 shows that the majority of the applicants were rejected, resulting in a severe pressure for complaints.
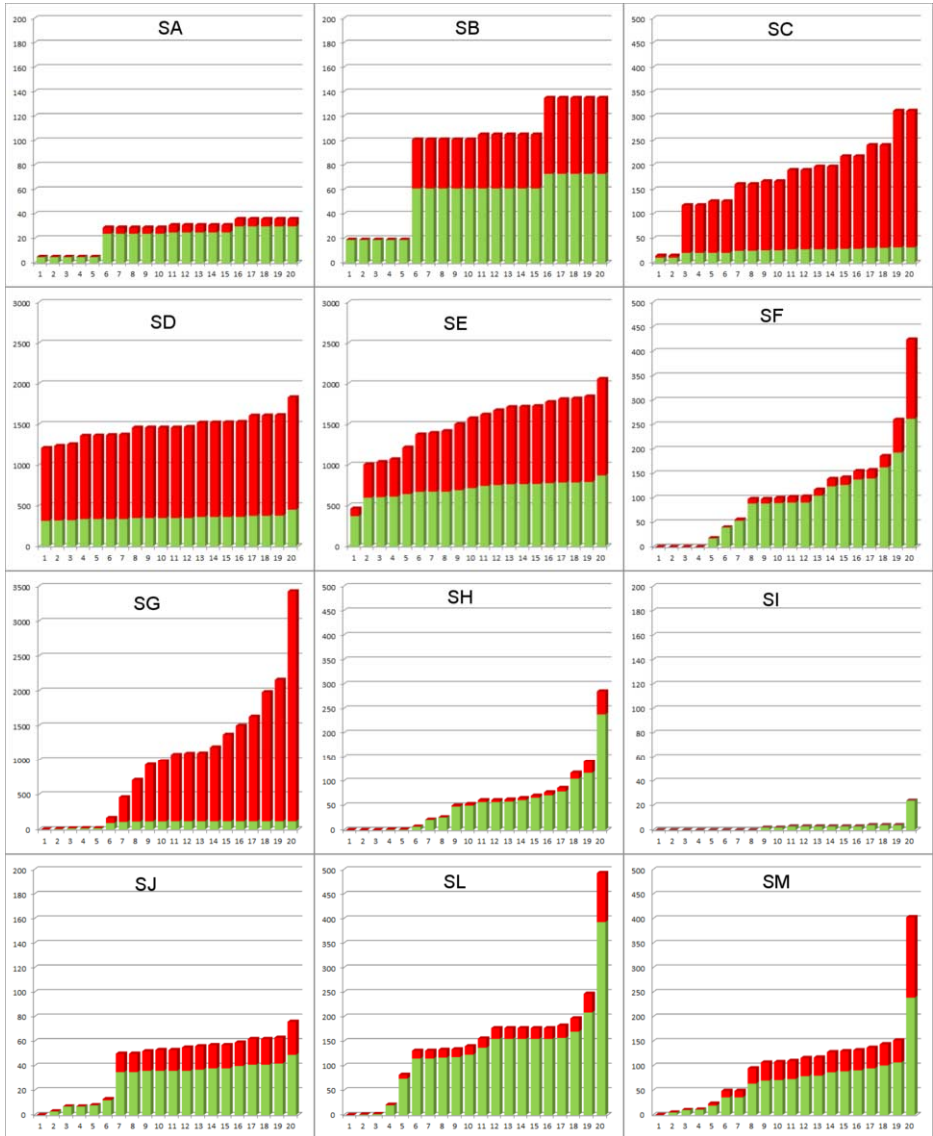
Figure 1. A comparison of the twelve tenders for selling apartments by the Fund showing number of approved (green) and rejected (red) applications according to priority ranks.

The current tender SM, which is still active in 2015, exhibits a different behaviour in terms of categories and application approval. There is a long tail visible in category 20; also, the number of approved applications in this category is quite high. The main reason for this can be found in the fact that after the first cycle, in which only a half of available apartments were sold, there were 7 additional cycles, where applicants with less privileges can also compete for the remaining apartments.

4. Conclusions

In this paper we analyze the main causes of changes in the policy used by the housing Fund for selling apartments under favorable terms to citizens in the last decade and investigate how they are reflected in the underlying conceptual and decision support models. The policy studies are typically concerned with the process and substance, since the outcomes require a longer perspective and different approach. This study describes the analysis of the policy outcomes of twelve consecutive tenders for selling flats in the period of 13 years. The results of our analysis indicate that the key critical success factor was to maintain a delicate balance between the sustainability of the models and changes in the policy.

In the context of distributing housing facilities, which is itself a strategic and delicate task, we have designed and described an approach to constitute and maintain a policy governing the task under study. The resulting policy remained stable in structure and contents through the years; at the same time, the proposed approach allows for transparent introduction and reflection of subtle changes in substance required by the management and state regulations.

The analysis described in this paper was carried out on twelve consecutive tenders for selling apartments by the housing Fund. For future work we plan to extend the scope and include the state projects of similar contents (distributing subventions, grants, etc.) and compare the findings between different fields.

References

[1] M. Bohanec, B. Cestnik, V. Rajkovič, A management decision support system for allocating housing loans. In: P. Humphreys (ed.), *Implementing systems for supporting management decisions: concepts, methods and experiences*, 1st ed. London: Chapman & Hall, 1996.

[2] Capegemini, Rand Europe, IDC, Sogeti and Dti, *Digitizing Public Services in Europe: Putting ambition into action - 9th Benchmark Measurement*, December 2010.

[3] B. Cestnik, A. Kern, IT for supporting the Slovenian National Housing Savings Schema: example in G2B interoperability, Eastern European e|Gov days, Prague, 2007.

[4] B. Cestnik, A. Kern, H. Modrijan, The housing lottery in Slovenia: e-Government perspective. Electronic Government, 6th International EGOV Conference, Regensburg 2007. Proceedings of ongoing research, project cont. and workshops. A. Groenlund, H.J. Scholl, M.A. Wimmer (Eds.), Trauner Verlag, Linz.

[5] Y.C. Chen, D.V. Dimitrova, Civic Engagement via E-Government Portals: Information, Transactions, and Policy Making. E-Gov Research: Policy and Management (ed. D. Norris), IGI Global, 2008.

[6] D. Graeber, *The Utopia of Rules: On Technology, Stupidity, and the Secret Joys of Bureaucracy*, Melville House, 2015.

[7] M. Keating, *The Government of Scotland: Public Policy Making after Devolution*, Second edition, Edinburgh University Press Ltd, 2010.

[8] A. Kern, B. Cestnik, Housing Subventions in Slovenia: Experience from the Second Year. Eastern European e|Gov days, Prague, 2008.

[9] D.S. Linthicum, *Next Generation Application Integration, From Simple Information to Web Services*, Addison-Wesley, 2004.

[10] J. Nograšek, M. Vintar, Observing organisational transformation of the public sector in the e-government era, *Transforming Government: People, Process and Policy*, Vol. 9: 1, 2015, pp.52 – 84.

[11] T.R. Peltier, *Information security Policies, Procedures and Standards: Guidelines for Effective Information Security Management*, Auerbach Publications, A CRC Press, 2002.

[12] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, 1988, pp. 283.

[13] M. Yildiz, E-government research: Reviewing the literature, limitations, and ways forward. *Government Information Quarterly*, Vol. 24, Issue 3, July 2007, pp 646–665.

[14] J.A. Johnson, M. Friesen, *The Success Paradigm: Creating Organizational Effectiveness Through Quality and Strategy*, New York: Quorum Books, 1995.

This page intentionally left blank

Open Government and Smart Cities

This page intentionally left blank

91

Delivery of Public Value to Multiple Stakeholders through Open Government Data Platforms

Gabriela Viale PEREIRA [a,1], Marie Anne MACADAR [a] and Maurício Gregianin TESTA [a]

[a] *Pontifical Catholic University of Rio Grande do Sul*

Abstract. Governments around the world are trying to realize the benefits of technology to make a real difference in people's lives. The use of data provided by open government data platforms has a great potential to enable new services, improve the lives of citizens and make government and society work better. This research proposes a conceptual model to explain how open government data platforms are used to enhance the access to and delivery of government information and services in order to make a real difference in people's lives. This study explores and illustrates the model by taking the perspective that Open Government Data (OGD), as a platform, influences the way city agencies are delivering information and services to increase feedback for stakeholders, including citizens, government agencies and employees. To analyze the impact on citizen's lives, government agencies and employees we used the public value perspective that can helps to determine the value of government activities from multiple stakeholders' perspectives. Specifically, this working paper presents preliminary results of our analyzes from the NYC Open Data portal. We aimed to understand how citizens, employees and other agencies are using OGD to improve their activities, the impact of that and unintended consequences.

Keywords. Open Government Data, e-government, Public Value.

1. Introduction

In the last decades the use of information and communication technologies (ICT) has made significant advances into diverse aspects of social life [1]. The application of ICT to government – or electronic government – has been considered a driver of social, economic, political and developmental changes such as government administrative reform, social transformation, organizational change and economic growth [1, 2]. However, the impacts and results associated with electronic government projects, including the organizational transformation capacity, are still not clear and well known [1].

The use of ICT in government is an innovative way for governments to offer rich opportunities for governments that significantly improve the delivery of services and to interact more openly with the stakeholders [3, 4, 1]. E-government can be defined as the transfer of government information and services among governments, their customers and suppliers [5].

---

[1] Corresponding Author.

Future research in this area might make the connection between e-government and the traditional concerns of public administration [5] including the improvement of citizen's lives.

One important way to improve the quality of services that governments provide to business and citizens is by a customer-orientation [6]. This focus on customer/citizens-orientation are shared by many researches to analyze user perceptions of e-government projects and intentions of use or expectations [7, 8, 9] or to focus on the relationship between the e-government stakeholders [1] and the coordination and connectivity between them [7, 10]. However, it's not clear how governments interact with the stakeholders to provide services and information that fits with what people really want. Helbig et al. [6] emphasizes the lack of attention on citizens' needs and questions if users actually want, or could they use, what government were given.

New technologies are making it possible to use data generated by public organizations and build services that automatically answer questions related to public administration issues. In this case, despite the fact that data required is not always available in a form which is easy to use, it has a great potential to enable new services, to improve the lives of citizens and to make government and society work better [11].

Considering the latest trends in technology (cloud computing, big data, mobile), future referrals in federal governments should embrace a digital strategy that comprises the opportunity to innovate more with less, and enables stakeholders to better leverage government data to improve the quality of services [12, 11]. Governments around the world are trying to realize benefits of technology to make a real difference in people's lives. The use of data provided by open government data platforms have a great potential to enable new services, improve the lives of citizens and make government and society work better. Government as platform is a model that democratizes the real time exchange of information and services, allowing people to use information to make important personal decisions, but also is increasing the quality of weather forecasts, transit information, and health alerts which are all generated from government data that further improve people's quality of life [13].

Open data is a way to mend the common separation between government and users, and is helping public organizations act as an open system that interacts with its environment [14]. However, due to the early stage of Open Government Data (OGD) initiatives, their impact and ramifications are still debated in professional and academic circles [15]. Existing research enrolls public value through open government and emphasizes the necessity of an analysis of government activities from multiple stakeholder perspectives [16], the public value resulted from e-government initiatives [17] and the generation of value from OGD initiatives [15, 18]. However, is not clear how government agencies are using OGD to increase the delivery of government information and services to other stakeholders; the feedback of this process; or the overall impact of it.

This paper addresses the following question: To what extent Open Government Data are helping governments' agencies to enhance the delivery of public value to multiple stakeholders? Therefore, this article aims to develop a conceptual model including a stakeholders' perspective of OGD in aspects such as application, contribution and impact of government information and services. To explore and illustrate it this study takes the perspective that OGD influences how cities agencies are delivering information and services and increasing the feedback for the stakeholders, including citizens, government agencies and employees. To analyze the impact on citizen's lives and on government agencies and employees we take the public value perspective which can help to determine the value of government activities from multiple stakeholders' perspectives [16].

This study partially addresses the gap suggested by Helbig et al. [6] who criticizes the lack of attention on citizen's needs. We argue that citizens are getting involved in

the processes of the government and are creating value for both. Additionally, we analyze how agencies interact with each other and they are using OGD to increase their own processes, as predicted by Goldsmith and Crawford [19]. To achieve this goal, we present preliminary results, based on our propositions and conceptual model, by analyzing the use of predictive policing in the New York Police Department (NYPD) and the New York City (NYC) Open Data Portal.

2. Open Government Data

Moving beyond the application of ICT to government, the Digital Government Strategy has a new key priority area in terms of the use of technology to better serve citizens by innovating ways to deliver better digital services using fewer resources [12]. Open Government Data is related to one of the problems that the Digital Government Strategy sets out to address--unlocking data and improve the quality of services for the people [12]. This means that the government seeks to enable the public, entrepreneurs, and their own government programs to better leverage the richness of federal data through inputs into applications and services by ensuring that data is open and machine-readable by default.

The open data movement has a principle that data should be available to all who have a use for the information and can be accessed and re-used for any purpose [19, 11]. Open data is defined by the Open Definition as "any data and content that can be freely used, modified, and shared by anyone for any purpose" (http://opendefinition.org). It means that open data must be available and accessible, must permit re-use and redistribution and allow universal participation, i.e. everyone can use open data without discrimination by fields, persons or groups [11].

Proponents of Open Government Data believe that the new role of the public sector as an information provider will strengthen democracy and improve the impacts of government work through increased transparency, participation and collaboration [15]. The principles of Open Government are: Efficiency – obtaining increased outputs or goal attainment with the same resources or with lower resource consumption; Effectiveness – increasing the quality of the desired outcome; Intrinsic enhancements – changing the environment or circumstances of a stakeholder in ways that are valued for their own sake; Transparency – access to information about the actions of government officials or operation of government programs that enhances accountability or influence on government; Participation – frequency and intensity of direct involvement in decision making about or operation of government programs or in selection of or actions of officials; and Collaboration – frequency or duration of activities in which more than one set of stakeholders share responsibility or authority for decisions about operation, policies, or actions of government [16].

Some of the areas where open government data is creating value include: transparency and democratic control, participation, self-empowerment, improved or new private products and services, innovation, improved efficiency of government services, improved effectiveness of government services, impact measurement of policies and new knowledge from combined data sources and patterns in large data volumes [11].

3. Conceptual Model and Discussion

To achieve the paper's objective, we have created a conceptual model that shows the relationship between OGP, e-government and its impact in citizens and other govern-

ment agencies. The Figure 1 illustrates the logic of our proposal, based on several different concepts.
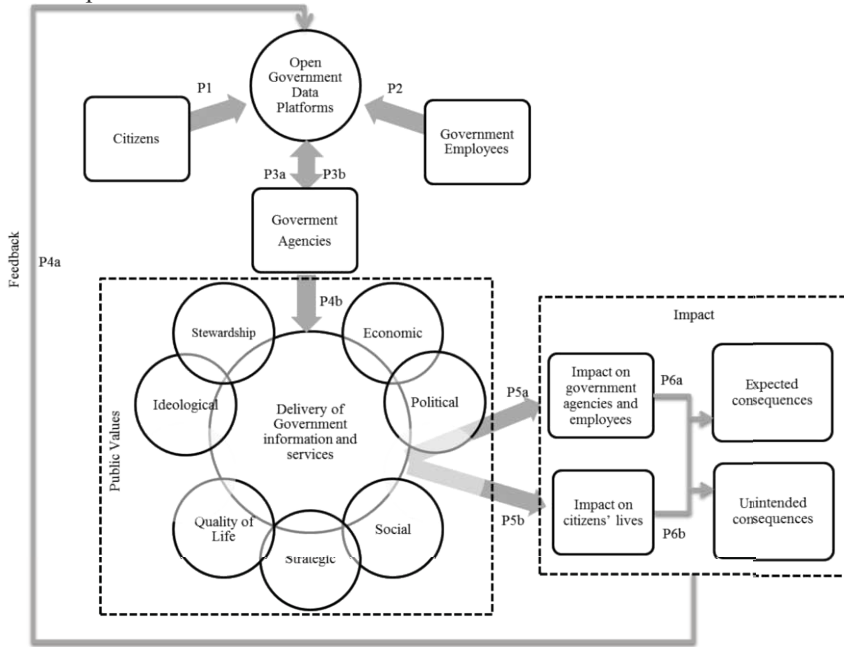
Figure 1. Conceptual model.

3.1. Open Government Data

With an analysis of the supply and demand side of e-government, we can examine the phenomena in terms of open government data initiatives. The supply side of e-government focuses on initiatives that create electronic services (like open government data platforms) and opportunities for participation from citizens, businesses and other stakeholders (government agencies or employees, for example). The demand side includes how people are using these services or other ICT to their advantage within society [6]. In the same way, citizens' demand for information and the role of government responsiveness are considered drivers of open data and represent the demand and supply side of OGD initiatives [20]. Rowley [21] brings the notion of stakeholder benefits and looks at them as consumer-side to try to understand their objectives and interests from e-government. In this way, it is possible to focus more directly on what stakeholders want from an e-government service, allowing governments to compare and integrate the perspectives of all stakeholders.

"Across the globe, people are uniting in a common struggle: to participate freely in the events and processes that shape their lives" [Mahbub ul Haq apud 22]. Open data platforms make it possible for people to get involved in the processes of the government and create value for both [20]. As more information is provided through OGD platforms, citizens will find information about government activities more easily and complain and communicate their opinion, representing an increase in their interest in actively participate in the government process. Thereby, **Proposition 1**: *Citizens are being engaged in the provision of services and getting involved in the processes of government through open government data.*

Digital solutions give city leaders the opportunity to empower government employees to promote better lives for citizens [19]. Departments that use predictive policing, for example, not only enable more efficient management of resources and manpower, but also empower street-level agents to make the best possible decisions. Also, OGD platforms can be used as a tool to search for relevant data, detect relationships, and analyzing millions of lines of data in real-time. This capability helps leaders address many urban challenges and make smarter and earlier decisions. Thereby, **Proposition 2**: *Government employees are working toward better lives for citizens through open government data.*

Local, national and international public and private organizations are publishing their data in open data infrastructures, resulting in OGD [23]. The other opportunities for city leaders through digital solutions are that they can engage with citizens in the provision of services and enable citizens to work with local government on shared solutions [19]. **Proposition 3a**: *Government Agencies are promoting public use of city data through open government data.*

Government responsiveness is the extent to which the government responds to issues raised by citizens [20]. Citizens with diverse social and economic background possess different needs and they pressure their government in distinct ways for new e-government services. Such demands must be controlled by governments to continually improve their services to satisfy all kinds of citizens' needs [24]. Also, the OGD platforms can help governments learn from the experiences of other agencies and city governments in their attempts to improve the way in which governments can deliver better services [4]. Thereby, **Proposition 3b**: *Government Agencies are using open government data to respond to the issues raised by the citizens and to learn from other government agencies and employees.*

3.2. Delivery of Government information and services

"OGD advocates argue that it enables greater government efficiency through an information infrastructure that allows for better data re-use within the public sectors and inter-agency coordination" [15, p.2]. Considering digital government as a system, coordination and connectivity are characteristics which evolve the ways agents in a system connect and relate to each another, forming patterns from the interactions and connections [10, 7]. The same relation can be applied to analyse the relationship between the stakeholders in e-government initiatives and how they interact through OGD. Also, considering that there is often a gap between what government officials initiate, what they expect, and what is really delivered to the citizens, government officials need external and objective feedback on their e-government efforts and effects [24].

"Public value focuses attention on the collective and societal interests that are served by particular institutional arrangements and actions by government" [16, p.2]. The seven types of value that capture the range of possible results of government actions are: economic, political, social, strategic, quality of life, ideological, stewardship. Thereby, **Proposition 4**: *Government Agencies are using open government data to enhance a) the access to (feedback) and b) the delivery of Government information and services to the citizens, other agencies, and other Government entities (Public value).*

3.3. Results of the Delivery of Government information and services through OGD

One of the factors for the success of e-government initiatives is the coordination and integration of inter-governmental agencies at all levels, and that roles, processes and policies

are reflected in systems and in the interaction of government with stakeholders [7]. If solutions in terms of e-government are investigated by looking at the context-use-access perspective, it is possible eliminate the unintended consequences of narrowly defining the problem [6]. Technology has dramatically altered the way in which government and citizens relate. Information flow has increased in terms of speed and capacity between the two entities and the expanded use of database and web technologies has changed the horizontal and vertical integration of government and citizens [25]. Thereby, **Proposition 5**: *The more Government Agencies use open government data as inputs to Delivery Government information and services, a) the greater the results on the way that government agencies and employees work (government-to-government/employees); and b) the greater the results on citizens' lives (government-to-citizens).*

3.4. Expected and unintended consequences

The changes faced by governments as a result of the digital revolution may result in a number of possible negative effects, such as information misuse, information inequity, and privacy violations, all enabled by the increasing use of advanced information technologies [25]. Regardless of the change of focus in e-government efforts, the e-government phenomenon can also present some pitfalls which should be known by the public, such the incompatibility between a security-oriented perception of e-government and at least three of the original founding principles of the e-government phenomenon:, open government, transparency and responsiveness [5].

Just two of the concerns regarding the topics of security and privacy consider the privacy impact of constantly increasing amounts of information available online and, consequently, the possibility of misuse of this information or in a security breach resulting in the loss of confidential data [25]. Because of the complexity involved with the publication and use of open data, is not easy to predict who, how and when users use open data [23]. Due to this uncertainty, some agencies avoid publishing open data to prevent unintended consequences of OGD. An especially prominent topic is the use of government databases to centralize medical and criminal records, in order to maintain data consistency and web technology to facilitate data transfer, but also presents issues such as safety, convenience, and security for citizens [25].

Finally, considering the knowledge citizens needs to have the ability to use this technology [7] and access to data might be limited, [10] OGD programs can further contribute to the digital divide. Thereby, **Proposition** 6: *The results on (a) the way that government agencies and employees work and (b) the results on citizens' lives from the use of open government data as inputs to Delivery Government information and services by Government Agencies may result in unintended consequences rather than the expected results.*

4. Research Strategy and preliminary results

This section aims to test the conceptual model previously proposed, based on e-government and OGD literature and secondary data based in city governments' agencies in United States; We collected secondary data from an open government data initiative in United States: The NYC Open Data Portal and we analyze how the NYPD are using to become more responsive.

New York City has a thriving data community and one of the strongest open data portals in the nation. "A comprehensive data strategy would build on past successes,

improve the city economic and social well-being, and help city government better meet the needs of its citizens" [26]. The Mayor's Office of Data Analytics (MODA) worked to provide insight to NYC leadership and agencies, to lead the City's expansion of the Open Data Portal, and to liaise with outside data organizations [27]. MODA continued to focus on its greatest legacy to NYC operations, DataBridge, the City's single repository of operational data, integrating the data in a way that makes it accessible to analysts across the City, as well as to the public through the Open Data portal [28].

4.1. Open Government Data

In NYC, new directions in terms of OGD will allow new employees to assume the processes of collection of data, service delivery portals and automated licensure processes to personalize and improve the quality of services [19]. In this case, the data are being used to solve problems nominated by the cities, helping other government employees to see the benefits of sharing what they know in digital formats, enhancing the data-drive solutions [19].

Mike Flowers as MODA's chief analytics officer and chief of open platform officer lead the effort on open data to make government data available to the people. But more than publishing open data for public use, his focus is especially in how these data can be used for internal government operations, considering that at the same time that data are being released to the public, is being released to sister agencies [19]. According to Flowers, to increase the platform and make it became more valuable, at some point agencies will have to feeding the platforms with its own data to be able to take advantage of the other agencies data [19].

4.2. Delivery of Government information and services

The NYPD is completely joining the open data movement, sharing data in more usable formats with city agencies and incorporating many other data streams to give partners more clarity concerning what's really going on in underreported crime-prone environments [19], i.e., delivering better services to the society. The use of analytic capacity crosswise agencies to solve big issues are increasing and making cities getting close to the ideal of responsive government [19].

This case clearly generates at least two public values: stewardship and quality of life. The first one enhances the public's view of government officials as faithful guardians of the value of the government in terms of public trust, integrity, and legitimacy because of the greater responsiveness of government [16]. The second one impacts individual and household health, security, satisfaction, and general well-being [16].

There are many ways that cities are using data to create a better city and delivery better services to the citizens, especially in NYC. According to The New York City [29], the MODA, the Department of Information Technology and Telecommunications (DOITT), and NYC Digital are working together to collect, analyze, and share NYC Data, to create a better City supported by data-based decision making, and to promote public use of City data. Among other things, they are collecting data to measure government performance, creating a comprehensive City-wide data platform that serves as a record of City activity, and a foundation for NYC Open Data and using the DataBridge platform to reduce safety risk in the City, deliver daily services more efficiently, and enforce laws more effectively.

New York City government, as a pioneer in OGD, has unlocked important public information, enabling technologists to build tools that help citizens in daily activities [30]. The idea is that NYC will further expand access to services, which will enable

innovation that improves the lives of New Yorkers, and, also, increase transparency and efficiency. Among the proposals previewed in the Roadmap for the digital city in 2011, is the development of the NYC Platform, an Open Government framework featuring APIs for City data, engaging and cultivating feedback from the developer community and introduce visualization tools that make data more accessible to the public. One example of that is this data visualization which displays vehicle collisions aggregated by time of day using recently released NYPD motor vehicle collision data, which, according to Mayor de Blasio, help to achieve the goal of eliminating traffic fatalities in New York City [29]. New directions in NYC include plans to expand the use of data in several areas, from policing to reduce traffic fatalities [19].

4.3. Results of the Delivery of Government information and services through OGD

The NYPD and other police departments around the world are embracing the use of data to, among other things, use maps to track crimes by neighborhoods [19]. The second step is apply predictive analytics (the use of past data to forecast future developments) to solve or prevent future urban issues. Also, the use of data-driving policing could help to deal with real-time concerns of crime or safety, make employees work smarter, allowing a better alignment between mission and resources, including community support in the fight with crime [19].

NYC Open Data makes the wealth of public data generated by various New York City agencies and other City organizations available for public use. This catalog offers access to a repository of government-produced, machine-readable data sets which anyone can use them to participate in and improve government by conducting research and analysis or creating applications, thereby gaining a better understanding of the services provided by City agencies and improving the lives of citizens and the way in which government serves them [29]. The Figure 2 summarizes the preliminary results of initially analyzes based on secondary data from this case study. In future studies, besides improving the data analyses we also intend to bring to light primarily data from stakeholders' interviews.

Figure 2. Preliminary results

Dimension	Variable	Findings	Indicators
Open Government Data	Citizens	As more information is provided through OGD citizens will find information about government activities more easily and complain and communicate their opinion, representing an increase in their interest in actively participate in the government process	Access to information Communication Participation
	Government Employees	Empower government employees to work seeking to promote a better live for citizens	Tool that help making smarter and earlier decisions
	Goverment Agencies	Local, national and international public and private organizations are publishing their data in open data infrastructures	Data can be used for internal government operations
Delivery of Government information and services	Use of OGD	Seven types of value that capture the range of possible results of government actions: economic, political, social, strategic, quality of life, ideological, stewardship	NYPD case resulted values: Stewardship Quality of life

Dimension	Variable	Findings	Indicators
	Feedback	Engaging and cultivating feedback from the developer community and introduce visualization tools that make data more accessible to the public	The NYPD is completely joining the open data movement, sharing data in more usable formats with city agencies
Results on government agencies and employees	Expected consequences	Are useful to help city governments to become more responsive Can be spread for cities around the world, allowing governments to learn from one another Can help to solve basic urban issues	Greater responsiveness of government Reduce safety risk in the city Deliver daily services more efficiently Enforce laws more effectively
	Unintended consequences	Misuse of data by government employees (privacy)	Capacity to use data with good judgment Preserve primary human values Maintain human dignity and autonomy
Results on citizens' lives	Expected consequences	Context-use-access perspective Technology has dramatically altered the way which government and citizens relate	Help in daily activities Increase: transparency, participation and Collaboration
	Unintended consequences	Misuse of data by citizens (criminals or terrorists)	Maintain human dignity and autonomy

Source. [19, 16, 29].

5. Concluding Remarks

The purpose of this exploratory study is to capture, describe, document and conceptualize the open government data phenomenon as well as explain how government are using OGD to enhance the access and delivery of information and services and impact on citizens' lives. Considering the emergent nature of the field, we start our study with the identification of definitions and relevant concepts and preliminary analysis of the case. Then, based on e-government and OGD literature we propose a conceptual model to understand how open government data platforms are being used to enhance the access to and delivery of government information and services in order to make a real difference in people's lives. We analyze secondary data and show preliminary results, based on the NYC Open Data portal, to understand how citizens, employees and other agencies are using OGD to improve their activities. Also, we discuss some of the public values resulted from e-government projects, as well as unintended consequences of that, especially the misuse of data, considering that agencies does not know what people are doing with open data or can do in the future.

For future research we must explore deeply practical implications of the conceptual model. To that end, at the beginning, considering the richness of the case study, we will continue studying the case, specially including primary data from stakeholders by making deep interviews that could allow us understand in a global view the NYC Data portal and other city agencies. Also, we aim to analyze the conceptual model in different contexts, comprising cities around the world which presents innovative cases of OGD.

Acknowledgements

References

[1]   Luna-Reyes, Luis Felipe, J. Ramon Gil-Garcia, and Georgina Romero. "Towards a multidimensional model for evaluating electronic government: Proposing a more comprehensive and integrative perspective." *Government Information Quarterly* **29.3** (2012): 324-334.

[2]   Faik, Isam, and Geoff Walsham. "Modernisation through ICTs: towards a network ontology of technological change*." Information Systems Journal* **23.4** (2013): 351-370.

[3]   Halchin, L. Elaine. "Electronic government: Government capability and terrorist resource." *Government Information Quarterly* **21.4** (2004): 406-419.

[4]   Irani, Zahir, Peter ED Love, and Ali Montazemi. "E-government: past, present and future." *European Journal of Information Systems* **16.2** (2007): 103.

[5]   Yildiz, Mete. "E-government research: Reviewing the literature, limitations, and ways forward." *Government Information Quarterly* **24.3** (2007): 646-665.

[6]   Helbig, Natalie, J. Ramón Gil-García, and Enrico Ferro. "Understanding the complexity of electronic government: Implications from the digital divide literature." *Government Information Quarterly* **26.1** (2009): 89-97.

[7]   Irani, Zahir, Tony Elliman, and Paul Jackson. "Electronic transformation of government in the UK: a research agenda." *European Journal of Information Systems* **16.4** (2007): 327-335.

[8]   Gil-Garcia, J. Ramon, InduShobha Chengalur-Smith, and Peter Duchessi. "Collaborative e-Government: impediments and benefits of information-sharing projects in the public sector." *European Journal of Information Systems* **16.2** (2007): 121-133.

[9]   Verdegem, Pieter, and Gino Verleye. "User-centered E-Government in practice: A comprehensive model for measuring user satisfaction." *Government Information Quarterly* **26.3** (2009): 487-497.

[10]   Janssen, Marijn, Soon Ae Chun, and J. Ramon Gil-Garcia. "Building the next generation of digital government infrastructures." *Government Information Quarterly* **26.2** (2009): 233-237.

[11]   Open Knowledge Foundation (OKF). *Open Data Handbook* version 1.0.0 (revised November 14, 2012), (2012). Available at http://opendatahandbook.org/pdf/OpenDataHandbook.pdf

[12]   Digital Government. *Digital Government: Building A 21St Century Platform to Better Serve The American People*. (2012). Available in: https://www.whitehouse.gov/sites/default/files/omb/egov/digital-government/digital-government-strategy.pdf

[13]   Hecth, B. *Big Data Gets Personal in U.S. Cities*. New Cities Foundation (2014).

[14]   Janssen, Marijn, Yannis Charalabidis, and Anneke Zuiderwijk. "Benefits, adoption barriers and myths of open data and open government." *Information Systems Management* 29.4 (2012): 258-268.

[15]   Jetzek, Thorhildur; Avital, Michel; and Bjørn-Andersen, Niels, "The Generative Mechanisms of Open Government Data" (2013). *ECIS 2013 Completed Research*. Paper 156.

[16]   Harrison, T., Pardo, T., Cresswell, A., & Cook, M. *Delivering Public Value through Open Government*. Center for Technology in Government (2011).

[17]   Grimsley, Mike, and Anthony Meehan. "e-Government information systems: Evaluation-led design for public value and client trust." *European Journal of Information Systems* 16.2 (2007): 134-148.

[18]   Huntgeburth, Jan C. and Veit, Daniel J., "A Research Agenda For Evaluating Open Government Initiatives" (2013). *ECIS 2013 Completed Research*. Paper 112.

[19]   Goldsmith, Stephen, and Susan Crawford. *The Responsive City: Engaging Communities Through Data-smart Governance*. John Wiley & Sons, 2014.

[20]   Agrawal, Deepti, William Kettinger, and Chen Zhang. "The Openness Challenge: Why Some Cities Take It On and Others Don't." *Twentieth Americas Conference on Information Systems*, Savannah, 2014.

[21]   Rowley, Jennifer. "e-Government stakeholders—Who are they and what do they want*?." International journal of Information management* 31.1 (2011): 53-62.

[22]   UNDP. *The 2013 human development report*–"The rise of the south: human progress in a diverse world", (2013).

[23]   Zuiderwijk, Anneke, et al. "Special Issue on Innovation through Open Data: Guest Editors' Introduction." *Journal of theoretical and applied electronic commerce research* 9.2 (2014): i-xiii.

[24]   Huang, Zhenyu. "A comprehensive analysis of US counties'e-Government portals: development status and functionalities." *European Journal of Information Systems* 16.2 (2007): 149-164.

[25]   Heinze, Nathan, and Qing Hu. "e-Government research: a review via the lens of structuration theory." *PACIS 2005 Proceedings* (2005): 75.

[26]   Korte, T; Misra, J. *Big data can be a big help to city*. CRAIN'S, New York Business (2014). Available in: http://www.crainsnewyork.com/article/20140413/OPINION/140419954/-op-ed-big-data-can-be-a-big-help-to-city#article_tab.

[27]   Goldsmith, S. *Digital Transformation*: Wiring the Responsive City. Civic Report 87, 2014.

[28]   Bloomberg, M. R.; Flowers, M. *NYC by the Numbers. Annual Report – 2013*. NYC – Mayor's Office of Data Analytics, 2013.

[29]   The City of New York (2014). *NYC Open Data Portal*. Available in: https://data.cityofnewyork.us/.

[30]   The City of New York (2011). Road Map for the Digital City: Achieving New York City's Digital Future.

Fusepool P3: A Linked Data Platform for Open Government Data

Luigi SELMI[a,1] and Alessia NEURONI[a]

[a]*Bern University of Applied Sciences*

Abstract. Many public administrations have been publishing data sets as open data for the last few years through portals based on CKAN or other platforms. Applications have been developed using, most of the time, one single data set as developers, data journalists and small businesses cannot afford the cost of re-using and integrating different data sets. This is due to many factors as the different formats used for the data, lack of documentation, different metadata and lack of public, trustworthy registries for entities of interest that could ease the task of connecting information provided by different authors about them. The Linked Data principles offer the guidelines to solve all these issues by leveraging the Semantic Web standards for describing resources on the Web. Moreover the Linked Data Platform specification, a recent W3C Recommendation, provides a realization of the guidelines defining some basic interactions between a client and a server to manage resources using the HTTP protocol. The purpose of this article is to present the Fusepool P3 platform that extends the LDP specification to support services which transform raw data sets into RDF format and enrich them.

Keywords. Linked Data, Linked Data Platform, Data Integration, Application Integration

1. Introduction

In the last years there has been an increase in the awareness of the importance of citizens' engagement and participation in society. Governments have been often criticized for being opaque and the mainstream media are seen by many mostly at the service of interested parties. Participation and engagement must be based on information and above all information produced by public administrations. Furthermore, opening the public sector information assets could result in economic gains estimated up to €40 billion a year in the EU [1]. Citizens, SMEs, tax payers accustomed to use the Web to communicate, buy stuff, learn, share information, do business, want to have access to the facts not just opinions and have boosted the request for the administrations to provide their information and services online. A strong impulse to publish raw data sets as they were produced by the public administrations has come from the European Commission directive of 2003 on the re-use of public sector information. The subsequent request for open data was raised to address two main issues: the formats in which the data sets were provided and the licenses for re-use. Clearly open, non-proprietary formats like CSV or XML are easier

---

[1] Corresponding Author: Senior Researcher, Bern University of Applied Sciences; E-mail: luigi.selmi@bfh.ch

to extract information than Excel spreadsheets or pdf files but cannot help in overcoming the semantic interoperability issue that arises when trying to merge different data sets. A solution to this last issue is offered by the Semantic Web standards in particular when they are used with the Linked Data principles as guidelines. Briefly, the proposal of the Semantic Web is to give any entity of interest a global identifier and a description using the RDF data model and terms from shared vocabularies. The Linked Data principles add some clarification such as the use of HTTP URI as global identifier and to provide links to other entities. Opening the data is the first step to a data driven society that requires huge changes in workflows and roles within each public administration and a commitment to deliver useful and reliable information to the citizens. Taking the next step towards the Semantic Web vision requires a stronger commitment to provide, maintain and make available name registries of entities of public interest using shared vocabularies. The idea is that an investment in such direction will provide further indirect returns to governments.

This paper discusses the issues that have to be addressed to merge the growing number of open data sets delivered by public administrations at different level and presents the Fusepool P3[2] platform that is being developed for publishing Linked Data from raw data sets. The paper is organized as follows. In the next Section 2 a short description of the Linked Data principles is introduced. Section 3 describes the Linked Data Platform, a formal specification of the principles. Section 4 describes different stages of a resource in the process of being transformed from raw to RDF format and how the Fusepool P3 platform enables the data transformation process and the publication of the result as Linked Data. In Section 5 are described some of the platform components that are being developed to transform and enrich the data sets. Last, in Section 6, the conclusions and future work.

2. Linked Data Principles

Linked Data [2] is a proposal published by Tim Berners-Lee in 2006 as a new way to publish data on the Web. It is based on four rules

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)
4. Include links to other URIs so that they can discover more things.

The Linked Data principles aimed at providing some implementation guidelines to the realization of the Semantic Web. URI and RDF, the two main standards of the Semantic Web used to name and describe distributed resources, are quite abstract and do not enforce developers to use any other Web standard such as the HTTP protocol to access and use those resources. In the Semantic Web a URI is a mean to name an object, real or abstract, but it does not mandate for a representation of the object to be retrievable as a resource on the Web, not even when it is a HTTP URI. An application or software agent cannot rely on the availability of a description on the Web of any of the component of a RDF triple. The Linked Data principles do enforce the use of

---

[2] http://p3.fusepool.eu/

dereferenceable HTTP URIs, namely URLs, as name for things and the provision of information in RDF format when dereferencing any URI named thing. The guidelines have been successful as many research projects have been started from them. One of the main projects is LOD2[3] in which a complete set of tools have been integrated in the LOD2 Stack distribution to support the whole life cycle of linked data sets from data modeling to semantic lifting, interlinking, information extraction and data storage. Many RDF data sets, that have been published as Linked Data and interlinked, form the Linked Open Data Cloud[4], a growing RDF graph of interconnected data sets available online.

3. Linked Data Platform

As the Linked Data principles promote integration of data sets at the Web scale through standards such as dereferenceable HTTP URI, RDF, SPARQL and others, they do not provide a formal specification on how a client and a server should communicate in order to create a resource, or organize collections of resources. Most of the websites that publish data as Linked Data provide only read access to resources through a SPARQL endpoint or RDF data dumps following best practices [3] and patterns [4]. No write access is usually provided. The main effort of many research projects was to produce a reasonable amount of RDF data to bootstrap the Semantic Web. Once the data sets were published on a SPARQL 1.0 endpoint or in form of zipped RDF files, the job was mostly done. In 2009 Tim Berners-Lee wrote a follow-up [5] to the Linked data principles focusing on write access to web resources. HTTP, WebDAV and SPARQL Update protocols were foreseen as good candidates to support write access to Linked Data. The first attempt to implement a full read-write platform based on the Linked Data principles has been done by IBM as they were investigating the Linked Data approach for integrating their requirements engineering tool IBM Rational with other development tools. IBM then joined the W3C to work on a specification. The Linked Data Platform specification [6] has become very recently a W3C recommendation. A Linked Data Platform can handle all type of resources such as plain text, semi-structured text and multimedia like images, video or audio. A resource is named by a HTTP URI and can be a RDF source or any other type of resource, namely a Non-RDF source. A LDP container is a RDF source into which other resources can be put. A LDP container can be further specialized as basic, direct or indirect. A basic container has a containment relationship with its resources while the relationship between a direct or indirect container with its resources can be further specified. An HTTP server that implements all the mandatory requirements listed in the LDP specification becomes a LDP platform. A client can look up a representation of a resource in a LDP platform by simply sending an HTTP GET request to its URI. A client can create a LDP resource sending an HTTP POST request to a LDP container with a representation of the resource. If the creation of the resource is successful the server will send a response message with the "Location" header containing the URI of the created resource. A client can create a new container sending an HTTP POST request to a container with a RDF description of the type of container. A client can also retrieve a list of all the resources within a container. An LDP container can be seen as a

---

[3] http://lod2.eu/

[4] http://lod-cloud.net/

folder in a file system that can be accessed to read and write LDP resources from the Web.

4. Fusepool P3: An Extended Linked Data Platform Implementation

The European FP7 project Fusepool P3 aims at extending the LDP specification to address the needs of open data publishers and consumers. It adds to LDPs a new type of container that can be configured to host the result of a data transformation processes.

Data owners and publishers that want to get four or five stars of the Tim Berners-Lee's deployment scheme[5] for their data sets must send them through one of more transformation steps such as semantic lifting, reasoning, interlinking and information extraction. Starting from a Non-RDF resource the data can be transformed into RDF through a process called semantic lifting. Creating a RDF representation of a Non-RDF resource implies putting in an explicit form the semantics of the resource metadata. This task is performed designing a model of the resource based on ontologies and defining the rules to map the resource metadata to the terms in the ontologies. The mapping rules can be defined for example by XSL transformations or other mapping languages. During this process new entities are given a URI following a pattern defined in the XSL stylesheet. When more data sets are transformed in RDF, terms coming from different vocabularies can be used that define the same meaning or there might be an interest in inferring new relations among the entities described in different data sets. As an example, in a data set some entities can be described as being hotels, according to an ontology like schema.org[6], while in another data set other entities can be described as restaurants and a user might want to know all the local businesses in both data sets. A local business can be defined in the same ontology as an upper class of both hotel and restaurant but an application cannot answer the user query as the fact that each hotel or restaurant is also a local business has not yet been made explicit [7]. Most of the time these inferences are not stored in a database as they can be made at runtime by a reasoner on behalf of a user agent. There might be cases in which different URIs are used in a data set to denote the same entity. Finding and replacing such duplicates is a well-known issue in any data integration project [8]. The task is performed comparing descriptions of the entities through functions that return a numeric value depending on the similarity of the descriptions. When the returned value is higher than an upper threshold, the entities are automatically denoted as equal and a triple is added to the RDF data set to state the inferred fact between the two entities, often using the *owl:sameAs* relation from the OWL ontology. When the returned value falls below a lower threshold, the entities are different and no action is performed. In a third case, when the returned value falls between the thresholds, a human intervention can be required to state whether two entities are equivalent or not. The same task can be performed to interlinking entities described in different data sets in order to merge their descriptions. Further enrichments can be added to a data set extracting information from plain text provided in the description of a resource through NLP processes or enabling manual text annotations by experts.

In Fusepool P3 a transformation process can be performed by a single component or a chain of components. A component is defined as a "transformer" in the project

---

[5] http://5stardata.info/
[6] http://schema.org/

documentation and must implement a REST API. The input data can be sent via a HTTP POST request to be transformed or enriched. The input and output formats supported by a transformer can be retrieved sending a HTTP GET request to it. A transformer can take the address of another resource as a parameter to be used as rules to adapt its behavior to the input data. As an example, a transformer has been made available to transform data from XML to RDF. The transformer needs a XSL stylesheet to perform a transformation from a specific XML file into RDF using terms from one or more vocabularies. Different transformers of the same type can be instantiated through a factory changing the parameter value sent in the request, namely the URL of the XSL stylesheet. The architecture based on the REST style and HTTP protocol enables two or more transformers to be used in chain. A pipeline transformer can be configured as an ordered list of transformers that will send the data through each of them. A basic check of the compatibility between the output format of a transformer and the input format of the following one is performed before a pipeline can be instantiated by its factory. A transformer responds synchronously to all requests. The response message can contain the result of the transformation if the transformer has been configured to process the input data synchronously or it contains the URI of the resource that will be created as a result of the transformation if the transformer has been configured to process the data asynchronously. The Fusepool P3 platform has been tested with two LDP implementations, Apache Marmotta [7] and OpenLink Virtuoso [8] which both provide also a triple store for persistence of the RDF data and a SPARQL endpoint. The Fusepool P3 platform adds to the LDPs a proxy to which a client can send a HTTP POST request to create a transforming container. As an example in the tourism domain we considered an application developer that wanted to transform a data set about cultural events in RDF in order to merge the result with information about points of interest in the same area. The data is provided by the local tourism office on a daily basis as an XML file at the URL http://wonderland.com/events.xml. A transformer that is able to interpret XSL stylesheets can be instantiated from its factory and made available at the URI http://sandbox.fusepool.info:8164 pointing to a XSL stylesheet at the URL http://tourismtoday.com/wonderevents.xsl. The full URI of the transformer instance will use the encoded URL of the stylesheet

http://sandbox.fusepool.info:8164/?xslt=http%3A%2F%2Ftourismtoday.com%2Fwonderevents.xsl

A client can send an HTTP POST request with the XML document to the transformer to get the result of the transformation or it can create a transforming container using the URI of the transformer instance. A transforming container can be created sending to the LDP proxy a message with a description of the container in which is stated the URI of the transformer that will be used whenever a client will send some data to it. The description of the transforming container must be provided as RDF

---

[7] http://marmotta.apache.org/

[8] http://virtuoso.openlinksw.com/

```
@prefix ldp: <http://www.w3.org/ns/ldp#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix dcterms: <http://purl.org/dc/terms/> .
@prefix eldp: <http://vocab.fusepool.info/eldp#> .

<> dcterms:title "Wonderland Events Transformer"@en ;
    dcterms:description "Wonderland Events Transformer Container" ;
    a ldp:DirectContainer ;
        eldp:transformer
<http://sandbox.fusepool.info:8164/?xslt=http%3A%2F%2Ftourismtoday.com%2Fwon
derevents.xsl>
```

A new transforming container with relative path /ldp/wonder/ can be created in the root platform container /ldp/ or in other containers. After creating a file "mycontainer.ttl" with the description of the container a client can send the request to create it to the LDP proxy whose address is http://sandbox.fusepool.info:8181

curl -i -X POST -H "Content-Type: text/turtle" -H "Slug: wonder" -d @mycontainer.ttl
http://sandbox.fusepool.info:8181/ldp/

If the request is successful the proxy will send a response message with the URI of the created container in the "Location" header. Once the transforming container has been created a client can send to it a request with the XML data to transform

*curl -i -X POST -H "Content-Type: application/xml" -d @events.xml *
http://sandbox.fusepool.info:8181/ldp/wonder/

The LDP proxy will send the data to the transformer that was linked in the container description. The proxy will create two new LDP resources in the container, a Non-RDF source with the input data and a RDF source containing the result of the transformation. The application developer might also be interested in enriching her RDF data linking the entities to those in DBpedia [9], a multilingual RDF knowledge base extracted from Wikipedia, or she might want to extract the entities named in textual parts and link them to the event. The developer will have to create an instance of the pipeline transformer with the list of the transformers' instances needed for the tasks and then create a new transforming container with a reference to the pipeline URI in the description.

5. Fusepool P3 transformers and administration tools

The Fusepool P3 platform provides a growing number of transformers, based on tools developed in open source projects or online web services. Some transformers that are being developed are shown in Table 1. They can be used in different phases of a transformation pipeline: semantic lifting, interlinking, information extraction, text annotation.

Table 1. Fusepool P3 transformers

Transformer	Scope
P3-BatchRefine	Semantic lifting
P3-Xslt	Semantic lifting
P3-DataTxt	Information extraction
P3-DictionaryMatcher	Information extraction
P3-Silkdedup	Interlinking
P3-Pundit	Annotation

The P3-BatchRefine transformer is based on OpenRefine[9] and its RDF extension. It can be used to transform tabular data in RDF. A JSON resource file with transformation rules specific for the input data must be provided as a parameter to the transformer. The P3-Xslt transformer can be used to transform XML files into RDF format as discussed in Section 4. The P3-DataTxt transformer has been implemented as an Apache Stanbol enhancer and provides a REST API like any other NLP engines available in Apache Stanbol[10]. It is based on an online web service and can be used to extract named entities from text. The P3-DictionaryMatcher extracts named entities in text that match with concepts in a SKOS taxonomy. The P3-Silkdedup transformer is based on SILK [10] and it can be used in deduplication or interlinking tasks. A configuration file with the linkage rules to compare the descriptions of the entities in the source data set and in the target one must be provided as a parameter. The P3-Pundit transformer supports a user interaction mechanism to assist experts in manually annotating text. The Fusepool P3 platform integrates some of the tools available in the LOD2 stack as distributed components that can communicate through REST API without requiring the installation of any of them in the same machine. Other components can be developed in Java or any other programming language, and added in a pipeline to easily implement specific use cases. As an example, many data sets contain geographic information for which further transformations can be required such as coordinates transformation and geocoding. User interfaces are being developed to easily register transformers, create pipelines and import data sets to be transformed and retrieved via LDP protocol or SPARQL.

6. Conclusions and Future Work

The Linked Data Platform specification is a new W3C recommendation that formalize the Linked Data guidelines with which a client can create, read, update and organize resources on a server using the HTTP protocol and standards like HTTP URI to name them and the RDF data model to provide a representation of the resource and its relationship with other resources on the Web. The Fusepool P3 platform aims at adding to it a mechanism to help open data providers and consumers to address the semantic interoperability issue and boost the re-use of open data. The distributed nature of the platform enables different players to collaborate. Developers and SMEs can provide software components that offer services to be integrated in the pipeline of a client. Data analysts and data journalist can set up pipelines to extract information hidden in raw, fragmented data sets. The future work in the development of the Fusepool P3 platform will be focused on the implementation of an authentication service based on WebID

---

[9] http://openrefine.org/
[10] https://stanbol.apache.org/

and the integration with the CKAN platform. The platform documentation [11][12] and software is available on the Github repository[11].

Acknowledgements

References

[1] European Commission, Open Data, An engine for innovation, growth and transparent governance, 2011.
[2] Berners-Lee, T., Linked Data, http://www.w3.org/DesignIssues/LinkedData.html, 2006.
[3] Heath, T., Bizer, C., Linked Data, Evolving the Web into a Global Data Space, Morgan&Claypool Publishers, 2011.
[4] Dodds, L., Davis, I., Linked Data Patterns, http://patterns.dataincubator.org/book/, 2012.
[5] Berners-Lee, T., Read-Write Linked Data, http://www.w3.org/DesignIssues/ReadWriteLinkedData.html, 2009.
[6] Linked Data Platform Working Group, Linked Data Platform 1.0, http://www.w3.org/TR/ldp/, 2015.
[7] Allemang, D., Hendler, J., Semantic Web for the Working Ontologist, 2nd Ed., Morgan Kaufmann, 2011.
[8] Fellegi, I.,P., Sunter, A.,B., A Theory for Record Linkage, Journal of the American Statistical Association, 1969.
[9] Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontakostas, D., Mendes, P., Hellmann, S., Morsey, M., van Kleef, P., Auer, S., Bizer, C., DBpedia – A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia, 2012.
[10] Volz, J., Bizer, C., Gaedke, M., Kobilarov, G., Silk – A Link Discovery Framework for the Web of Data, 2009.
[11] Gmür, R., The Fusepool P3 Platform, http://fusepool.gitbooks.io/the_fusepool_p3_platform/
[12] Fernandez, S., Frank, J., Westenthaler, R., Blakeley, C. Idehen, K., Gmür, R., Gschwend, A., Fusepool P3 Deliverable 5.1, Technical platform specification and basic implementation.

---

[11] https://github.com/fusepoolP3

Characteristics of business intelligence and big data in e-government: Preliminary findings

Rikke GAARDBOE[a], Tanja SVARRE[a,1] and Anne Marie KANSTRUP[a]

[a] *eLearning Lab, Department of Communication and Psychology, Aalborg University, Rendsburggade 14, DK-9000 Aalborg, Denmark*

Abstract. Business intelligence and big data represent two different technologies within decision support systems. The present paper concerns the two concepts within the context of e-government. Thus, the purpose of the paper is to present the preliminary findings regarding publication patterns and topic coverage within the two technologies by conducting a comparative literature review. A total of 281 papers published in the years 2005–2014 were included in the analysis. A rapid increase of papers regarding big data were identified, the majority being journal papers. As regards business intelligence, researchers publish in conference proceedings to a greater extent. Further, big data journal papers are published within a broader range of journal topics compared to business intelligence journal papers. The paper concludes by pointing to further analyses that will be carried out within the 281 selected papers.

Keywords. Big Data, Business Intelligence, E-government, Literature review

1. Introduction

According to Gartner [1], business intelligence/analytics is the top-priority issue for CIOs and investments. The public sector could increase the effectivity of administration by using big data. An estimate is 250 billion EUR each year, which is equivalent to 0,5% annual productivity growth in the European Union [2]. The two terms applied to describe the software for business analytics are 'business intelligence' and 'big data'. Business intelligence and big data are used for supporting online political participation, e-government service delivery, process transparency and accountability, and is based on opinion mining, social network analysis and data from the accounting system [3]. The term 'decision support system' originated 50 years ago and covers computer-based tools for sense-making and decision-making. The terms used have changed over time, and among others, the term 'business intelligence' is now more commonly used. However, although new terms appear, they cover the same purpose[4]. Some researchers consider big data business intelligence 3.0 [3] as a new generation of business intelligence and analysis, while other researchers consider it as a paradigm shift [5].

The two technologies share the purpose of delivering decision support in a changing world. However, despite terminological overlaps and shared ambitions about

---

[1] Corresponding Author.

decision support, business intelligence and big data are not identical. Pedersen [6] has identified seven differences between business intelligence and big data which are included in Table 1.

Table 1 Differences between BI and Big Data, based on Pedersen [6]

	Business Intelligence	**Big Data**
Data types	Structured (Mostly)	Unstructured (a.o)
Data sources	Mostly internal	Mostly external
History	Essential	(Often) less relevant
Users	Manager/controller	Data scientist
Precision	Exact results	Approximate results
Privacy	Not critical	Critical
Control over data	Almost full control	Little or no control

The main difference between business intelligence and big data is the data source, as it is crucial to further architecture. Business intelligence uses data structured in relational form, preferably from the organization's internal systems; this means that there is full control of data and that data are structured and validated. In contrast, big data applications often use external data, where there is no control [6]. Consequently, the potential of the two types of technologies in the public sector varies.

Olszak [7] emphasized effectiveness as an important potential of business intelligence in the public sector, especially with respect to the planning and process improvements in an organization. Manyika et al. [2] highlighted the potential of big data, emphasizing the ability to combine structured and unstructured data. In the public sector. That has the following potentials: creating transparency with accessible related data; discovering needs; improving performance; customizing actions for suitable products and services; decision making with automated systems to decrease risks and innovating new products and services. Furthermore, in the health care clinical decision support systems, individual analytics can be applied for patient profiles, personalizing medicine, performance-based pricing for personnel, analyzing disease patterns and improving public health [8].

The idea for this article arose from a PhD study about business intelligence in the public sector. Terminological confusion and important differences between the two types of technologies require researchers to study the literature on business intelligence and big data in e-government. The purpose of the present review is to compare the extent, growth and nature of the technologies in government research. Subsequently, through analysis, we strive to understand and clarify the differences and overlaps between the technologies in research.

This paper presents the preliminary results from our literature study on business intelligence and big data in e-government. These initial results indicate trends and implications for literature studies within these areas. As with all literature reviews, we aim to analyse a given topic and identify research trends and potential gaps that can lead to new studies [9]. Below we present the first steps in this direction.

2. Method

The present paper builds upon a comparative, systematic literature review. Taking this approach, the study is based on informetric methods to quantitatively measure and compare the development of the two subject fields of business intelligence and big data in an e-government context [10], [11]. A common approach in literature reviews is to

limit the sources of papers to core journals within the field in question. However, to be able to test whether research papers on the two technologies in question are published in a variety of channels, we are not limiting the sources of potentially relevant papers to specific journals or conferences. Instead, in collecting relevant references and delimiting the subject field, we searched the following three databases covering computer science (journals, monographs *and* conferences): Web of Science (ISI), Scopus (Elsevier) and ABI/INFORM Complete (Proquest). Another reason for searching these databases was their advanced search interfaces, enabling complex queries for retrieval of documents.

In each database two queries were defined: one for business intelligence and one for big data. To be included in the pool of relevant papers, papers needed to be peer reviewed and written in English. The e-government domain was represented as follows:

(municip* OR govern* OR council* OR ministr* OR "public administration" OR "public sector" OR egovern* OR e-govern*)

The purpose of this representation of the e-government query component was to ensure a broad representation of the domain in question, and to enable the inclusion of papers concerned with government in general. The underlying assumption was that the majority of governments in the Western world will have some level of digitalization within the specific period under investigation. Lastly, the year of publication was defined to include papers within the time window 2005–2014. In sum, the queries carried out consisted of the following search components:

SUBJECT ((BI OR BD) AND e-government) AND DOCUMENT TYPE AND PUBLICATION YEAR AND LANGUAGE

Each query was adjusted to the specific command language of the three databases searched. Subsequently, the retrieved documents were entered into an access database, where duplicates were removed; as a result of this process, the access database now contains 99 papers on business intelligence and 182 papers on big data. No overlap was found between the two pools of papers. The variables characterizing the reviewed papers comprise various bibliographical data, abstracts, Dewey classification codes (for journal papers), and for some papers, full text versions and/or cited references. In the analysis below, we report on and compare the full collection with respect to publication years and document types. In addition, we have added Ulrichsweb's (https://ulrichsweb-serialssolutions-com) Dewey classification codes for journals to the part of our papers covering journal papers. The purpose of adding the Dewey codes where possible is to investigate the extent of topics covered by business intelligence and big data, respectively. However, conferences are not covered by Ulrichsweb, meaning that this particular analysis will be carried out solely for journal papers.

3. Preliminary findings

In this section we present the preliminary findings of the comparative review. The results are presented in two sections; in the first section, the results are analyzed for the whole population based on the 281 collected articles. The second part of the results specifically reports findings regarding journal papers (cf. Section 2 on Dewey codes).

Table 2 Number of e-government research publications concerning Business Intelligence and Big Data in the period of 2005-2014

Year	Business Intelligence	Big Data
2005	3	-
2006	1	-
2007	6	-
2008	12	1
2009	10	2
2010	6	-
2011	17	1
2012	19	19
2013	11	46
2014	14	113
Total	99	182

Within e-government business intelligence, there is a general increase of publications from 2007 to 2012 (see Table 2). After 2012, the number decreased again. The trend in big data publications differs, as the number of publications has grown exponentially since 2011 (see Table 2). From 2013 to 2014, the number of publications more than doubled. Thus, in the period business intelligence represent a minor, but permanent interest within research, whereas big data denotes a newcomer with an increasing focus. With the introduction of NoSQL databases around 2009, an increase in performance and flexibility in handling big data appeared [12]. INMemery technology was also launched for commercial use that year; for example, SAP launched their system HANA, which is an INmemery data management system that improves the performance of analytical and transactional applications [13]. Both technologies support big data by enabling processing of large data volumes. The introduction of new technologies may partially explain the significant increase in the number of research publications on big data and e-government. However, further analyses are needed to fully understand the nature of the growth.

Table 3 Publication types applied for research dissemination by researchers

Publication type	Business Intelligence	Big Data
Conference proceedings	58 (59%)	32 (18%)
Generic	1 (1%)	-
Journal article	25 (25%)	139 (76%)
Monograph	15 (15%)	11 (6%)
Total	99	182

Another interesting finding is the type of communication channels chosen by the authors. Table 3 shows the distribution of publication types across the two fields. As illustrated in the table, business intelligence has the largest share of conference papers (59% of all publications), whereas big data has an even larger share of journal papers (76% of all publications). As previously stated, conferences tend to publish earlier research results, whereas scientific journals publish more completed work [14]–[16]. Considering the publication patterns of the two fields, it is surprising that the youngest of these fields has the largest share of journal papers. An expected distribution would be a larger share of conference papers for the younger field and a larger share of journal papers within the older, more established field of research. We will look into possible causes in additional analyses.

Table 4 summarizes the subject distribution of the journals chosen for publication by the authors. For both technologies, the three most frequent categories comprise Computer Science, Information and General Works, Social Sciences and Technology.

Papers concerning big data and e-government are also published in the categories of Science, Arts and Recreation, History and Geography, Language and Philosophy, and Psychology.

Table 4 Dewey main classification of journals (as added by Ulrichsweb) containing the journal papers of the collected papers

Dewey main classification	Business Intelligence	Big Data
Arts & recreation	-	1 (1%)
Computer science, information & general works	10 (40%)	41(29%)
History & Geography	-	2 (1%)
Language	-	1 (1%)
Philosophy & Psychology	-	2 (1%)
Science	-	9 (6%)
Social Sciences	5 (20%)	43(31%)
Technology	10 (40%)	40(29%)
Total	25 (100%)	139 (100%)

Thus, despite overlaps in the most frequent categories, big data research apparently has a more diverse scope of topics compared to business intelligence. The more varied distribution of big data may partially be explained by a considerably larger number of journal papers written about big data (139) compared to business intelligence (25 papers). For the e-government research field, this means that broader perspectives are needed in terms of subject coverage to ensure a sufficient reflection of the field. However, it is nonetheless an interesting finding that we will examine more thoroughly in a follow-up study.

4. Future analysis and perspectives

Above we have presented the preliminary results of the distribution of papers within the two research fields under investigation. We now have an indication of the distribution of publications with respect to publication year, type, volume and journal subject. However, in some cases, we have been puzzled by the distribution. In their review typology, Paré et al. [17] identified various ways of composing a review. To explain the distributions reported in the present paper, we plan to carry out a follow-up study. The follow-up study should be designed as a scoping review based on the comprehensive search strategy also applied here. In a scoping review, the methods applied for analysis include content and thematic analysis. To conduct these analyses, we intend to examine different aspects of the papers.

One aspect of the papers that should be investigated further is the author-supplied keywords added to the papers. As opposed to the controlled Dewey codes reported above, author keywords represent an uncontrolled way of indicating the content of a paper. Uncontrolled elements in a subject description are characterized by not having any delays in incorporating new terms and reflecting the mindset of the author [18]. As an extension of the Dewey code analysis, and to avoid being restricted to journal papers, we intend to conduct an analysis of the author keywords by means of text mining [19]. Using this technique, the most frequent keywords can be identified across authors. We will also carry out citation analyses to identify core authors, papers and journals within the fields.

In addition, we want to carry out more thorough coding of the content of the papers gathered. In the first round of coding, we will identify the use of methods in the papers by means of the typology put forward by Dwivedi [20]. Another interesting coding to perform is to identify the unit of analysis. To do this, we will apply Joseph's [21] categorization.

The method applied for collecting papers presents several benefits, including that we have not limited the scope of the two technologies to a specific journal. Therefore, a broad collection of papers have informed the review. However, we may discuss the number of years selected for the review. Regarding big data, it appears from the number and distribution of papers that we have framed the significant years of publication. However, business intelligence dates further back than 2005. This means that from a methodical perspective, the full range of business intelligence papers is not included in the present comparison of technologies. This condition should be taken into consideration in the assessment of the findings.

5. Conclusion

The publication of business intelligence articles in the context of e-government has steadily increased since 2005, whereas the publication of big data articles has grown exponentially since 2011. Despite big data being a new technology, a lot of research findings have been published as journal papers. On the contrary, the more established field of business intelligence were to a larger extend disseminated through scientific conferences. The purpose of a literature review is to identify a research area in depth, meaning that journal articles and conferences must both be involved as common procedure. Therefore, 281 publications were analyzed in this study. However, we have made several interesting findings. To be able to understand the reasons behind, we need to dig deeper into the publications and apply additional analyses in a follow up study.

References

[1]   Gartner, *Flipping to Digital Leadership: Insights from the 2015 Gartner CIO Agenda Report*, Gartner, Stamford, US, 2014.
[2]   J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers, *Big Data: The next Frontier for Innovation, Competition, and Productivity*, McKinsey Global Institute, 2011.
[3]   H. Chen, R. H. Chiang, and V. C. Storey, Business intelligence and analytics: From big data to big impact., *MIS Quarterly*, **36** (2012), 1165–1188.
[4]   S. Alter, A work system view of DSS in its fourth decade, *Decision Support Systems*, **38** (2004), 319–327.
[5]   A. McAfee and E. Brynjolfsson, Big data: The management revolution. (cover story), *Harvard Business Review*, **90 (**2012), 60–68.
[6]   T. B. Pedersen, Managing big multidimensional data, *EGC 2015. Revue des Nouvelles Technologies de l'Information*, vol. RNTI-E-28, 2015.
[7]   C. M. Olszak, Business intelligence and analytics in organizations, in *Advances in ICT for Business, Industry and Public Sector*, vol. 579, M. Mach-Król, C. M. Olszak, and T. Pełech-Pilichowski, Eds. Cham: Springer International Publishing, (89–109), 2015.
[8]   S. Sagiroglu and D. Sinanc, Big data: A review, in *Collaboration Technologies and Systems (CTS), 2013 International Conference on*, (2013), 42–47.
[9]   Y. Levy and T. J. Ellis, A systems approach to conduct an effective literature review in support of information systems research, *Informing Science: International Journal of an Emerging Transdiscipline*, **9** (2006), 181–212.

[10] J. Tague-Sutcliffe, An introduction to informetrics, *Information Processing & Management*, **28** (1992), 1–3.

[11] L. Egghe, Expansion of the field of informetrics: Origins and consequences, *Information Processing & Management*, **41** (2005), 1311–1316.

[12] M. Stonebraker, SQL databases v. NoSQL databases, *Communications of the ACM*, **53** (2010), 10-11.

[13] F. Färber, S. K. Cha, J. Primsch, C. Bornhövd, S. Sigg, and W. Lehner, SAP HANA database: data management for modern business applications, *ACM Sigmod Record*, **40** (2012), 45–51.

[14] W. D. Garvey and B. C. Griffith, Communication and information processing within scientific disciplines: Empirical findings for Psychology, *Information Storage and Retrieval*, **8** (1972), 123–136.

[15] W. D. Garvey, N. Lin, C. E. Nelson, and K. Tomita, Research studies in patterns of scientific communication: I. General description of research program, *Information Storage and Retrieval*, **8** (1972), 111–122.

[16] B.-C. Bjork, A model of scientific communication as a global distributed information system, *Information Research*, **12** (2007).

[17] G. Paré, M.-C. Trudel, M. Jaana, and S. Kitsiou, Synthesizing information systems knowledge: A typology of literature reviews, *Information & Management*, **52**, (2015), 183–199.

[18] C. P. R. Dubois, Free text versus controlled vocabulary, *Online Review*, **1 (**1987), 243–253.

[19] W. Fan, L. Wallance, S. Rich, and Z. Zhang, Tapping the power of textmining, *Communications of the ACM*, **49** (2006), 76-82.

[20] Yogesh K. Dwivedi, An analysis of e-Government research published in Transforming Government: People, Process and Policy (TGPPP), *Transforming Government*, **3** (2009), 7–15.

[21] R. C. Joseph, A structured analysis of e-government studies: Trends and opportunities, *Government Information Quarterly,* **30** (2013), 435–440.

Publication and Reuse of Linked Data: The Fusepool Publish-Process-Perform Platform for Linked Data

Adrian GSCHWEND [a] Alessia C. NEURONI [a] Thomas GEHRIG [a]
and Marco COMBETTO [b]

[a] *Bern University of Applied Sciences, Switzerland*
[b] *Informatica Trentina, Italy*

Abstract. Linked open government data can be a catalyst in the production of value-added services and products. The vision of the presented R&D-project is to make publishing and reuse of linked data as easy as possible for the end user thanks to a thriving marketplace with data publishers, developers, and consumers along the value chain. The project develops tourism-related applications and software components that support data owners and open data enthusiasts in transforming legacy data to Linked Data. It does that by providing tightly integrated components and supports the multilingual data value chain from data exploration, extraction, enrichment, and delivery. In the paper we present first results of the project and reflections on the methodology in order to increase sustainability by focusing on users-driven development.

Keywords. Linked Open Data, Linked Data Platform, RDF, extract transform load (ETL), integrated access, analytical processing, user requirements, sustainable methodology

1. Introduction

Organizations and companies are progressively adopting Linked Data for publishing knowledge as Open Data. Data sources are often internal systems that have little or no relationship to other data sources and most of the time these silos use long established technologies (relational database systems, XML or even simple solutions such as spreadsheets) to curate and maintain data.

Publishing this data as Linked Open Data increases the reusability and value of the data; but it requires a certain amount of work to capture the knowledge of the data owner and domain specialist and expressing it as Linked Data. As of today this work is cumbersome and not easy to automate. There are many different tools available that can support the data owner in this process but they are not easy to chain or automate.

The here presented platform Fusepool P3 considerably speeds up this process by streamlining the publish, process, perform workflow through a common user interface and API.

The paper is structured as follows. It starts by introducing the Fusepool P3 project and its publish, process, perform workflow in Section 2. Then, it gives an overview of the

Open Data and Linked Open Data development and related work. Additionally it outlines how the team handles the development process in Section 3. In Section 4 it explains the Open Data experience of the project stakeholders and how their experience and use cases are integrated into the development process, followed by emphasizing the case for Linked Data. Finally, it talks about lessons learned and further activities in this still ongoing project.

2. P3: Sketching the EU-Project

To improve publishing and reuse of public data, Fusepool P3 develops software components that support data owners and open data enthusiasts in transforming legacy data to Linked Data. It integrates existing and newly developed software components and provides the missing glue and user interfaces to perform the necessary tasks in a sustainable, repeatable way.

Within the Fusepool P3 project the development team interacts with, and supports the stakeholders in the following ways:

Publish: identify data sources to promote businesses
> When revealing data, the team works with data providers or third-party data developers to use structural and descriptive metadata to make the data more explicit to facilitate finding, consolidation, and interrogation. The goal is to make publishing and reuse of linked data as easy as possible for the end user thanks to a thriving eco-system with data publishers, developers, and consumers along the value chain.

Process: transform and enrich data to make it actionable
> Data becomes machine-readable so that software programs can find and consume that data without the user intervening at every step of the data refinement process. It involves natural language processing to extract entities and classifications from raw text or the mapping of structured data sources, their cleaning, and conversion into the Resource Description Framework (RDF)[2] model; then data interlinks with other data. This involves reasoning based on the available metadata associated with the data to infer possibly related data resources.

Perform: deliver data for many use cases and devices
> The main tangible outcomes of the project are a set of integrated software components within the open-source Fusepool P3 platform.

The platform refers to the actual installation of the components in a concrete instance, local or in the cloud. To ensure stakeholders get the most out of Fusepool P3, additional goals were defined:

- The platform, including all necessary components, can be easily deployed within any other appropriate software environment.
- It provides an ecosystem, which contains value-adding components such as data transformers and enrichers.
- App developers can query the data in a standardized way; the data sets are unlocked and can be accessed and queried across their boundaries.
- The Fusepool P3 community consists of the main players along the data value chain as well as current and future stakeholders.

Fusepool P3 goes beyond existing related projects because its focus is on creating a consistent and integrated software stack rather than individual applications. This enables data publishers to stream updated data within pipelines and app developers to reuse that data with minimal manual intervention.

3. Creating an Added Value: An Overview

In the past few years the term Open Data became more widely known and in many countries it triggered movements to make data from government organizations available to the public[3]. This is not necessarily new *per se*; in Switzerland the Federal Statistical Office has made its census data available since the 19th century, at first on paper but since the 1960ies also in digital form[1]. Interestingly, with the advent of the Web there was a counter development where corporations started to get exclusive licenses on publicly funded data and started to lock down access behind paywalls. The formal definition of Open Data aims at changing this by defining a common understanding of what Open Data means and by pressuring governments and organizations to liberate their data under appropriate licenses and in useful formats. According to the Open Knowledge Foundation[2], common reasons for open data include transparency, releasing social and commercial value as well as promoting participation and engagement.

3.1. (Re-) Using Linked Data

From a technical perspective, many government organizations currently publish Open Data on portals such as CKAN[3]. CKAN is a data management system aimed at data publishers wishing to make their data open and available. It provides tools to facilitate this publishing step and helps to locate and use data. The data quality depends wholly on the data provider. There is no additional work done on the data sets except the addition of some meta information. The data that gets pushed into the system is the data which is made available to the user.

App developers need to download the raw data and process it using their own ETL (Extract, Transform, Load) processes. With every update of the raw data the ETL process has to be triggered for every single application where it is used. If the format of the raw data changed, the process has to be adjusted and cannot be automated and with every new data source, maintenance complexity of these open data sets and its apps increases.

There is far more knowledge in these data sets available than what is visible and accessible to the open data developer. However, in its scattered form, this knowledge is not accessible. Inspired by the ideas of holism we come to the conclusion that a data set should be viewed as a *gestalt*, not as collections of parts. This leads to synergy, which is the creation of a whole that is greater than the simple sum of its parts. Linked Data enables synergy for separate, unconnected data sets. If Open Data is about political and administrative interoperability then Linked Open Data is about technical and organizational interoperability.

---

[1]http://www.bfs.admin.ch/bfs/portal/en/index.html
[2]http://okfn.org/opendata/
[3]http://ckan.org/

3.2. Related Work

3.2.1. Linked Open Government Data (LOGD)

Linked Open Government Data can be a catalyst in the production of value-added services and products developed and exploited by the private sector. Services and products based on government data provide more choice and value for money to consumers as well as increasing government transparency, improving public administration functionality, and providing social value. The value chain of reusing government data covers production, collection, storage, distribution, adoption, marketing and lastly the provision of services to end users[4].

Publishing data as linked data is mainly about publishing structured data in RDF using URIs; the approach makes government data available on the Web through a RESTful API, SPARQL search endpoints and/or RDF dumps instead of making the data available as downloadable files in formats such PDF, Excel XML etc.[5]. Linked Open Data offers the possibility of using data across domains or organizations for various purposes including statistics, analysis, maps and publications; by linking this knowledge, correlations can be understood quickly, and new conclusions can arise. The best practices in the usage of Linked Open Government Data can be summarized in the following types: a) portals for the supply of LOGD-data sets, b) portals providing raw data sets of LOD for Public Administrations along with technical tools or developer kits for understanding, interpreting, or processing the provided data, c) already existing portals, acting as showrooms for LOGD best practices and d) mobile apps for smartphones using LOD for Public Administrations[6].

It is important to enable data owners to publish their data as Linked Data. Providing the necessary tools that motivate data owners to invest the extra effort needed for Linked Data publication can facilitate this.

3.2.2. Users-Driven Focus and Process Streamline

The vision of this R&D-project is to make publishing and reuse of linked data as easy as possible for the end user thanks to a thriving marketplace with data publishers, developers, and consumers along the value chain. Fusepool P3 is part of Fusepool projects[4], which digest and turn data from different sources into linked data to make data interoperable for reuse.

Previous work focuses on technology intelligence and scouting[5]: the project SME provides data for better product/service development and research. It integrates patents, journal articles, funding descriptions, and descriptions of people and organizations (profiles) and publishes Linked Open Government Data along the R5 framework: Reveal, Refine, Reuse, Release, and Run[7]. During the Fusepool project a considerable amount of time was invested in transforming legacy data to Linked Data, which led to the conclusion that building and maintaining such a workflow needs to be automated whenever possible and so this gave rise to the Fusepool P3 pipeline concept.

The proposed solutions of the project Fusepool P3 center on tourism-related applications and go beyond the state-of-the-art of different aspects of dealing with data (e.g. selection and preparation, extraction and mapping, enrichments and semantics, delivery,

---

[4]http://www.fusepool.eu/
[5]http://sme.fusepool.eu/

platform and storage). The innovation of Fusepool P3 from an organizational or. methodological point of view is the user-driven focus and the effort to involve users especially (public) data owners in all the development process phases. From a technological perspective Fusepool projects are innovative because they integrate components for various user-defined data that streamline the process of data extraction, transformation, indexing, interlinking (and visualization)[8].

3.3. Agile Development as a Methodology for Increasing Sustainability

Based on experience in earlier projects the project team decided to take a different, more agile approach in developing Fusepool P3[9]. Prior to this, tasks were split among the involved project partners and at some point, usually after several weeks or months of development, integrated into the main platform. This led to much wasted time and effort since often expectations were not clearly enough specified or software components had to be adjusted as the initial concept did not work out. This had negative influence on the schedule and outcome of the project and led to extra work.

For Fusepool P3, an initial meeting of all developers defined the rough development roadmap for the project runtime of two years with platform release dates. Within these releases the team executes 2-3 week sprints. For these sprints the product owner defines and communicates product requirements in close collaboration with future users of the platform. The developer team accepts and assigns these requirements and implements them within the time frame. To make sure deviations are noticed fast, the development team synchronizes via conference calls at least once a week, involving all partners. This allows the team to feel and act as one and help each other out where necessary. In the beginning this required some extra effort from the developers as they come from multiple companies with different working cultures. After an initial period this agile approach started to work out well and interaction between developers greatly improved.

Given that the ultimate goal of the Fusepool P3 project is to provide a platform, which is useful for current and future stakeholders, providing an initial platform as soon as possible was one of our top priorities, thus ensuring that stakeholders can start using it long before the project ends and provide important and valuable feedback that can be integrated into upcoming development cycles. The goal is not just to successfully finish the project but also to provide additional value for our stakeholders and a lasting user- and development community.

4. Integrating Stakeholders and Creating Use Cases

Integrating stakeholders as platform users at an early stage in the project was one of the key lessons learned from the first Fusepool project. While it was done in the first project it only took place during the last six months. At the final dissemination events the stakeholders gave great feedback that triggered many new ideas for the platform but at that stage there was no real community outside the core development team and project time for supporting stakeholders was over.

In Fusepool P3 the approach was different: The initial proposal already brought two stakeholders as project partners on board, which should ensure that Fusepool P3 creates a real value as soon as possible for both the involved project partners and future

stakeholders. It is essential to understand their motivation and needs in an early stage of the development process.

Talking to stakeholders generates more work initially and it takes time to integrate them and their requirements into the development process. Once this step is reached the whole team benefits as use cases have a real world background and can be tested on real data supplied by stakeholders. This greatly improves the value for the software platform.

4.1. Experience in Publishing Open Data

The Fusepool P3 project partners Provincia Autonoma di Trento (PAT)[10] and Regione Toscana (RET) have been publishing open data and supporting the development of applications and services in the tourism domain for several months. During this time both partners have gained valuable experience in data creation, maintenance and publication.

Public administrations have a long history of providing services for citizens. With the introduction of information technology in the seventies, many of these services were automated and more centralized. Moreover, with the advent and later success of the Internet, many administrations now provide vertically focused web portals, mobile apps, etc.

At the same time administrations realized that they could not provide a service for everything by themselves, even if they had many more data sets in internal systems that might be useful to the public; it would simply be too expensive to do all of this themselves. The Open Data movement at the same time started propagating the opening up of locked data sets to the public. By publishing data administrations enable third parties to provide additional services that might be useful for the public.

After publishing first open data sets the administrations realized that there could be additional value in the data if the datasets from different sources and systems were (inter-)linked with each other. In this paper we show examples of this interlinking in the field of tourism, however, the concept of interlinking also applies to many more available data sets.

PAT and RET first started publishing data sets which were considered of strategic importance. In Italy in general but also in the two regions Tuscany and Trentino the most important business is clearly tourism and related industrial activities around tourism. Thus the regions are struggling with one particular question: How can they support and push tourism by changing their daily operations and if necessary, themselves.

Based on this question Trentino 2012 started to publish data as open data, particularly related to tourism, culture, and geographical data. In July 2012 the Tuscany Region started opening up data, involving authorities, citizens and businesses and publishing it on a public platform, mainly released under a Creative Commons license.

Starting in January 2013, the Region of Tuscany officially committed to the implementation and management of an open data platform with the goal of disseminating data from the region and other regional governments.

4.2. The case for Linked Open Data

However, the regions have realized that publishing data itself is not enough; they also need to link them in better ways. There is far more knowledge in the data that could be exploited by companies and startups if this knowledge is easily accessible.

The regions also hope to change the culture in administrations in positive ways. Currently many administrations seem to be afraid of opening up data, for example because one could find errors in the public data sets. By showing them that third parties start to use their data in a different, productive way they can slowly change this culture and motivate administrations to release even more data. Providing the right set of technologies that facilitates the open data publishing process also supports this culture change.

Linked data addresses some of the mentioned problems users have with open data. It does that by providing a standardized format (RDF) that provides benefits like unique identifiers (URIs) for each piece of information (data), generalized access to such a resource via the web (HTTP), standardized vocabularies that describe the meaning of the data and allow relational links between information, support for multiple languages in its core and a standardized querying language (SPARQL).

The Linked Data technology stack[11] provides many ways to interact with data in RDF format[12]. This greatly reduces the time and technical effort needed to integrate data sets into apps and thus increases the value of open data. However, one needs new ETL processes to transform raw data to Linked Data. While it facilitates data usage for app developers, Linked Data requires initially more work by the data owner and publisher.

In the past few years many powerful tools have been developed or extended to support the creation and maintenance of Linked Data. Also new W3C standards and vocabularies have been developed to facilitate the transformation of legacy data into Linked Data. Many of these tools and standards have been developed, extended or implemented by Fusepool P3 project partners. The emerging Linked Data Platform Standard (LDP) provides standardized means of making collections of linked data resources accessible.

What is lacking is an integration framework that combines the data transformation to RDF, necessary enhancement steps and the publishing of the Linked Data. Fusepool P3 will provide such an integration framework along with User Interface tools that serve both to model the data publication process as well as to coordinate the human interactions that might be required while the data is processed.

This framework will integrate state-of-the-art tools like OpenRefine, OpenLink Virtuoso, Apache Stanbol, and Pundit. The framework is developed and tested based on the requirements defined by our project partners PAT and RET. Both partners have started working with the platform in an early stage and feedback is directly integrated into Fusepool P3's agile development process.

5. Lessons Learned and Further Activities

During our work with Fusepool P3 stakeholders we learned that publishing Open Data is not to be underestimated. Publishing existing data sets on a portal in well-known data formats is not too much work but also of limited use as it requires a lot of repetitive ETL work by users of the data. Publishing Linked Data gives access to the knowledge of an expert and data owner in open, standardized formats developed by the W3C. Initially this requires more work by the data owner as it becomes their responsibility to choose the right vocabulary for the semantic description of the data and publish it on an appropriate platform. Once this is done the data becomes much more accessible to users and easier to integrate and scale. Fusepool P3 can facilitate this Linked Data publishing step by

providing state-of-the-art tools integrated under a common user interface and pipelining API. Creating a sustainable Fusepool P3 platform that will exist long after the initial funding ends, is essential. To get there it is important to have a vibrant community of users and code contributors who develop and support the platform in the future. This can only be reached if the user involvement and community building processes already start during the initial period of the project.

Acknowledgements

References

[1] N. Mihindukulasooriya and R. Menday, "Linked data platform 1.0 primer," 2014.

[2] G. Schreiber, Y. Raimond, F. Manola, E. Miller, and B. McBride, "Rdf 1.1 primer."

[3] A. C. Neuroni, R. Riedl, and J. Brugger, "Swiss executive authorities on open government data – policy making beyond transparency and participation," in *Proceedings of the 2013 46th Hawaii International Conference on System Sciences*, HICSS '13, (Washington, DC, USA), pp. 1911–1920, IEEE Computer Society, 2013.

[4] T.-M. Yang, J. Lo, H.-J. Wang, and J. Shiang, "Open data development and value-added government information: Case studies of taiwan e-government," in *Proceedings of the 7th International Conference on Theory and Practice of Electronic Governance*, ICEGOV '13, (New York, NY, USA), pp. 238–241, ACM, 2013.

[5] E. Kalampokis, E. Tambouris, and K. Tarabanis, "On publishing linked open government data," in *Proceedings of the 17th Panhellenic Conference on Informatics*, PCI '13, (New York, NY, USA), pp. 25–32, ACM, 2013.

[6] S. Consoli, A. Gangemi, A. G. Nuzzolese, S. Peroni, V. Presutti, D. R. Recupero, and D. Spampinato, "Geolinked open data for the municipality of catania," in *Proceedings of the 4th International Conference on Web Intelligence, Mining and Semantics (WIMS14)*, WIMS '14, (New York, NY, USA), pp. 58:1–58:8, ACM, 2014.

[7] M. Kaschesky and L. Selmi, "Fusepool r5 linked data framework: Concepts, methodologies, and tools for linked data," in *Proceedings of the 14th Annual International Conference on Digital Government Research*, dg.o '13, (New York, NY, USA), pp. 156–165, ACM, 2013.

[8] "Fusepool publish-process-perform platform for linked data. annex i - description of work," 2014.

[9] R. Pichler, *Agile Product Management with Scrum: Creating Products that Customers Love*. Pearson Education, 2010.

[10] P. Shvaiko, F. Farazi, V. Maltese, A. Ivanyukovich, V. Rizzi, D. Ferrari, and G. Ucelli, "Trentino government linked open geo-data: A case study," in *Proceedings of the 11th International Conference on The Semantic Web - Volume Part II*, ISWC'12, (Berlin, Heidelberg), pp. 196–211, Springer-Verlag, 2012.

[11] T. Berners-Lee, "Linked data - design issues."

[12] C. Bizer, T. Heath, and T. Berners-Lee, "Linked data - the story so far," *Int. J. Semantic Web Inf. Syst.*, vol. 5, no. 3, pp. 1–22, 2009.

[13] M. Kaschesky, A. Gschwend, G. Bouchard, P. Furrer, S. Gamard, and R. Riedl, "Aid to regional development agencies: Finding and matching research funding opportunities," in *Proceedings of the 13th Annual International Conference on Digital Government Research*, dg.o '12, (New York, NY, USA), pp. 2–10, ACM, 2012.

Establishing a MDM function: First Steps in the Master Data Management Architecture Design

Riikka VILMINKO-HEIKKINEN[1]

Department of Information Management and Logistics, Tampere University of Technology, Finland

Abstract. Public sectors need to manage data for internal and external use is growing. Master data management aims to manage the core data that affects generally the data quality in large extent. The objective of the research is to observe the factors that affect the architectural decisions when establishing a MDM function. This is done though prior research and a case study in a municipality. Business needs and the existing IT environment indicate the best usage scenario for MDM. Scenario indicates the suitable architectural implementation style and also helps the organization to comprehend what they should emphasize in their development.

Keywords. Master data management, MDM, data management, Business intelligence, Open data

1. Introduction

Open data, big data and the growing need for predictive analytics set demands for public sector. Better data management practices are needed to ensure high quality of data for internal use, external use, and re-use. Master data management (MDM) aims to manage an organization's core data (i.e., master data). It tackles the data quality issues through process improvement by using organizational and technical aspects.

The focus in prior research has been in implementation styles and application architecture on private sector. Establishing MDM and the first steps of MDM architecture design have gained less attention. MDM development is a relevant topic in public sector in Finland. From September 2011, the Act on the Direction of Public IT Governance has mandated the use of an Enterprise Architecture (EA) Framework. The National Enterprise Architecture (NEA) steers the EA development [8]. Also MDM is a part of the NEA and the National MDM reference architecture was finalized in 2013. The objective of the research is to observe the factors that affect the architectural decisions when establishing a MDM function. This is done though prior research and a case study in a municipality.

The paper is organized as follows. First, we identify how prior literature has observed the architectural questions in MDM. The research methods and settings are then described. Empirical part follows and concludes with a discussion and recommendations for future research.

---

[1] Corresponding Author.

2. Related research

MDM should be seen as a process improvement plan to identify, assess, and implement methods and tools for assuring good quality data for the decision process [6]. The usage scenario defines the method of use for MDM.

In an operational scenario [10], [5], all interactions and transactions are applied to the master version, and consistency requirements must be strictly enforced [10]. Analytical scenario is relevant when the need is primarily to produce and maintain master data for data warehouse (DW), reporting purposes, analytics, and big data systems [10]. Enterprise scenario combines analytical and operational scenarios [1] and is similar to reference information management. The focus is on the importing of data into the master data environment and the ways that the data is enhanced and modified to support the dependent downstream applications. The collaborative usage scenario emphasizes achieving an agreement on a complex topic among a group of people [5]. This usually includes workflows and multiple tasks. Social MDM is a more recent perspective on MDM. It focuses on providing a platform for gathering, integrating, and stewarding a wider set of customer and product data, and for making them available through-out the organization [13]. Some of the unique characteristics of each usage scenario are presented in Table 1.

Table 1 Unique characteristic of different usage scenarios for MDM

Usage scenario	Characteristics
Operational	— Operational systems must execute their transactions against the master data environment instead of their own data systems [10], [5] — Integrates operational applications (e.g., ERP, CRM) in upstream data flow [15] — Individual and application access is closely monitored [10] — The MDM repository is considered as the authoritative source [11]
Analytical	— Applications are more likely to use than create master data [10], [14] — Applications can effect classifications/categorization of master data records [10] — MDM systems key role is to be a provider of consistent data for BI [5], [11], [10] — Intersection between the MDM and BI [5], mainly for reporting purposes [17] — MDM systems may include some key features for analytics [5] — Classifying master data is done through analytics [10] — Resembles customer data integration (CDI) [1] — Uses a unidirectional flow of data to the master record, using extract, transform and load (ETL) processes before importing the data [14] — Direct modifications to the master data can be made [10] — Least intrusive approach and most used [3]
Enterprise or reference information management	— Master records are created directly [10] — System coordinates users and systems to reach agreement on a data set [5] — Data in incorporated into the master environment, where it is available for publication to client applications [10] — Ensuring the quality of data in the entry [10] — Direct modifications to the master data can be made [10] — Classification made directly according to predefined rules [10]
Collaborative	— Collaborative authoring of master data: creation, definition, augmentation, and approval [8, 15]
Social MDM	— Analytics derived relationships [13] — Especially affiliated to big data capabilities [13] — Customer centricity (people and organizations) [13]

Repository, registry, and hybrid are common models for implementing MDM architecture. In repository, the complete collection of master data is stored in a single database

[10]. The global attributes of the data set are always created in central master data system [9]. In registry model, data sets are created, maintained, and distributed by different applications [9]. The hybrid model includes features of both of these approaches [10]. Also a consolidation model has been identified [16]. Fragments of master data are authored in a distributed fashion and stored in the source systems, but the central MDM system creates a composite golden record. Architectural models complement each other, and several models are sometimes used to fulfill the needs of an organization [18].

Master data application architecture provides complete overview of how the architecture is deployed by using different applications [2]. It contains applications for creating, storing and updating instances of the master data attributes defined by the conceptual master data model [19].

Many of the case studies addressing MDM architecture design (e.g. [3, 20, 21, 23]) have been focused on the private sector, the target of this research is a municipality, which offers an interesting opportunity to observe the architectural design decisions in a public sector organization. The focus will be on different MDM usage scenarios that have not received much attention in prior research.

3. Research Methods

An ethnographic case study (following the instructions of Myers [12]) was conducted in a municipality comprising 220,000 inhabitants and approximately 14,500 employees. The case organization consists of central administration, purchasing unit, welfare services, municipal corporations and several subsidiaries. The MDM projects and development were mainly conducted in the central administration.

The data collection period lasted from November 2010 to June 2013. During this period, two MDM development projects were carried out, and MDM development became rooted as a part of routine operations. The data was collected by participating in all project-related meetings and informal discussions in both projects, and in the project preparation and procurement phases of the second project. The author was actively involved in the first project as a member of both the steering group and the expert group. In the second project, the author acted as a project manager and as a member of the project group and the steering group. These positions offered unique opportunities to observe and thoroughly understand MDM implementation.

Diary entries were made weekly and whenever MDM-related issues were observed. In addition to observations, also questions that emerged and impressions were documented. To complement the diary, different kinds of project documentation were also utilized: procurement documentation, project plans, monthly status reports, and a set of memos from the working group, steering group, project portfolio group, stakeholder groups, and kick-off and closing seminars. Between the two projects, memos from the IT development group and the architecture group were also used. Finally, some internal documents were utilized, such as information management strategy.

The data was analyzed by following the interpretive research approach and the principles of Klein and Myers [7]. The data was re-viewed by observing it throughout and identifying what kind of objectives were presented during the process for establishing the MDM function. These objectives were categorized under several themes. Then these themes were used to review to data again. This time all issues supporting the themes were gathered. As a result, a list of the objectives was formed and discussion around these objectives was analyzed.

Ethnographers need to balance subjectivity and objectivity. Ethnographies are expected to meet standards of objectivity even when ethnographic research is highly dependent on the individual's unique knowledge and experience, and his/her actions as a thinking agent who brings his/her subjectivity to bear on the construction of information and knowledge [20]. All materials were analyzed in their entirety at the end of the overall data collection in June 2013. The idea was to gain some distance between the researcher and the context, and to keep data entries as neutral as possible so that they were not limited or affected by the analysis of earlier entries. This was done to minimize unintended entry manipulation, as one may easily make subconscious decisions about what to record.

4. Findings

The organization's centralized IT unit has been in operation since 2007. Previously, the units acquired information systems separately, with the exception of some organization-wide systems (e.g., Enterprise Resource Planning, ERP). As a result, it has approximately 400 information systems from different operating areas. The motivation for starting the MDM process was problems with data quality in main business processes. It was assumed that the problems originated from both maintenance processes and applications. The business objectives for MDM were identified for the first time in 2008. These were more efficient work, improved reporting, and service oriented architecture (SOA) interoperability.

The organization has several different external stakeholders (e.g. government agencies) that the organization is obligated to report regularly and also ad hoc (diary: 1/12). Data quality was considered one of the barriers to generating high-quality reports (BI report: 1.2.13). One of the clearest objectives became first supporting internal and external reporting (Diary 11/11) and later on supporting the BI comprehensively (Diary: 2/13).

MDM also had a role in harmonization of data structures, which would ease the difficulty in combining information (Diary 9/11). The last of the business objectives "service oriented architecture (SOA) interoperability" was not current later on in the development. The organization made the decision (EA principle) not to use SOA in the development of new applications in general.

The organization has been struggling with problems regarding data maintenance. Formal processes were inadequate and employees had invented additional ways to maintain the data to solve problems. One of the basic objectives is to simplify the process for data life-cycle management and to automate functions that had been performed manually. A large amount of the master data was still stored manually in Excel sheets, making maintenance difficult and error prone (diary: 3/12). For example, organization's products and services are managed manually (Diary 10/11).

Several steps should be followed when creating, changing, or deleting data (Business Workshop: 27.9.12). It is important that checkpoints for changes are in place and those checkpoints are automated (Diary: 1/12). Workflows are particularly important when there are several tasks in the workflow or when a task is performed less frequently (Diary: 4/12). The data quality validation should be done while creating the data. History of the data should be also stored. Data standards were seen as an important issue in tackling data quality issues (Steering Group memo: 11.10.11). Several roles

were identified to enable a finely divided control of the data. Roles were strictly limited to the need to make changes (Security and Privacy Work-shop: 13.9.12).

The organization's master data included sensitive attributes. The ongoing situation was that this information is often in several applications and there were problems with access management. Data is also imported from external sources. Usually this was done separately to the individual applications and the data would be obsolete in this respect in other applications (Diary 4/12). MDM should support compliance and provide a reliable foundation to support changes and updates in policies to help avoid penalties or other regulatory actions (Diary 8/12). The MDM system´s log should make it possible to identify problems relating to data misuse or other issues. For example, there are examples about data of organizational unit being deleted incorrectly (Diary: 9/11).

Organization´s master data domains include several hierarchies. Maintenance of these should be coherent (Diary: 1/12). The hierarchies should match those used at government level, and the maintenance should be synchronized (Project Group memo: 30.1.12). Master data is also affected by the definitions of government and government agencies and the need for a change often comes from a stakeholder. These affect the modification needs of the data models, attributes, and also the metadata (Business Workshop: 27.9.12). Master data objects should be categorized in different ways to present the perspectives that the data is used and observed. This serves especially the needs of the BI.

Several ways for enhancing data quality were identified (IT Workshop: 28.8.12, 4.9.12, 7.9.12, and 18.10.12). External sources should also be used for validating the data (IT Workshop: 28.8.12). The need for methods to continuously monitor the data quality was also identified (Project Group memo 28.9.2011; Diary: 4/12; Diary: 9/12).

The organization outlined an initiative to open data in a machine-readable format as part of their operations (Diary 5/12). Master data has high value in terms of re-use (also for commercial use) (Diary: 10/11) and in making the organization's operations transparent. It was seen as a problem that the data was scattered. For this purpose the structure of the data as well as easy access was essential. Master data often includes attributes that cannot be opened, it is important that there is attribute level access control that helps to control what data is published and by whom. Also, the understanding in the organization about this is part of the data governance and data privacy (Diary: 4/12; Security and Privacy Workshop: 13.9.12).

5. Discussion

The case organization's business needs were observed and classified. These were categorized under different usage scenarios. Summary is presented in Table 2.

Table 2 Business needs categorized by usage scenarios

Reference/Enterprise	Analytical	Operational	Collaborative
Control and formalize data creation	Automate infrequent tasks	Streamline work processes and the organization	Workflows, approval points
Function-oriented instead of system-oriented doing	Generate automatic reports of changes for stakeholders	Control and formalize the creation of data, Reduce errors in data entry	Decentralized data creation
Reduce errors in data entry	Automated and secure publishing	Access control to sensitive data	Divided control of the data
Direct data entry through one UI	Publishing machine-readable open data	Comply with security and data privacy rules	

Easier master data quality management and enhancement	High quality data source for DW (relationships between the concepts)	Better interoperability and view of the organization's core data
Supporting data maintenance	Data classification	High-quality data for process-es
Data standards, classi-fication based on predefines rules	High-quality data for reporting and enabler for BI	Easier master data quality management and enhance-ment
Modification of metadata		Improve the data quality
Data categorization		Use of data standards
Detect errors before they affect functions		Approval function and formal process for maintenance
Comply with security and data privacy		Supporting the maintenance of data

Many of the objectives responded to analytical scenario's characteristics. Despite this; it was also quickly evident that the scope was wider. Reporting was seen as an important area, but the main problems affecting it were issues with data quality in op-erational systems. The development of BI was ongoing and the clarifications in the BI development also shed some light to MDM development. Big data development was observed in the last phase of the data collection. The analytical scenario was not effi-cient for the organization in long run. As seen in Table 2, business needs support al-most all of the scenarios in some extent. Collaborative scenario seemed very separate from the other scenarios. It seems to be more a perspective on the other scenarios than a distinct method of use. Social MDM approach would acquire a more mature phase of big data development and use. The vision for BI included, e.g., idea to use social media data to enrich the customer data, but it there was no clear development plan for big data.

Number of the objectives supporting the scenario was not adequate to refer to the suitable scenario. Cervo and Allen [3] emphasize indirect objectives, such as cost sav-ings, as assessment criteria, but in the case organization these were not considered at this stage. Cost savings were much more seen as a result of the business needs, instead of a separate factor. Because of this, organization and IT-landscape were also observed as factors that emphasize the suitable scenario option.

The model of centralized IT would support the operational scenario. Organiza-tion's IT environment is complex and for the legacy systems, publication to other ap-plications as defined in reference architecture would be more interoperable with these. Operational usage scenario seemed to echo to the business needs and IT environment most. Because of this and the needs for analytical usage, the organization's usage sce-nario was enterprise MDM.

The MDM usage scenario narrows the options for architectural implementation styles. Certain styles respond better to certain usage scenarios [3]. At the final stage of the data collection, the organization decided the hybrid model as an implementation style. This style is suitable when the organization is looking for a method to improve and manage the data quality, completeness and consistency of master data across sev-eral systems and the organizational commitment and the re-sources support proper data governance activities [16]. Dreibelbis et al. [5] have also identified the characteristics of legacy systems as constrains for choosing an architecture pattern for MDM. The organization's ERP was seen as a master system for two of the identified master data domains. Hybrid model has been identified as a good fit for ERP environments [11].

The current IT-landscape was the main influencer for making the decision of the MDM application architecture. The large number of legacy systems was one of the reasons why MDM architecture included a separate MDM system. The organization struggled to make the decision between a dedicated MDM system and an operational system converted to a master data source. The amount of the existing applications was the reason why they were reluctant to acquire a new system. Also the cost of a new solution was seen problematic. In the end, the extent of the desired MDM development resulted in the need for MDM system to enable the required elements and functions.

Prior research has not addressed MDM in the public sector or especially in the municipalities, nor has it made generalizations about the common or distinguishing features between the public sector's and private sector's MDM. The type of the organization affected the MDM development in many ways, but did not seem to affect significantly the first choices made with the architectural design. However, the MDM application architecture was affected by the existing IT-landscape.

6. Conclusions

The objective of the research was to observe the factors that affect especially the first architectural decisions when establishing a MDM function. The prior research names three layers in the MDM architecture design. First step includes identifying the method of use for MDM. Second step is to determine the right architectural implementation style for MDM. Last step, the application architecture, defines the technical architecture in detail. The first step has gained less attention and because of this, it was observed more closely through the case study.

The factors affecting the first architectural decisions were complex. The maturity stage of the organizations BI responded well to analytical usage scenario. For an organization that has higher maturity, social MDM could respond well to the needs. Operational view contemplated the demands for streamlining work processes, and enhancing data quality in operational systems and processes across the organization, and the needs that open data places. It also responded well to the data security and privacy demands. Because there were characteristics of multiple of the approaches, the usage scenario could be identified as enterprise MDM. Business needs should indicate primarily the chosen architectural scenario, but also other factors have an effect. IT environment was the final factor in identifying the suitable usage scenario. This also indicates that the architectural implementation style should respond to different use scenarios.

The MDM architectural design is complex and challenging to design. To start the process by identifying the usage scenario for MDM through business objectives and IT environment sets the architectural path to right direction. Usage scenario indicates the suitable architectural implementation style and also helps the organization to comprehend what they should emphasize in their development. MDM usage scenarios have been dismissed in prior research for not being elaborate (e.g. [14]), but it seems that they present a good starting point for designing MDM architecture. Usage scenarios imply the applicable architectural styles and further on also the suitable application architecture. They also clarify how objectives are translated into design decisions for the organization.

Public sector's master data management practices have gained very little attention in research. Still, these should be emphasized more, because their data is transforming into public data. Public sector organizations set interesting research settings because of

their complex IT landscapes and diverse data domains. MDM has some unique characteristics when established in public sector (e.g. data privacy issues). This research was a single case study, and caution should be exercised with regard to generalizations. For future research, there are several different paths to explore. The effects of usage scenarios to the design of governance might offer interesting perspectives on the overall concept of MDM. The social MDM in big data development also offers an interesting viewpoint.

Acknowledgements

References

[1] C.G. Apostol, Enterprise Master Data Management Trends and Solutions. *Informatica Economica*, 3(43) (2007), 35-39.

[2] E. Baghi, S. Schlosser, V. Ebner, B. Otto, H. Österle, Toward a Decision Model for Master Data Application Architecture. *In: 47th Hawaii International Conference on System Science*, (2014), 3827 – 3836.

[3] D. Cervo, M. Allen, *Master Data Management in Practise: Achieving True Customer MDM*. Hoboken, New Jersey: John Wiley and Sons, 2011.

[4] T.P. Das, R.M. Manas, A Study on Challenges and Opportunities in Master Data Management. *Int. J. of Database Management Systems*, **3**(2) (2011), 129-139.

[5] A. Dreibelbis, E. Hechler, I. Milman, M. Oberhofer, P. van Run, D. Wolfson, *Enterprise Master Data Management: an SOA Approach to Managing Core Information.* Westford, IBM Press, 2008.

[6] M. Fung-A-Fat, Why is Consistency So Inconsistent? The Problem of Master Data Management. *Cutter IT journal*, 20(9) (2007), 23-29.

[7] H.K. Klein, M.D. Myers, A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems. *MIS Quarterly*, **23**(1) (1999), 67-93.

[8] J. Lemmetti, S. Pekkola, Enterprise Architecture in Public ICT Procurement in Finland. *Electronic Government and Electronic Participation* (2013), 227 – 236.

[9] C. Loser, C. Legner, D. Gizanis, Master Data Management for Collaborative Service Processes. *In: International Conference on Service Systems and Service Management* (2004), 1-6.

[10] D. Loshin, *Master Data Management*. Burlington, MA, Morgan Kauffman, 2008

[11] A. Maedche, An ERP-centric Master Data Management Approach. *In: 16th Americas Conference on Information Systems*, Paper 384 (2010).

[12] M.D. Myers, Investigating Information Systems with Ethnographic Research. *Communications of AIS*, **2** (1999), 1-20.

[13] M. Oberhofer, E. Hechler, I. Milman, S. Schumacher, D. Wolfson, *Beyond Big Data: Using Social MDM to Drive Deep Customer Insights*. Pearson, Indiana, 2014.

[14] B. Otto, How to Design the Master Data Architecture: Findings from a Case Study at Bosch. *Int. J. of Information Management*, **32**(4) (2012), 337-346.

[15] R. Silvola, O. Jääskeläinen, H. Kropsu-Vehkaperä, H. Haapasalo, Managing One Master Data - Challenges and Preconditions. *Industrial Management & Data Systems,* **111**(1) (2011), 146-162.

[16] A. White, D. Newman, D. Logan, J. Radcliffe, *Mastering Master Data Management*. Stamford, Gartner Research, 2006.

[17] A. White, J. Radcliffe, *Four dimensions of MDM*: Understanding the complexity. Stamford, CT: Gartner research, 2007.

[18] R. Vilminko-Heikkinen, S. Pekkola, Establishing an organization's master data management function: a stepwise approach. *In: 46th Hawaii International Conference on System Sciences* (2013) 4719-4728.

[19] R. Winter, R. Fischer, Essential Layers, Artifacts, and Dependencies of Enterprise Architecture, *J. of Enterprise Architecture*, **3**(2) (2007), 7-18.

[20] U. Schultze, A Confessional Account of an Ethnography about Knowledge Work, *MIS Quarterly*, **24**(1) (2000), 3-41.

Open Skies with Cloud Computing

Mauricio SOLAR[a,1] and Mauricio ARAYA[a]

[a] *Universidad Técnica Federico Santa María, Chile*

Abstract. This article shows that high costs to obtain data in astronomy justifies the use of cloud computing. In such a field, like others in science, there are several benefits when this data is freely available to the whole scientific community. The availability of these data will foster innovative applications to exploit and explore data. These applications could be offered as a service in the cloud. A National Data Center for e-Science could take advantage of these high costs data and offer open data to the community to maximize data accessibility.

Keywords. open data, data science in government, cloud computing in science.

1. Introduction

The concept of open access to scientific data was institutionally established with the formation of the *World Data Center* system, in preparation for the International Geophysical Year of 1957–1958 [1]. The *International Council of Scientific Unions* (now the International Council for Science) established several World Data Centers to minimize the risk of data loss and to maximize data accessibility, further recommending in 1955 that data be made available in machine-readable form (http://www.icsu-wds.org).

More than 50 years later, in 2007, the Science Ministers of all nations of the *Organisation for Economic Co-operation and Development* (OECD), which includes 34 most developed countries of the world, published the "OECD Principles and Guidelines for Access to Research Data from Public Funding" as a soft-law recommendation [2].

In recent years, the flood of data problem is visible in both, science and business environments, becoming more relevant the proper use and management of what has been called Big Data. This concept encompasses the management research of large amounts of information, and that these are not easy to process using traditional tools and procedures. When the data volume reaches TeraBytes (TB) to ZetaBytes (ZB) ranges, algorithms and procedures must be adapted for its use in new high-performance computing platforms, with Cloud tools, in a distributed manner, and on-line. Additionally, we are not only dealing with large stationary volumes of data, but also with the data-generation frequency, and the heterogeneous nature of data sets; this creates new challenges in developing solutions, as are the storage, variability in the format, and response time.

Big Data is not a technology in itself, but rather a work approach to obtain value and benefits of these large volumes of data generated nowadays. These are some of the features to consider:

- How to grasp, manage and make use of them
- How to ensure, check authenticity and reliability
- How to share and obtain improvements and benefits
- How to communicate, simplifying the decision-making and subsequent analysis.

One of the domains where the Big Data problem is approaching its turning point is astronomy. Its state-of-the-art facilities in operations, such as the *Atacama Large Millimeter/submillimeter Array* (ALMA) will generate over 1 TB of data per observation day [3], and those under construction, as the *Large Synoptic Survey Telescope* (LSST), and the *Square Kilometer Array* (SKA) will produce large-scale data [4]. The plan for the year 2020 is to have more than 60 PetaBytes (PB) of accessible information for the astronomical community. The LSST initial computer requirements are estimated at 100 TFLOPS of computing power and 15 PB of storage, rising as the project collects data (full operations for a ten-year survey commencing in January 2022), (http://lsst.org/lsst/science/overview).

The cost to design, build and operate such mega-observatories is very high, so the cost of on-site observation time and the end-to-end cost to obtain astronomical data is also very high. Therefore, it would be very profitable if the obtained data could be freely available to the whole scientific community.

In Section 2 we show the costs involved in the construction of observatories. In Section 3, we show the *International Virtual Observatory Alliance* (IVOA) [5] as a good example to follow in other fields of science in the world. We also show the development of the Chilean Virtual Observatory (ChiVO) based on the standards and protocols defined by IVOA. In Section 4 we show some challenges to install a data center to offer e-science data and cloud services to the scientific community. Finally, we conclude in Section 5.

2. Big Data, Big Costs

Computationally, the concern is focused on the rapid growth of the generated data volumes, which had passed from the GigaBytes (GB) to the TB level in the past decade, and will pass from TB to PB in the near future. For example:

- *Galaxy Evolution Explorer* (GALEX): the first orbiting telescope in the space generated 30 TB of data in the first 3 years of operation [6].
- *Sloan Digital Sky Survey* (SDSS): in its 12th version gathers 116 TB of data published as images, catalogs, and other products (www.sdss.org/dr12/data_access/volume/).
- *Panoramic Survey Telescope & Rapid Response System* (Pan-STARRS): its goal is to characterize objects that will come closer to the earth, as asteroids or comets. 10 TB of data is expected per night of observation [7].

We could continue listing more observatories, but in summary, there is a growing *avalanche* of astronomical data, changing the way that astronomy is done nowadays. From the computational point of view, it is necessary to implement standard services that allow to access, process, and modeling the data produced by each observatory (which generates public data).

2.1. Cost of Observing

The privileged atmospheric conditions make Chilean skies one of the most favorable places for astronomical scientific research: over 330 clear sky nights per year. Chile hosts also the world's closest ground observatories to space, at an altitude of 5,000 meters above the sea level (masl). Clear skies, easy access, communication infrastructure, isolation from urban settlements, and effective protection of sighting sites from light/luminal contamination, preserve the area for astronomical purposes. For these reasons, many observatories have settled in the Regions of Antofagasta, Atacama and Coquimbo.

There are more than a dozen wide sweeping astronomical facilities throughout Chile; for example, the already mentioned ALMA, the "*Very Large Telescope*" (VLT), and in the coming years, the "*European Extremely Large Telescope*" (E-ELT). As a result of Chilean clear skies and appropriate weather conditions, by 2018 about 68% of the global astronomical infrastructure will be settled in Chile [8]. Table 1 shows the investment (over USD 5 billion) in new astronomical observatories installed or to be installed in Chile [8].

Table 1. Investments in observatories

State	Telescope	costs US$
Operating	Magellan Telescopes	100M
	Gemini	300M
	VLT	700M
	ACT	40M
	ALMA	1.500M
In construction	E-ELT	1.500M
Projected	GMT	800M
	TAO	100M
	LSST	500M
	CCAT	200M
	Total	**5.740M**

One of the conditions established as country is that 10% of the observing time should belong to the Chilean astronomical community; thus, it is justified the need to develop a national astro-informatic platform, for an intelligent management, and analysis.

According to the investment indicated in **Table 1**, the estimated cost of the 10% of the observation time of astronomical observatories in Chile is (considering operational costs as 5% of total investment):

- About US$30.000K in the year 2011.
- About US$87.000K in the near future (2018-2020);
- Between US$ 25.000 and US$ 30.000 is valued one hour of observation time in ALMA.

2.2. Other (Big) Costs

The current and future large-scale data generated by the astronomical observatories placed in Chile, have created new needs and requirements, which serve as an opportunity for the development of a Data Center (DC) to offer public data to the scientific community. To have a general idea of the amount of data that will need to be stored and processed, we can consider the ALMA project: the fully operational

observatory (with the complete array) will generate over 1 TB of data per observation day. The handling of large-data volumes can generate complications and high costs in the following issues:

- **Storage**: it is necessary to have a DC capable of storage according the data-consumption needs, without ignoring the physical space of the equipment and the architecture behind the storage procedure.
- **Access**: the open data requirements of astronomical data sets stress the necessity of a reliable and fast access from anywhere and for anyone. A web-based system under a high-speed networking environment is an appropriate mechanism to supply this requirement.
- **Processing**: data processing is an important area of applied computer science, which is concern on understanding the nature and the structure of the data, as well as the development of tools and techniques that can be used to carry it out. When the processing to accomplish deals with large volumes of data --- whether it is correcting, calibrating, analyzing, etc.---, it requires more time than usual, and of course, more computational power, resources that a user usually does not have.

Thus, the need of a system with these characteristics, as a solution to public access and advanced manipulation of the large-scale astronomic data, is an excellent good to maximize the availability of public high-cost data. The idea of a Virtual Observatory (VO) was born in the year 2002, and the standards and protocols of how to inter-operate between different resources are in charge of IVOA [5], [9], [10].

The VO paradigm is an international initiative that enables universal access to astronomical data, under the responsibility of specialized centers for its storage and processing, at which both astronomers and regular people may access. With the standardization of methods and information is possible to study the astronomic registries without requesting new observations, reducing the physical requirements of instruments and locations, and minimizing data duplication.

2.3. Benefits of VO

The cost of acquiring new astronomical data is too high. Moreover, astronomers use their data to proof, test or discard their theories once, and after that they usually don't need more those data. This is why observatories publish this kind of data as a public data after one year (or more depending on the observatory's policy).

ALMA is an astronomical interferometer of radio telescopes using two or more radio antennas. When combining these signals to analyze them, it is possible to obtain detailed information of the source of emission with unprecedented resolutions. From the standpoint of data production, this process generates 3D cubes given by two position shafts in the celestial vault, and one shaft of frequency spectrum (Figure 1). The uniqueness of these data cubes is their size, since the spatial and spectral high resolution that this observatory provides generates large-scale data cubes (GB and TB).

An astronomer will receive a big file containing more information than he needs, because he wants to proof, test or reject his theory with a very small portion of the data cube he received. Usually, only a small part of the image is actually analyzed (see the square on the left side of Figure 1), or only a slice of the cube is used. Even though this is correct from the astronomer point of view, all the rest of information could be very helpful to other astronomers.
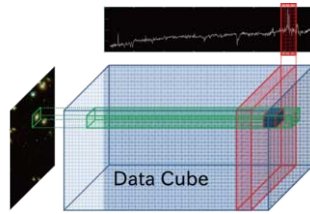
Figure 1. ALMA data cube

Open data allows reusing these very expensive data, and making it available to other scientists. In fact, the use of this open data increases the number of publications in astronomy: about 3 per researcher per year [11]. The GAVO Millennium database service is been referred to in more than 10% of the "Millennium" publications [12].

In the case of the observatories installed in Chile, it is common that big data is transferred from the southern hemisphere to the northern hemisphere. This obviously takes long time: in most cases moving big files will take hours if not days.

3. Virtual Observatory as an example of e-science

The first important concept to understand is that the VO is not a software package that allows user access to data like a web repository, despite it is exactly that what is expected from those two words. Technically, a VO can be described as an integral architecture, which formalizes in each application level, necessary protocols and standards to interoperate between VO partners worldwide. Therefore, when an observatory wants to publish data, it will not be the need to invent a new architecture, as they can use the existing one.

IVOA's basic mission is focused on coordination and collaboration between facilities, which is the main entanglement to enable global and integrated access to data collected by the international astronomical community. In IVOA, 21 VOs are currently engaged, and is composed by several working groups that discuss about the creation and versioning of standards and protocols, including its architecture.

The Data Access Protocol (DAP) defines a family of access services interfaces to the astronomical data available through VO. The DAP describes how the data providers share the information to users, and how users retrieve information. There are several standards involved, some of these are:

- **Simple Image Access** (SIA) [13]: A query defining a rectangular region on the sky is used to query for candidate images.
- **Simple Cone Search** (SCS) [14]: The query describes sky position and an angular distance, defining a cone on the sky.
- **Simple Spectral Access** (SSA) [15]: defines an interface to remotely discover and access one-dimensional spectra.
- **Table Access Protocol** (TAP) [16]: defines a service protocol for accessing general table data. The access is provided for both database and table metadata as well as for actual table data.

Each service returns a list of candidate images, astronomical sources, data or metadata formatted as a *VOTable* [17].

3.1. Chilean Virtual Observatory (ChiVO)

Although Chile is one of the most dynamic countries in its astronomic activity in the world, it did not have a VO until a year ago. For the moment, nor has ALMA services to support the protocols and standards of other VOs; therefore, it was a challenge to pose the needs and requirements of this type of platform.

Though the international community has been working and refining VOs and their standards, every new telescope and instrument imposes new challenges and opportunities of development. This is particularly true for the ALMA data case, which introduces the Big Data problem as a current problem. This means to equip the VO that is hosting its data, with last generation technology and frontier research in its tools.

The development of the ChiVO aims the development of several VO services following IVOA standards. The ChiVO will host the data of the observatories located in Chile, adhered to the interoperability standards of IVOA.

Broadly, the architecture of IVOA has three layers:

- **Users**: astronomers and scientists in general, interested in the data published by the observatories.
- **Resources**: observatories and centers producing astronomical data.
- **Intermediate layer**: defines what a VO is; that is, how users and resources communicate using protocols and standards to search and access data.

3.2. Requirements

The creation of the ChiVO required the identification of the current needs of the national astronomy community, which can be summarized in:

- **Discover:** Find astronomical data of an object or instrument on a high dimension specific region of the space, based on the spatial, temporary, spectral shafts, red shift, polarization parameters, etc., either by search or exploration.
- **Obtain**: A download link of the required data in different formats, either through the VO or through an external service.
- **Compare**: Information crossing of the data obtained between the different sources of information.

A multidisciplinary team participated in this process (astronomers, engineers, scientists, experts in ALMA data, etc.), while the astronomy community was defining its requirements and use cases, the IT team contrasted it with international standards and designed the following architecture and development model (Figure 2).

Abstraction layers: **Clients.** This layer represents the final user, and how communication between the user and data can be simplified. In this layer the user conducts queries through the access protocols offered by ChiVO, or through an advanced form, using compatible applications with VO and its web portal. Once the query is performed, the system returns to the user a list describing the objects or observations found (metadata), and provides access to them through a downloading link associated to each result.

Abstraction Layer: Applications. In this layer, we find the programs that process the queries between users and data. The server-side analysis tools are essential for ChiVO efficiency, because data are usually large and transfer it is expensive. Bringing

the analysis and processing tools closer to the place where the data is stored solves this problem as cloud computing does. Moving the algorithms to the data is faster than moving the data to the algorithms.

Abstraction Layer: Data. This layer contains resources that describe the data and metadata. This is basically a relational database that stores metadata linked to the data model recommended by IVOA (ObsCore DM).
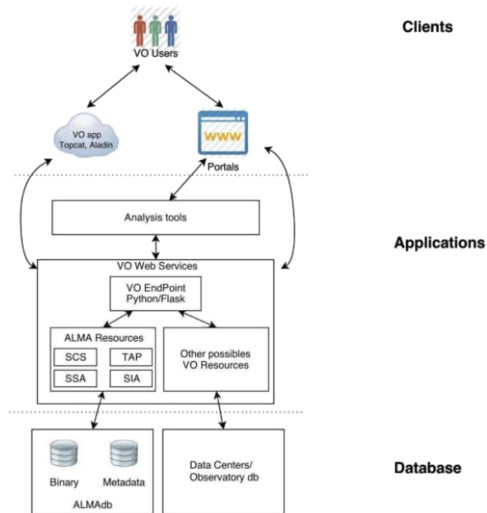
Figure 2. ChiVO architecture

4. A Data Center (DC) or Cloud Service for Open e-Science Data

Observational sciences need to secure their data because experiments are usually expensive and difficult to reproduce. In astronomy is even more important because each observation have a unique timestamp, so each experiment is irreproducible from the time domain point of view. Therefore, experimental sciences need to store their data in the order of PB of data indefinitely (ideally). Unfortunately, the cost of this storage requirement is high. In this line, there are different challenges that should be faced in Technological, Organizational, and Economical terms.

Besides those high-level challenges, Big Data science also imposes new challenges as: (a) data are being produced at an exponentially growing rate, which is currently exceeding the capabilities of local storage and processing capabilities; (b) data production and collection is expensive and it must be accessible and preserved for a very long time during operations (many decades), and even beyond the end of the specific project that provided the funding; and (c) historical data could be more important than the actualized data, so there is no clear archiving mechanism that can ease the previous challenges.

A DC for Science and Industry shall provide the capacity to store 1 EB (1000 PB) worth of data. The ChiVO DC is starting with 1 PB, will be connected to a 10 GB network, and will have a reduced environmental impact, a Power Usage Effectiveness (PUE) less than 1.3.

5. Conclusions

It is a fact that the cost of obtaining astronomical data is too high, so it does not seem reasonable to use just a portion of the data and discard it without sharing it with other scientists. With these datasets we can make further researches and even produce more and higher-level knowledge. A cloud service for storing data should provide search capabilities to the scientists that allow finding and downloading, if necessary, the required files containing the data. The offer of cloud services allows scientists to access data, and to take advantage of computing capabilities required to process big data.

Acknowledgements

References

[1] National Academy of Science. *Earth Observations from Space: The First 50 Years of Scientific Achievements Committee on Scientific Accomplishments of Earth Observations from Space*, National Research Council. ISBN: 0-309-11096-3, 142 pages. (2008)
[2] OECD. *OECD Principles and Guidelines for Access to Research Data from Public Funding*. http://www.oecd.org/science/sci-tech/38500813.pdf, (2007)
[3] Atacama Large Milimeter/sub milimeter Array (ALMA). ALMA Basics - Welcome to the Science Portal at NRAO. https://almascience.nrao.edu/about-alma/alma-basics, (2013).
[4] Ball N.M., Brunner R. J. Data Mining and Machine Learning in Astronomy. *International Journal of Modern Physics D*, **19**:1049–1106, (2010).
[5] International Virtual Observatory Alliance. IVOA.net. http://www.ivoa.net/, (2014).
[6] Martin et al. The Galaxy Evolution Explorer: A Space Ultraviolet Survey Mission. *The Astrophysical Journal*, **619**:L1–L6, January, (2005)
[7] Denneau L. et al. The Pan-STARRS Moving Object Processing System. http://arxiv.org/pdf/1302.7281v1.pdf (2013)
[8] Addere Consultores. *Capacities and Opportunities for Industry and Academy in Activities related to or derived from Astronomy and Large Astronomical Observatories in Chile*. Chilean Ministry of Economy. http://www.economia.gob.cl/wp-content/uploads/2012/06/OPORTUNIDADES-ASTRONOMIA-EN-CHILE-INFORME-FINAL.pdf (in Spanish) (2012)
[9] Borne K. Virtual Observatories, Data Mining, and Astroinformatics. In H.E. Bond (ed.), *Planets, Stars and Stellar Systems*, volume 2, pages 409-443. Springer, Dordrecht, (2013).
[10] Hanisch R., Quinn P. The International Virtual Observatory. Retrieved from http://www.ivoa.net/about/TheIVOA.pdf, (2003).
[11] European Southern Observatory (ESO)/Government of Chile Joint Committee for the development of Astronomy. 10 Years Exploring the Universe. http://www.eso.org/public/archives/books/pdfsm/book_0040.pdf, (2005)
[12] GAVO. German Astrophysical Virtual Observatory. http://www.g-vo.org/pmwiki/Documentation/Documentation (2013).
[13] Tody, D., Plante, R., and Harrison, P., "Simple image access specification version 1.0," IVOA Recommendation 20091111 (2004).
[14] Williams, R., Hanisch, R., Szalay, A., and Plante, R., "Simple cone search version 1.03," IVOA Data Access Layer WG Reccommendation (2008).
[15] Tody, D. et al., "Simple spectral access protocol," http://www.ivoa.net/Documents/REC/DAL/SSA-20080201.pdf. IVOA Standards (2008).
[16] Dowler, P. et al., "Table access protocol version 1.0," IVOA Recommendation, March (2010).
[17] Ochsenbein, F. et al., "IVOA recommendation: VoTable format definition version 1.2," arXiv preprint arXiv:1110.0524 (2011).

Defining Smart City Architecture for Sustainability

Leonidas ANTHOPOULOS
Associate Professor, Business School
TEI of Thessaly 41110 Larissa, Greece
lanthopo@teilar.gr

Abstract. Smart city has been evolved since late 1990s to a rapidly emerged domain, where the academia, the industry and the government have mutual interest in transforming cities to innovation-based sustainable spaces. However, this evolution has come to a critical point of argument, where existing efforts are being developed mainly with public spending, which serve industrial purposes. As such, governments focus on smart city standardization in an attempt to clarify the smart city domain. Such standardization concerns smart city architecture too, which has to serve all potential innovations. This paper aims to define a common smart city architecture, which serves government purposes for innovation and sustainability, while it utilizes experiences from prestigious cases and corresponding theoretical context.

Keywords. Smart city, architecture, standardization, governments.

1. Introduction

Various scholars have attempted to define smart city from different lens [1; 2; 3; 4; 5; 6; 7] and now the smart city domain is close to a common definition, which concern innovation –not necessarily based on the information and communications technologies (ICT)-, which aim to enhance urban performance in terms of people, governance, mobility, economy, environment and living. The smart city domain has emerged tremendously since its initial appearance in the late 1990s and engages al-most all the business sectors. The smart city niche market is estimated to reach the amount of $3 trillion by 2025 and exceed the size of all traditional business sectors. This estimation was grounded earlier by Simon Giles [8] from Accenture, who sees the source of this money on embedded operational efficiency, as well as on new entrepreneurship. Moreover, [9] predicts that the amount of €38.9 billion will be spent on smart cities in 2016 alone. However, today, the smart city market race is led by public investments [10; 11], which show that enterprises are still reluctant to invest directly on smart cities and they seek to secure their entrance with standardization and business models.

Standardization is welcome by governments too, which fund smart city initiatives. As such, almost all international organizations are under the process of developing corresponding standards: International Standards Organization (ISO) [12], British Standards Institute [13], International Telecommunications Union (ITU) [14] and the US National Institute of Standards and Technology (NIST) [15] are only some of the organizations that develop smart city standards, which contain specifications for urban performance of for various solutions that can be incorporated in smart cities (i.e., energy,

water, waste, telecommunications, buildings etc.). However, no standardization has been developed so far, which sees smart city as a whole system, although various scholars and developers have suggested alternatives.

Architecture refers to the abstract representation of a system or structure [16]. As such, smart city architecture defines the organization and interrelation of all potential sub-systems and elements, which deliver all expected smart city services to its audience. The aim of this paper is to deal with the above observation and aims to answer the following research question: RQ1 – *what is the structure of a smart city architecture that could define a corresponding standard?*

The answer to the above question is crucial for both governments and the private sector, which will have a common and agreed "picture" on a smart city and will know how to deliver individual solutions within the urban space. Moreover, governments will hold a precise material to deal with recent arguments, which criticize the use of smart city concept and potential and claim that the smart city is the outcome of vendors' marketing campaigns [17]; others say that smart cities reflect little more than usual urban innovations [18]; while Brown [19] criticizes the whole concept of smart city by questioning their effectiveness.

The remainder of this paper is structured as follows: background section 2 classifies smart city projects. Section 3 presents existing approaches to smart city architectures, according to literature findings and some good practices and concludes on a common architecture. Finally, section 4 contains conclusions and some future thoughts.

2. Background

Smart city has risen from the urbanism phenomenon, according which the proportion of the international population that will live in cities will exceed 70% in 2050. The previously given smart city definition considers smart city [14, 6, 20] to be an urban space with innovative –not necessarily based on ICT- features, which are grouped in the following dimensions:

- *People*: in terms of discovering and meeting today and future requirements;

- *Living*: enhancing quality of life and social coherency, as well as efficiency regarding energy, food, water etc.

- *Environment*: protection, waste and emissions control and resilience against climate change;

- *Governance*: in terms of ensuring urban utility and service availability;

- *Economy*: in terms of sustainable growth and city competitiveness;

- *Mobility*: addressing transportation and traffic management issues.

Other approaches discuss resilience in terms of resistance against natural disasters, pandemics, terrorist attacks, accidents etc. [21], which are considered by the above environment and governance dimensions; innovation in terms of urban innovation [22], which is mainly disruptive, although the smart city becomes steadily a social innovation [23]. The above characteristics illustrate the complex nexus, where the identified six dimensions co-exist and interrelate. Smart Cities can be classified according to the smart

infrastructure type and corresponding development stage [24] to the following categories:

- *Hard infrastructure based:* this category refers to city innovations, which target the efficiency and technological advancement of the city's hard infrastructure systems (i.e. transport, water, waste, energy).

- *Soft infrastructure based:* city innovations, which address the efficiency and technological advancement of the city's soft infrastructure and the people of the city (i.e. social and human capital; knowledge, inclusion, participation, so-cial equity, etc.).

With regard to the city development's stage they're classified in the following groups:

- *New cities* (Greenfield or 'cities from scratch' or 'planned cities'): they concern smart city projects where the entire city is being developed from ground zero, even urban planning addresses the above smart city dimensions and innovative solutions are embedded in the city. Various cases of this type are under development around the world, such as Songdo (South Korea), Tianjin (China), Masdar (United Arab Emirates) etc.

- *Existing cities:* they concern smart city projects where the innovative solutions are installed in existing infrastructure. Representatives of this category concern all the cities, which develop various types of innovative solutions (i.e., Barcelona, Amsterdam, Vienna, Copenhagen etc.).

- *Smart plants:* they concern from-scratch projects, which are developed inside existing cities (i.e., new neighborhoods, new blocks or harbors etc.). Indicative cases of this category concern the Kentucky Harbor, Kista (Stockholm) etc.

The above categories analyze the smart city in the following components:

- Soft infrastructure: people, knowledge, communities

- Hard infrastructure: buildings, networks (transportation, telecommunications), utilities (water, energy, waste)

- ICT-based innovative solutions: both hardware and software solutions, which address the above hard and soft infrastructure.

- Other innovative solutions (beyond the ICT): technological innovation that addresses smart city dimensions (i.e., open spaces, recycling system, smart materials, organizational innovations in government etc.)

- Natural environment: it concerns the physical landscape and the corresponding characteristics, where the city is installed (i.e., ground, forests, rivers, lakes, mountains, flora etc.)

3. Existing smart city architecture approaches

An analysis was performed with literature review, with findings from the following sources: international standards organizations for smart city documents; and SCOPUS, with searches only in journals that publish smart city articles [25], with the combination

of terms "smart city" and "architecture". Article crawl was performed within the period of 1997 (appearance of smart city concepts in literature) to early 2015. More than 200 articles were returned from this crawl, where screening was used to leave out irrelevant publications (like "city architecture").

The most important articles that discuss smart city architecture where from [11, 26, 27, 28]. Anthopoulos and Fitsilis [11] explored various smart cities around the world and concluded that the architecture that is preferred by well-managed managed cases is the multi-tier (Table 1), which is applied in new, existing and smart planting cases, while it addresses both soft and hard infrastructure, while it considers natural environment and the evolving Internet-of-Things (IoT) in terms of sensor installation.

Module definition for a smart city is an extremely complex process and it has to consider both the type and the architecture. According to the above analysis, soft urban infrastructure (people, data and applications) is flexible and can easily extend and interconnect. Difficulties rise from requirements, which deal with hard infrastructure and environment. Various attempts illustrate modular smart city approaches [26, 27, 28]. A modular architecture approach to smart city has been inspired from Al-Hader et al. [27] and can be utilized to the following:

1. Networking Infrastructure and Communications Protocol: this module addresses the necessary infrastructure to deploy smart services and enhance living inside the city. Cities from scratch are based on innovations (both ICT-based and non ICT-based), which are embedded on city's hard infrastructure. For instance, in New Songdo (Seoul) a waste disposal, recycling and tele-heating factory is installed and interconnected with buildings inside the city (Clever rubbish). In the same case, fiber-optic networks connect all local buildings with a central operating center, while smart buildings are accessible by their inhabitants via specific applications. In existing cities on the other hand, corresponding SSC cases integrate innovation with existing hard infrastructure with the IoT and basically with sensors that exchange data with specific applications. Moreover, protocol defines the codification for information inter-change in SSC.

2. *Applications*: this module concerns all the smart applications, which are available inside the smart city. A well method for analyzing this module could be the classification of applications in the four smart city dimensions, including a separate group of mobility (i.e., intelligent transportation applications).

3. *Business*: it addresses business groups, which are available in the city.

4. *Management*: this module contains all rules and procedures for managing a smart city: processes, people, resource, land and information are the primary elements and could be controlled centrally or individually with the appropriate set of standards.

5. *Services*: this module concerns all type of smart city services, offered with the contribution of the ICT from the supplier-side users (smart city stakeholders' users) and requested by the demand-side users (inhabitants and smart city stakeholders' users). Smart city stakeholders have been identified by [30] and concern organizations from the urban area (utility providers, NGOs, city service companies etc.); governments (local, state and national); international and multilateral organizations; citizens; the academia; urban planners and standardization bodies. On the other hand, services concern: transportation; e-government; e-business; safety and emergency; smart health; tourism; education; smart building; waste management; smart energy; and smart water.

An indicative n-tier ICT meta-architecture, where physical, utility and ICT environments coexist and interact, while people and businesses are also part of the eco-system and interact with the smart city via e-services is illustrated on (Figure 1). These 5 layers were chosen as the result from the above analysis, as well as an attempt to

address all United Nations Habitat (UN Habitat) key-performance indicators [14]. Service groups were also selected to meet these indicators accordingly.

Another architectural approach concerns the Service Oriented Architecture (SOA) [20], which is proposed for existing cities, where innovation mainly focuses on soft infrastructure, as well as where IoT is utilized [29]. Finally, event-driven architecture (EDA) is also discussed [11], but it has not been applied yet.

Table 1. Preferred architectures in various examined cases [11]

Case	Findings	
	Architecture	**Organization**
European Smart Cities	Urban Intelligence Measurement System	Project (various European Cities)
Two cities in Netherlands	SOA	State-Owned-Enterprise (SOE) run by the municipality
52 cities	n-tier architecture (4 layers): *Network, Content, Intelligence, e-services*	Public Organization (i.e., Gdansk (Poland), Masdar (UAE)) Public Private Partnership (PPP) (i.e., Amsterdam (Netherlands)) Private Companies (Malaga (Spain), New Songdo (Korea))
Helsinki, Kyoto	n-tier architecture (3 layers): *information, interface, interaction*	State-Owned-Enterprise (SOE) run by the Municipality
Dubai	n-tier architecture (3 layers): *Infrastructure, data, application*	Public Organization (Government)
Trikala, Greece	n-tier architecture (6 layers): *data, infrastructure, interconnection, business, service and user*	State-Owned-Enterprise (SOE) run by the Municipality
Barcelona	n-tier architecture (4 layers): *code, nodes, infrastructure and environment*	SOE run by the Municipality in cooperation with the local university
Blacksburg Electronic Village	n-tier architecture (3 layers): infrastructure, content, community	PPP between Bell Atlantic Telecoms, Virginia Tech, Municipality
Amsterdam	n-tier architecture	PPP between Municipality and Liander grid Operator
Singapore	n-tier architecture (4 layers): ICT infrastructure, Cognitive infrastructure, Services, Customers	Public Organization

Table 1 shows that architecture is independent to the smart city organization (Public organization, State-Owned-Enterprise (SOE), Project coalition or Private Company),

since the multi-tier architecture for instance is observed in various cases that follow alternative organization forms. On the other hand, layer selection in multi-tier architectures is not influenced by the ICT smart city selection (i.e., Trikala is a digital city, Kyoto is an online city etc.).

All the above information collected from literature and case studies, provide with answer the RQ1 and suggest that the architecture of the smart city must be multi-tier in order to be clear and sustainable, in terms of standardization and communication of these standards. This architecture (Figure 1) meets existing standardization efforts [12;13] and more specifically the UN Habitat key-performance indicators [14], which define a model for urban measurement at an international level. According to the examined cases, this n-tier architecture must utilize hard and soft infrastructure and must contain the minimum following layers (Figure 1) from top to bottom:
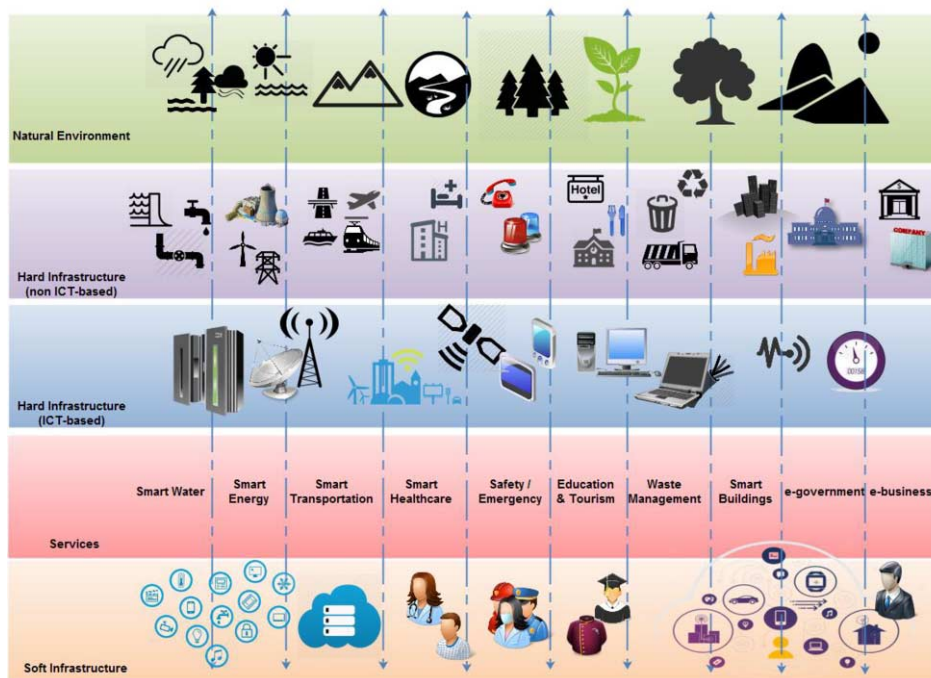
Figure 1. A generic multi-tier ICT architecture for smart city [31]

Layer 1) Natural Environment: it concerns all the environmental features where the city is located (landscape, rivers, lakes, sea, forests etc.).

Layer 2) Hard Infrastructure (Non ICT-based): it contains all the urban features, which have been installed by human activities and are necessary for city operation (buildings, roads, bridges, energy-water-waste utilities etc.)

Layer 3) Hard Infrastructure (ICT-based): it concerns all smart hardware, with which SSC services are offered (i.e., datacenters, supercomputers and servers, networks, IoT, sensors etc.)

Layer 4) Services: all types of smart city services, grouped in the smart city six dimensions and organized according to international urban key-performance indicators.

Layer 5) Soft Infrastructure: individuals and groups of people living in the city, as well as applications, databases, software and data, with which the SSC services are realized.

4. Conclusions and future thoughts

This paper addressed an important problem regarding the lack of standardization, which would consider smart city as a system and grounded a research question (RQ1), regarding the type and the structure of a smart city architecture that could define a corresponding standard.

In order to provide with answer RQ1, this paper used literature findings and combined them with data from well-known smart cities. Findings suggest that a common architecture must be multi-tier, consisting of five layers (natural environment, hard non-ICT infrastructure, hard ICT-infrastructure, services and soft-infrastructure). These layers address all potential smart city solutions, while it leaves space to incorporate services that are delivered across the urban space and grouped according to UN Habitat key-performance indicators. This architecture can be the baseline for smart city standardization, since it adopts internationally defined smart city requirements, it can fit to all city types (existing cities, new cities and city blocks) and it is easy to follow architecture principles like scalability, interoperability, security and vendor independence etc. [31]. However, limitations come from the missing of this architecture testing and corresponding validation. Future thoughts concern the testing of this architecture in alternative cases across the globe or the validation with responses from smart city experts.

Acknowledgements

References

[1] P. Neirotti, A. De Marco, A.C. Cagliano, G. Mangano, and F. Scorrano, Current trends in Smart City initiatives: Some stylised facts, *Cities* **38** (2014), 25-36.
[2] G. Piro, I. Cianci, L.A. Grieco, G. Boggia, and P. Camarda, Information centric services in smart cities, *The Journal of Systems and Software* **88** (2014), 169-188.
[3] K.C. Desouza, and T.H. Flanery, Designing, planning, and managing resilient cities: A conceptual framework, *Cities* **35** (2014), 88-89.
[4] W.-M. Wey, and J. Hsu, New Urbanism and Smart Growth: Toward achieving a smart National Taipei University District, *Habitat International* **42** (2014), 164-174.

[5] J.H. Lee, M.G. Hancock, and M-C Hu, Towards an effective framework for building smart cities: Lessons from Seoul and San Francisco, *Technological Forecasting & Social Change* (2013) http://dx.doi.org/10.1016/j.techfore.2013.08.033

[6] R. Giffinger, C. Fertner, H. Kramar, E. Meijers and N. Pichler-Milanovic, *Smart Cities: Ranking of European medium-sized cities*, 2007 [Online] http://www.smart-cities.eu/download/smart_cities_final_report.pdf

[7] H. Chourabi, N. Taewoo, S. Walker, J.R. Gil-Garcia, S. Mellouli, K. Nahon, T.A. Pardo, and H.J. and Scholl, Understanding Smart Cities: An Integrative Framework. *45th Hawaii International Conference on System Science (HICSS)* (2012), 2289-2297.

[8] M. Morris, M. Schindehutte, J. Richardson, J. and Allen, Is the business model a useful strategic concept? Conceptual, theoretical, and empirical insights, *Journal of Small Business Strategy* **17** (2006), 27-50.

[9] CISCO, The City of The Future, 2014, [Online] http://share.cisco.com/TheCityOfTheFuture/the-city-of-the-future.html?_ga=1.115393729.630037314.1415774790

[10] L. Anthopoulos, and P. Fitsilis, Smart Cities and their Roles in City Competition: a Classification, *International Journal of Electronic Government Research (IJEGR)* **10 (1)** (2014), 67-81.

[11] L. Anthopoulos, and P. Fitsilis, "Exploring Architectural and Organizational Features in Smart Cities, *16th International Conference on Advanced Communications Technology (ICACT2014)* (2014).

[12] International Standards Organization (ISO), *ISO 37120:2014 - Sustainable development of communities - Indicators for city services and quality of life*, 2014.

[13] British Standards Institute, *PAS 180 Smart city terminology*, 2014 [Online]. http://www.bsigroup.com/smart-cities/Smart-Cities-Standards-and-Publication/PAS-180-smart-cities-terminology/

[14] International Telecommunications Union (ITU), *Overview of key performance indicators in smart sustainable cities*, 2014 [Online]. http://www.itu.int/en/ITU-T/focusgroups/ssc/Pages/default.aspx

[15] National Institute of Standards and Technology, *Global City Teams Challenge Kick-off,* [Online]. Available: http://www.nist.gov/cps/global-city-teams-challenge-workshop.cfm

[16] J. McGovern, S. W. Ambler, M. E. Stevens, J. Linn, V. Sharan, and E. K. Jo, *A Practical Guide to Enterprise Architecture*, Prentice Hall PTR, 2003.

[17] O. Söderström, T. Paasche, and F. Klauser, Smart cities as corporate storytelling, *City: analysis of urban trends, culture, theory, policy, action* **18:3** (2014), 307-320, DOI: 10.1080/13604813.2014.906716.

[18] T. Nam, and T. Pardo, Smart city as urban innovation: Focusing on management, policy, and context, *5th international conference on theory and practice of electronic governance* (2011), Tallinn, Estonia.

[19] L.A. Brown, The city in 2050: A kaleidoscopic perspective, *Applied Geography* **49** (2014), 4-11.

[20] IBM Institute for Business Value, *How Smart is your city? Helping cities measure progress*, 2009. [Online] http://www.ibm.com/smarterplanet/global/files/uk__en_uk__cities__ibm_sp_pov_smartcity.pdf

[21] B. Pizzo, Problematizing resilience: Implications for planning theory and practice, *Cities* **43** (2015), 133-140.

[22] A. Khare, T. Beckman, T., and N. Crouse, Cities addressing climate change: Introducing a tripartite model for sustainable partnership, *Sustainable Cities and Society* **1** (2011), 227-235.

[23] T. Dixon, M. Eames, J. Britnell, G.B. Watson, and M. Hunt, Urban retrofitting: Identifying disruptive and sustaining technologies using performative and foresight techniques, *Technological Forecasting & Social Change* **89** (2014), 131-144.

[24] M. Angelidou, Smart city policies: A spatial approach, *Cities* **41** (2014), S3-S11.

[25] L. Anthopoulos, "Understanding the smart city Domain: A Literature Review," in Bolivar M.P. (Ed) *Transforming City Governments for successful Smart Cities, Public Administration and Information Technology Serie*s, Vol. 3, Springer Science+Business Media New York, 2015.

[26] G. Kuk, and M. Janssen, The Business Models and Information Architectures of Smart Cities, *Journal of Urban Technology* **18(2)** (2011), 39–52.

[27] M. Al-Hader, A. Rodzi, A.R. Sharif, and N. Ahmad, N. "Smart City Components Architecture". *International Conference on Computational Intelligence, Modelling and Simulation*, IEEE, 2011.

[28] G. Kakarontzas, L. Anthopoulos, D. Chatzakou, and A. Vakali, "A Conceptual Enterprise Architecture Framework for Smart Cities: A Survey Based Approach". In *ICE-B 2014 - International Conference on e-Business is part of the 11th International Joint Conference on e-Business and Telecommunications*, Vienna, Austria; 2014, SCITEPRESS.

[29] E. Bucherer, and D. Uckelmann, Business Models for the Internet of Things. In Uckelmann et al. (Eds) *Architecting the Internet of Things*, Springer Heidelberg Dordrecht London New York, 2011.

[30] International Telecommunications Union (ITU), "Technical report on "Engaging stakeholders or smart sustainable cities," 2014 [Online]. http://www.itu.int/en/ITU-T/focusgroups/ssc/Pages/default.aspx

[31] International Telecommunications Union (ITU), Technical Specifications on "Setting the framework for an ICT architecture of a smart sustainable city" (SSC-0345), 2015 [Online] http://www.itu.int/en/ITU-T/focusgroups/ssc/Pages/default.aspx

Cloud for eGov: the state of the art

Francesco DE ANGELIS, Roberto GAGLIARDI, Alberto POLZONETTI, and
Barbara RE
University of Camerino

Abstract. Specifically, the paper provides a detailed analysis of the state of the art regarding technological, regulatory and interoperability aspects, devoting particular attention to those projects (and related experience) in the field of cloud and distributed computing. The paper's aim is to evaluate benefits and risks of the applicability of solutions that are already in place for the specified areas of interest. The key factors that determine relevant regulations, with particular emphasis on the European target, have been listed in the paper. Finally, the main Italian projects and the situation in some European countries have also been mentioned.

Keywords. Cloud distributed computing, egov

1. Introduction

Any models for cloud computing implementation are bound to feature advantages and disadvantages. The model choice depends on the intended usage of the model itself and on the related requirements. Despite the huge saving of resources and the high potential, the widespread use of a "public cloud" is affected by a number of constraints, such as data control requirements or legal compliance issues. Especially in the e-Government field, criteria such as safety or data protection do play a crucial role.

Several studies have been carried out to investigate the use of cloud computing for public services [10]. Most studies recommend the private use of a cloud for public authorities and public clouds usage for the community. For example, because of the heterogeneity of national laws and regulations, also ENISA recommends these two models for e-Government applications. Basically, they state that "in terms of architecture, private applications and community clouds appear to be the most suitable solutions that are currently available to fulfill the needs of Public Administration bodies (PA), as they offer the highest level of data management, control and visibility " [4].

This paper aims to provide an analysis of the state of the art in relation to regulatory, technological, and interoperability aspects, with regard to existing projects and experiences of cloud computing.

Section 2 provides an analysis of the regulatory aspects. It may be useful to note that the cloud is a relatively new phenomenon, as well as a complex and structured one. Therefore, a careful study of the existing legislation and of similar legal implications may provide useful insights both in relation to the ability to introduce innovations fully supported by current regulations, and to guiding lawmakers in regulating new ones. Another important aspect that should be considered is the study of different types of contracts that may be suitable to define the underlying relationships between stakeholders and, consequently, to indicate the most appropriate framework to fulfil the

interests of the parties involved. A general overview will be provided concerning cloud application scenarios and/or some single aspects of the matter in question, such as open data, transparency, interoperability, federate identity, etc..

As regards the design experiences gathered from national and international projects and described in Section 3, we do not intend to provide a comprehensive guide of the state of the art, but rather to identify different approaches, success stories and solutions that are being developed or are already in place across Europe, and beyond, by providing information about the technical, economic, legal and political perspectives.

2. Analysis of regulatory aspects

Nowadays, many people argue that the introduction of increasingly innovative technology is changing economies and the ideological foundations of society. According to Peter Senge, theorist of the Learning Organization (P. Senge, "The Fifth Discipline: The Art and Practice of the Learning Organization." New York, 1990), the focal point in the evolution of the bases of modern economy is to be found in the replacement of fixed assets (as driving force for development) with information and knowledge. Moreover, that was accompanied by momentous changes that have characterized the western world in the last fifty years, such as the end of the welfare state, the fall of the Fordist model and globalization.

The result is the need for a "governance of change" in order to manage and fully implement the information and knowledge society. Governing the change also means setting forth guidelines aimed at ensuring stability and balance within society, as well as highlighting and exploiting the new opportunities offered by the development of the new society. The digital revolution we have been witnessing is going to transform the existing society into a "post-industrial" one, where technology development, information dissemination and knowledge sharing are going to play a decisive role.

The IT market occupies a significant position within the European economic scenario. According to Eurostat data,[1] the IT sector is amongst the top fifteen industries, in terms of both workforce and added value. When properly used within the public administration domain, Information Technology may provide an important element of selective regularization for the economic system.

Within such a scenario, besides promoting the public sector's efficiency and effectiveness, an e-Gov cloud project could be introduced as a proactive lever from the business world through the identification of initiatives to be implemented and to be developed with reliable timing and methods, in order to take advantage of the opportunities offered by the digitization process. Such initiatives may include: (i) opening of new market opportunities for ICT companies, products and services, which may also promote employment, in line with international trends; (ii) development of new ICT services to be acquired from the market to ensure the internal efficiency of companies and a better service for users.

As a growing phenomenon within the national and European scenario, Cloud Computing, seen as a set of technologies facilitating the use of IT resources (infrastructures, platforms, processing power, applications), such as services provided

---

[1] http://ec.europa.eu/eurostat/data/browse-statistics-by-theme

through Internet, is the great promise in the ICT world. This seems to have the potential to modernize the IT systems of Italian companies and PA bodies, besides making them less expensive and more efficient too, and to encourage small and medium enterprises to use IT more intensively, stimulating, therefore, an expansion of the ICT market.

Hence, Cloud Computing is going to be one of the most interesting scenarios within the public sector in the coming years. To rationalize and contain expenditure, PA bodies could start to develop a Cloud process, by responding to stimuli coming from the European Union within the G-Cloud. In March 2010, the European Commission tried to understand the state of "GovernmentCloud Computing" in Europe (amongst the members of the European CIO Network – through an informal roundtable for national representatives in the Management Committee of the European ISA programme - Interoperability Solutions for European Public Administrations[2]). The results obtained were as follows:

- Four EU member states already have a GovernmentCloud Computing strategy in place;
- Half of the member states stated that they had already used the Cloud for some centrally provided public services;
- Over half of the member states foresee to have some pilot projects ready by 2010-2011, national strategies in place by 2011 and public services in operation by 2012.

In actual fact, the route towards the Cloud is not free from obstacles. There are still many issues that may slow down the implementation of such systems. The main ones are related to protection and safety of the data stored in the Cloud, to a guarantee of service continuity, to regulations (that may not be the same as those in the country hosting the service), and to the actual interoperability/integration with internal systems. Other uncertainties are related to the types of contracts and to the service levels to be agreed upon, to the few best practice policies to be used as a comparison, to the fear of limitations imposed by third parties for critical activities, resulting in an excess of bargaining power that providers may enjoy. Through an analysis of the Italian context, the main problem to overcome seems to be the presence of a culture based on the physical possession of data, which hinders cooperation between the various PA bodies, in addition to some particular regulatory constraints. [2,9].

In addition to the above-mentioned concerns, there remains an opportunity for all businesses and PA bodies, whether large or small, to become more efficient and closer to their citizens. The Cloud concept is going to be increasingly pervasive and the "Cloud Economy" (i.e. productive systems based on Cloud usage) is a topic that's already being discussed. According to the CloudDividend report, in Italy, the benefits obtained through the Cloud Economy by 2015 were estimated as amounting to Euro 35 billion, in terms of increased productivity. [1,5]

---

[2] http://www.eurocio.org/

3. The current regulatory framework by areas of interest

3.1. Data and availability

This is the main problem that is encountered when any cloud computing activities/services are to be carried out or provided. Further issues are also linked to that, such as those concerning data access, sharing, free data usage by end users, besides the implications related to recent phenomena such as open data, open access, open government and all the "open" movements that, by now, seem to belong to the entire PA apparatus. This means that, nowadays, it's just impossible to disregard the need for instant access to data (either from a timing point of view or in relation to any intermediary intervention between users and data) and the ease of data "withdrawal", whenever requested by the user. [3,8]

With regard to the context illustrated above, the relevant basic regulations are listed here following:

a) Regulations concerning the plain and simple "Data Protection".

    i.   In Italy, general data protection has to be specified pursuant to the main relevant regulation, i.e. the Legislative Decree No. 196 of 2003 (the so-called "Privacy Code").

    ii.   In Europe, to date, it is still necessary to specify the compliance with the EC Directive 2006/24/EC on data storage and communication services (the so-called "*Data Retention*").

b) Data handling and safety.

    i.   In Italy, the basic relevant regulation is still the "Privacy Code" mentioned above, with some further requirements specified in the section concerning safety (see relevant section).

    ii.   In Europe, reference is made to the EC Directive 2006/24/EC on the Data Retention phenomenon.

c) Data accessibility, as regards the ease and speed of use.

    i.   In Italy, the accessibility of data and platforms, or of IT systems involved in the handling of data, is governed by the so-called "Stanca Law" (Law No. 4/2004) and by the Digital Administration Code (the Italian CAD) (Legislative Decree No. 82/2005), especially as regards PA websites.

    ii.   In Europe, the number of EU's accessibility policies continues to grow[3], together with many further references, such as the e-inclusion campaign, the EU's e-inclusion website, with some sections devoted to mature users, e-accessibility, broadband divide, e-Government for everybody's benefit, IT literacy and culture; in addition to the progress made in the field of e-accessibility in Europe.

---

[3] http://europa.eu/geninfo/accessibility_policy_it.htm

d) An additional problem concerning cloud computing may be the one originating from an excessive fragmentation of data, or rather from their extreme "delocalization". In fact, it may happen that data which have been "sprayed" somewhere else on the web have to be brought back to the local server. In such circumstances the regulatory function should also tackle the delicate aspect of interrupting business dealings with a given provider (in this particular case, with the provider-administrator of the "cloud" web server).

The most relevant regulatory tool for such cases, at both national and European level, remains a written contract, which should be particularly articulated and detailed and suitable to govern the relationships to be established between the parties involved in relation to the cloud computing service.

e) Another aspect (still related to data and their "usability") is the open data one. In this case too, there are regulatory indications by Italian and European lawmakers (and Italian lawmakers should adjust to the European provisions).

    i. In Italy, the emerging phenomena concerning open data, openGovernment etc. are regulated by the following laws: first of all, by Art. 97 of the Italian Constitution, which states the principle of good administration; then by Law No. 241/1990, which governs, amongst other things, the so-called access right; by the Legislative Decree No. 82/2005 (CAD); and by some other special regulations.

    ii. As regards Europe, the EC Directive 2003/98/CE on the re-use of information within PA bodies can be referred to.

3.2. Data privacy

Moreover, it should be pointed out that, in Italy, the Privacy Guarantor has tackled the cloud computing phenomenon through some special provisions. A particularly interesting one was titled "**Cloud computing**: indications for an informed use of services" (*"**Cloudcomputing**: indicazioni per l'utilizzo consapevole dei servizi"*) and was published on the following website: www.garanteprivacy.it/garante/document?ID=1819933 (please refer to such a provision).

At European level, several regulations have also been issued (data privacy directives). Amongst them the "old", forerunner Directive on data privacy stands out (95/46/CE) and, in spite of everything, it still offers a kind of legal protection that not all the countries outside the EU are able to ensure.

3.3. Copyright protection

Another "industry" that cloud services can reach is that concerning the distribution of data for which intellectual property rights can be claimed. In this case, a reasonable risk exists about difficulties in controlling user access to such data, as currently defined by law.

For example, that has already occurred with the Google book service which is quite "**cloudy**", (see http://books.google.it/). The service is very interesting and complex because of the implications it entails. The most immediate problem is the one associated with possible copyright infringement. Google's aim was to provide a tool for the promotion of books and not to break intellectual property laws, but the whole matter needs to be ascertained. In this context, the reference standards are as follows:

> i. In Italy, the **copyright law (Law No. 633 of 1941, better known in Italian by the acronym "l.d.a.", which aims at governing copyright infringements)** and an Industrial Property Code (Legislative Decree No. 30 of 10th February, 2005) that was developed to provide a general protection for intellectual/industrial property. The protection of the intellectual property and of the copyright holder is also widely "spread" over the Civil Code (Art. 27).
>
> **ii.** European legislation on copyright is part of the regulations set forth by the World Intellectual Property Organization (WIPO), which was created to achieve a double aim: protecting the financial interests of authors of art works, such as books, films, music, as well as those of databases, without hampering creativity and innovation.

3.4. Education and information rights

Another area where cloud computing services could be used is the education and training one. In Italy, cloud services have recently started being provided to students at Rome's La Sapienza University and University of Camerino. In fact, cloud technology may be used by university all over the world, as it's already happening in the United States where 66 out of 100 most prestigious universities, including Berkeley and Harvard, have implemented Google Apps for Education. Future applications may involve sharing **e-learning** platforms and **cloud computing** devoted to education. Reference regulations are listed here following:

- In Italy, the Constitution works as "guarantor" of personal rights:
    - o Right to inform and be informed (Art. 21);
    - o Right to education – research and teaching (Art. 33);
    - o Right to full personal development (Art. 2 – so-called "open provision" that should also include the recent right to Internet).
- In Europe, relevant regulations are mainly based on the Universal Declaration of Human Rights (UDHR) issued by the United Nations in 1948.

3.5. Safety

With regards to IT safety, our country has implemented the provisions against IT Crimes a long time ago, in compliance with European Directives already in place.

Therefore, the relevant regulations set forth by lawmakers to govern IT system violation issues may be summed up as follows:

      i. In Italy, Law No. 48 of 2008, Legislative Decree No. 109 of 30th May, 2008 – implementation of the Directive on data retention;

      ii. In Europe, the Budapest Convention of 23rd November, 2001, in addition to the above-mentioned Directive No. 24 of 2006 on data retention (…).

4. Existing solutions for PA bodies and best practices

4.1. The Italian experience with SmartGov, territory and taxation

As regards SmartGov, territory and taxation tools available online, a short description of some projects and related experiences is provided here following.

Mcloud Project[4]: Marche's Regional Government has been working for some time on an advanced cloud computing infrastructure named "Software As A Service", in cooperation with the National Nuclear Physics Institute (CNAF at Bologna and Perugia), the University of Camerino and the Polytechnic University of Marche (Information Engineering Department). Within such a context, the Mcloud pilot project was deployed in order to evaluate the most effective method to implement the cloud system and, at the same time, to provide a safe service for electronic medical reporting to patients and regional health system users.

CloudToscana Project[5,6,7]: Through the development of the new TIX (Tuscany Internet Exchange), Tuscany's Regional Government has achieved a most important goal, i.e. creating a new service oriented centre. The project's guiding principle was a vision based on the XaaS paradigm that would allow for the development of a solid bridge between the management of supply and demand for ICT by the Regional Government to satisfy its stakeholders.

CleanMoon Project[8]: The Clean MooN Project (Clean Mooring Network) OLBIA has been primarily applied within the *smart culture* and tourism sectors. In fact, it tackles the issue of handling the international nautical tourism in favour of cities which have tourist harbours. The project is aimed at offering special tourist routes that enhance the value of the surrounding territory by highlighting its peculiarities, typical products, culture and art. Local economy and companies are going to benefit from the project and, thanks to a smart governance of the tourist flow, the quality of hospitality and of the environment can be improved too, so that the tourist season may last longer and be more profitable.

---

[4] http://www.ecommunity.marche.it/AgendaDigitale/MCLoud/tabid/156/Default.aspx

[5] http://open.toscana.it/web/cloud

[6] https://www.tix.it/web/guest/Cosa-offre-il-TIX

[7] http://open.toscana.it/web/cloud

[8] http://nuke.valorizza.info/Portals/0/Monografie/Clean%20Moon.pdf

Cloud4eGov Project[9]: The project's mission is developing a cloud computing platform to enable the association and interoperability of open or proprietary cloud solutions, in order to support new means of interaction between citizens and PA bodies. Another goal of the project is to offer PA bodies an environment that should allow for the development of eGovernment applications for citizens and companies, and for the subsequent opening of the service to the outside world without having to set up dedicated IT and TLC infrastructures.

Prode Project[10]: The name is a synonym for "Dematerialization Project". This is an interregional project aimed at defining a dematerialization system model. The PA bodies involved have developed a first vision of the issue and, at a later stage, have also designed an actual project to define and share the specific techniques, as well as the subsequent planning of activities to be carried out throughout the territory in every Autonomously Governed Region and Province district, so that individual systems could adjust to the shared specifications.

4.2. National experience in the field of infomobility and video surveillance

Vi.So.Re. Project - Treviso[11]: The project includes the development of two integrated sub-systems – one for video surveillance, for both municipal authorities and the State Police, and one that allows reading car registration plates which is used by the State Police. Such systems enable users to collect, exchange and analyze information through a connecting, integrated network. It is just the connection between the two sub-systems that makes the project unique, to the point that it has become an experimental model nationwide.

Cagliari 2020 Project[12]: It is an initiative aimed at optimizing city transport, by reducing traveling time and, above all, by improving air quality. Such a project (that relates to smart cities and communities efforts) also involves Vitrociset (a company that is partly owned by Finmeccanica, and is specialized in innovation), Space Spa (Pmi), University of Cagliari (Electronic Engineering Department), CTM (the local public transport company), under the sponsorship of Cagliari's municipality, INFN (National Nuclear Physics Institute).

INFOCITY Project[13]: The project's objective is to develop an open, multimode, infomobility system that should be independent, integrated and interoperated, to provide statistical and dynamic information before and during the trip (pre-trip and on-trip), in relation to the overall transport offer and traffic conditions. The information

---

[9] http://www.toscana-notizie.it/documents/735693/1369400/Cloud4eGov.pdf

[10] http://www.progettoprode.it/

[11] http://www.infracom.it/it/categoria-comunicati-stampa/145-infracom-italia-l-hi-tech-della-videosorveglianza-evoluta-si-aggiudica-il-progetto-vi-so-re-di-treviso-insieme-a-kapsch-trafficcom

[12] http://www.unica.it/pub/12/show.jsp?id=18989&iso=913&is=12

[13] http://www.progettoinfocity.it/

should be provided in real time, and should be georeferenced and customized according to the actual needs of end users.

4.3. National experiences in the field of "Health"

Here following a brief description is provided of those projects, and related experiences, concerning the health sector and which are available online.

Mcloud Referti Service[14]: It is a reference pilot service within the Mcloud Project implemented by Marche's Regional Government. The service allows users to view the results of clinical tests carried out at laboratories belonging to the Regional Health Authority network and to the Ancona's Hospital Corporation.
G-Cloud Lazio Project[15]: The experimental phase of G-Cloud for PA bodies has been a technological opportunity that, by superseding expensive data-centres and by making the basic application platforms converge, facilitates the consolidation and concentration of IT resources. This translates into significant financial benefits for both local authorities and small-medium sized companies. It is available to Lazio's Regional Government bodies, companies and other bodies partly owned/controlled by Lazio's Regional Government, and Lazio's 12 Local Health Authorities.
Cloud4CARE Project[16] (Cloud4CAncerREsearch): It was the first "scientific-banking cloud project" implemented in Italy, in order to support cancer research through safe network connections and IT systems running within safe environments. It is a flexible and scalable solution that, since its first phase, has been enabling users to significantly reduce (by about 90%) the time devoted to the analysis of data concerning a research on ovarian tumors, thanks to faster calculation speed (around 4,000 CPU and one terabyte of RAM).

5. Conclusions

This paper analyzes the state of the art of the components of open-source cloud applications and of Smart City governance through an analysis of solutions and synergies amongst existing projects. Our investigation focused on specific interest areas, and on the analysis of issues, benefits and difficulties found in each area.
The analysis of regulatory aspects was articulated by considering two levels of legislative reference, i. e. the European and the national ones. With regard to the latter, special attention was devoted to the implementation of the European Digital Agenda and the related technical and technological solutions developed in pilot projects within individual local PA authorities.

---

[14] http://www.ecommunity.marche.it/AgendaDigitale/MCLoud/tabid/156/Default.aspx

[15] http://www.lazioeuropa.it/45_progetti_per_il_lazio-3/consolidamento_e_razionalizzazione_dei_data_center_regionali-35/

[16] http://www.01net.it/cloud4care-al-mario-negri-nuvola-privata-per-ricerca-pubblica/0,1254,1_ART_153999,00.html

Regulatory issues concerning the relatively new phenomenon of cloud computing (especially if involving PA) are going to evolve and change over time. Therefore, only a thorough and timely analysis of current regulations may guide lawmakers in their efforts to define new laws, and IT providers in the development of new services.

Over the last decade, several projects concerning distributed computing infrastructures (DCI) and cloud computing have been financed by national or European funds (with the involvement of various countries). In this paper, such projects were analyzed with the specific objective of collecting experiences and solutions that may be useful to make future projects successful.

To conclude, by analyzing different aspects of technology and regulatory issues concerning existing cloud computing solutions for PA bodies, this paper aimed at providing a solid starting base for the evolution of later projects and for the advancement of the respective Relational Objectives, by doing what follows:

- favouring the re-usage of standards and technological components;
- verifying whether current regulations may be suitable and up to date;
- highlighting experiences and solutions from other projects.

References

[1]   Banaian, K., & MacDonald, R. A. (2014). St. Cloud Area Quarterly Business Report, Vol. 16, No. 1.

[2]   Barras J, Brown P.F., Parker C., Transformational Government framework Version 2.0, 1 May 2014.

[3]   Chapman M., Durand J., Otto A., Piltz G., Rutt T., Cloud Application Management For Platforms Version 1.1, 27 August 2014.

[4]   ENISA. (2011). Security & resilience in Governmental clouds

[5]   Modica, G. D., & Tomarchio, O. (2015). A semantic framework to support resource discovery in future cloud markets. International Journal of Computational Science and Engineering, 10(1), 1-14.

[6]   Pérez San-José, P., de la Fuente Rodríguez, S., García Pérez, L., Gutiérrez Borge, C., & Álvarez Alonso, E. (2012). Study on cloud computing in the Spanish public sector (p. 178).

[7]   Robinson, N. (2010). Computing in the public sector: Rapid international stocktaking.

[8]   Sabo J., Willett M., Brown P.F.,. Jutta D.N., Privacy Management Reference model and Methodology (PMRM) Version 1.0, 26 March 2012.

[9]   Saldhana A., Nadalin A., Rutkowski M., Identity in the Cloud Use Cases Version 1.0, 8 May 2012.

[10]   Wyld, D. C. (2010). The cloudy future of government IT: Cloud computing and the public sector around the world. International Journal of Web & Semantic Technology, 1(1), 1–20.

Policy Compass: FCM-based Policy Impact Evaluation using Public Open Data

Youngseok CHOI, Habin LEE[1] and Uthayasankar SIVARAJAH
College of Business, Arts and Social Sciences, Brunel University,
Youngseok.Choi@brunel.ac.uk, Habin.Lee@brunel.ac.uk,
Sankar.Sivarajah@brunel.ac.uk

Abstract. This paper presents how Fuzzy Cognitive Map (FCM) technique can be applied by a policy maker to support the policy impact evaluation using the example of Policy Compass, an EU research project. The practical usage example on interest rate policy shows the potential of FCMs as a policy impact modelling tool. Through the provision of a more intuitive and easier means of using open data based on FCM techniques, the Policy Compass project can play a critical role for both policy maker and lay public to evaluate the policy impact and prepare for future policy making.

Keywords. Open Data, Fuzzy Cognitive Map, Policy Impact Modelling, Policy Impact Evaluation, Policy Compass

1. Introduction

The demand for open data analytics to support policy making is-increasing along with the growing importance of utilizing big and open data. Developments in open data solution and platform are used to open up governments, engage stakeholders, creates value from the huge amounts of data, and inform decision and policy making. Policy impact and evaluation is one of the representative areas that can benefit from data analytics. Open data such as historical records, social indicators, and even survey results from citizens can be an important standard to measure the success of past policies and be the basis to develop future policies. Since the United Kingdom (UK) government introduced the concept of evidence-based policy making in the 1990s [1], the importance of the use of public open data as evidence to support the decision on and evaluation of policies has been emphasized as significant more than ever.

In this regard, policy making has become a more complex process that needs to consider environmental and political variable factors. Policy makers are now are under the situation where they should check not only the political dynamics, but also the evidence of the past policy based on enormous data for their future policy making. This makes analysing the policy impact more difficult and complex. However, analytics tool or platform for facilitating this complex policy making situation can be hardly found in the research area as well as existing project.

Given the lack of availability of such tools, Policy Compass aims to be an innovative analytics platform for policy impact evaluation and potential policy making.

---

[1] Corresponding Author.

Based on Fuzzy Cognitive Maps (FCMs), Policy Compass seeks to enable the policy maker and the lay public to evaluate the impact of a causal policy model according to various open data in an intuitive manner. As FCMs are particularly suited to modelling complex and dynamic social problems [4], it has the potential to transform an interactive policy ecosystem with vague conceptual terms into a specific concrete form (i.e., the FCM) that can be readily understood by both policy maker and the general user.

This paper uses the example of the Policy Compass to explore how FCMs can be applied to support the policy maker in evaluating the impact of a policy and provide guidance for future policy development. The paper is organized as follows. The fundamental concept of FCM techniques and their formal definitions are briefly introduced in Section 2. The possibility of FCM as a policy impact modelling tool is outlined in Section 3. In Section 4 we show how the policy maker can use FCM as a policy impact evaluation tool and finally, we conclude in Section 5 with future research.

2. Preliminaries – Short Introduction of FCM

FCMs are fuzzy signed graphs with feedback [8]. They consist of nodes representing concept C_i and the relationships w_{ij} between concepts C_i and C_j. An FCM models a dynamic complex system as a collection of concepts and cause–effect relationships between concepts [6]. A simple illustration of an FCM consisting of five node concepts is depicted in Figure 1 below.
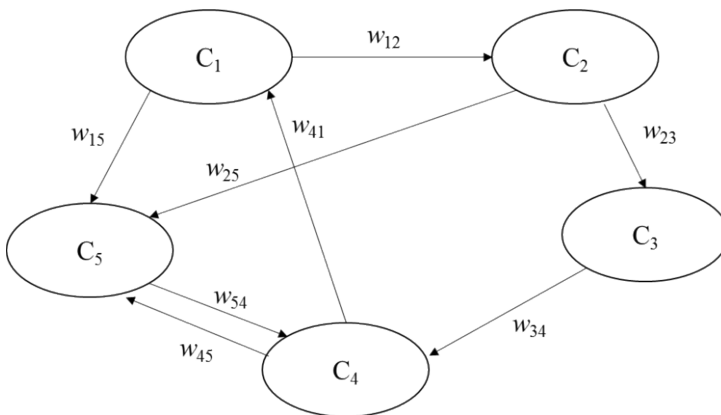
Figure 1. A Simple Fuzzy Cognitive Map

A weight w_{ij} describes the strength of causality between two concepts. Weight should take a value in the interval [-1, 1]. The sign of the weight indicates positive causality if $w_{ij} > 0$, which means that an increase in the value of concept C_i will cause an increase in the value of concept C_j. Similarly, a negative value of w_{ij} indicates negative causality. When no relationship exists between concepts, then $w_{ij} = 0$. The value of concept is usually fuzzyfied by mapping linguistic measure (i.e., very low, low, middle, high, and very high for 5 scale measure) to fuzzyfied value in the interval [0, 1]. According to the scale of the fuzzyfication scheme, every value of a fuzzyfied concept is a given fuzzy value. A fuzzyfication of linguistic measure enables the transformation

from qualitative measure to quantitative value so that we can use cognitive map as a multivariate time series prediction model.

The value of each concept at time *t* is calculated by applying the calculation rule of the equation below, which computes the influence of other concepts on the target concept:

$$x_i(t) = f(\sum_{j=1, j \neq i}^{n} x_j(t-1) w_{ji}) \tag{1}$$

where is the value of concept C_i at time *t*,   is the value of concept C_j at time t-1,   is the weight of the relationship between concept C_j and C_i, and *f* is the activation function. At each time step, the values of all concepts in FCM change and recalculate according to this equation. The value of concepts in FCM at time *t* can also be expressed as a matrix form. Assuming that vector **X**(*t*) is the *n* by 1 vector that gathers the value of *n* concepts, then the matrix **W** is an *n* by *n* matrix representing the weights between *n* concepts:

$$\mathbf{X}(t) = f(\mathbf{W}^{\mathrm{T}} \mathbf{X}(t-1)) \tag{2}$$

The concept of activation function is borrowed from artificial neural networks. It is a function that calculates the output of a concept based on its inputs, usually on the total sum. The output of activation function has its upper/under bound as +1 / −1. The most common type of activation function in FCM is the sigmoid function, which is a reciprocal of negative natural logarithms with few parameters. In addition to this function, many different type activation functions are applied in various contexts.

Based on the definition of equation and activation function, the state vector can be calculated, which contains the values of all concepts at time *t*. In the simulation of FCM, the calculation of the state vector is iterated until the steady state is reached, indicating that no changes occurred in the state vector after one iteration. Not every simulation result can reach an idle state. Fluctuation of every concept value can be found according to the FCM model.

3. Advantage of FCM as a policy impact modelling tool

FCM is widely used to analyse the impact of policy or strategy changes, including social science, political systems, and engineering systems. Different types of impact analysis models and tools have been proposed in the existing literature[2], each with their own pros and cons in terms of functionalities and analyst needs. In the case of FCMs, it is useful for modelling systems that cannot be explained entirely mathematically or that need to represent both qualitative and quantitative information or model both tangible and intangible issues [7]. Thus, FCMs are able to represent domain or process or problems that can be considered to be complex, vague and even incalculable [3]. The advantages of FCM can be summarized as the following:

- Easy to use and parameterise
- Be able to model casual relations that are not known
- Easily understandable/transparent to non-experts and lay people

---

[2] Literature review on policy impact modelling and analysis can be found in [5].

- FCM can be used as a rich body of knowledge by combining views of experts or stakeholder from different information sources banding them in structural/understandable form
- FCM is a dynamic system capable of capturing the dynamic aspect of system behaviour

The next section discusses how these advantages can be applied for evaluating the impact of policy with an example of interest rate policy.

4. Strategic Usage of FCM as a Policy Impact Evaluation tool – Use Case for Interest Rate Policy

4.1. Description of the use case problem

Assume that a policy maker in the government wants to know the future impact of change in interest rate to stimulate productive investment. However, in this context, the side effect of increasing the interest rate remains to be the problem. The present level of interest rate is "low", which can be fuzzifyied into a value of 0.4 in 5 scales fuzzyfication scheme (very low: 0.2, low: 0.4, middle: 0.6, high: 0.8, very high: 1.0)A policy maker estimates what will happen in the future by following different situations:

1. Situation 1: If the interest rate is **kept in the same level in the future**
2. Situation 2; If **the interest rate decreases to date**
3. Situation 3: Or if **the interest rate increases**

Considering these situations, the questions is how can FCM modelling and simulations help the policy maker to make better decision for interest rate?

First, he/she lists the factors related to the change in interest rate. Based on the listed factors, he/she can make the causal relationship diagram between the factors. Assuming that occupation and inflation are considered for the estimation of the relationship between interest rate and product investments, the policy maker can model the FCM with appropriate causality between them. Figure 2 illustrates the FCM model among the four related concepts and their causal relationships.
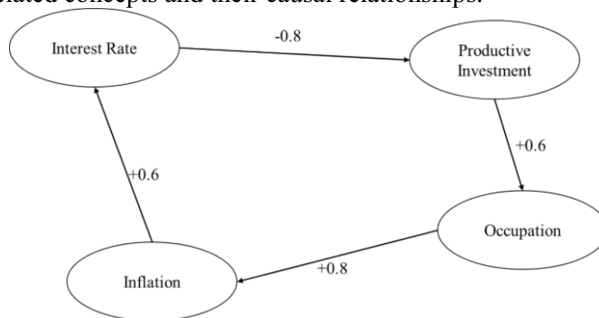
Figure 2. FCM for interest rate policy

The initial state of four concepts can be fuzzyfied with the 5-scale fuzzyfication scheme: ***Interest rate: 0.4 (low), Productive Investments: 0.2 (very low), OccupationT 0.8 (high), Inflation: 0.2 (very low).*** Finally, the possible future with this scenario can be analysed. Usually, the initial value of concepts and weights among concepts can be determined based on the discussion of panels and domain experts. In the context of

policy making and evaluation, policy makers and related experts can evaluate the current level of given values based on linguistic measure and fuzzyfication scheme. Policy Compass platform provides option for defining initial values and weights. Also, the weights that can explain the given FCM model with historical value can be automatically calculated using data mining techniques.

4.2. Simulation results[3]

Situation 1. If the interest rate is __*kept in the same level in the future*__

In this scenario, we can simulate the model with a given initial value. The sigmoid activator is used for all concepts. The simulation result can be obtained using the equation for calculating the next state vector (see the figure below).
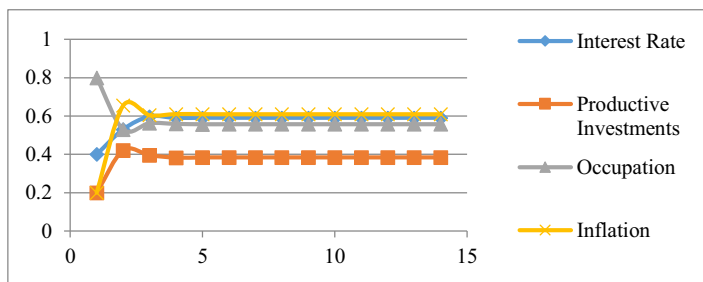
Figure 3. Simulation Result – Low interest rate, without fixed output effect

The model converges to an idle state after a number of iterations; each value of the concept is shown in the final state. In the final state, productive investment increases to 0.384, whereas occupation decreases to 0.557. Inflation also increases to 0.609. Even the future interest rate is estimated to increase to 0.59. However, this simulation has a **critical mistake**. Interest rate should be seen as a policy variable, which can be determined by policy rather than independent variable affected by other variables. Assuming that the interest rate is **"kept"**, this model should be in a totally different manner. In this case, we can enjoy the "fixed output" option for the concept value of "interest rate." If we consider the interest rate as a fixed output variable, it always returns fixed output regardless of the input value in a more realistic manner.
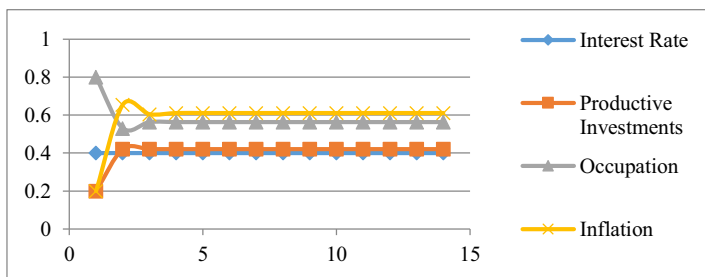
Figure 4. Simulation Result – Low interest rate, without fixed output effect

---

[3] We use the sigmoid function as an activator function.

The comparison table below shows that, if we do not consider the fixed output for interest rate, the simulation result underestimates the productive investment in the final state. This simulation result shows that, if the interest rate is to be maintained at its present level, productive investment will improve even if the inflation level increases.

Table 1. Comparison of the final state between with and without fixed output (interest rate: AS-IS)

	Interest Rate	Productive Investments	Occupation	Inflliation
w/o fixed output	0.5904428	**0.3840592**	0.5573553	0.6096602
with fixed output	0.4	**0.4206757**	0.5627685	0.6106903

Using the fixed output option for interest rate, we can simulate other scenarios in a more realistic manner.

Situation 2. If *the interest rate decreases to date*

The possible scenario for the stimulation of productive investments is to decrease the interest rate. This simulation allows the policy maker to estimate the potential impact to other important economic factors such as inflation. For this simulation, the interest rate is set at a very low level (0.2).
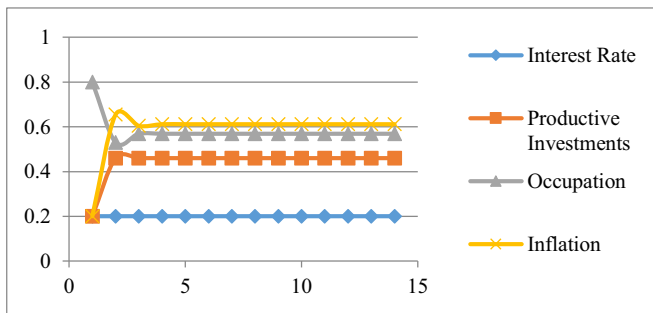
Figure 5. Simulation Result – Very Low interest rate, with fixed output effect

Figure 5 shows that inflation is getting worse (maximum 0.6547) in the first few iterations, and then stabilizes around 0.61 under very low interest rate condition. In terms of productive investments, the final value reaches 0.46, which improved compared with its current situation. To confirm the effect of lower interest rate on improvement in productive investment, we can assume the opposite, i.e., higher interest rate situation.

Situation 3. If *the interest rate increases*

In the third situation we assume that if the interest rate increases and is maintained at a higher level, how will this affect the simulation results? By comparing this result with the prior simulation (lower interest rate scenario), decision makers can understand the impact of altering the interest rates.
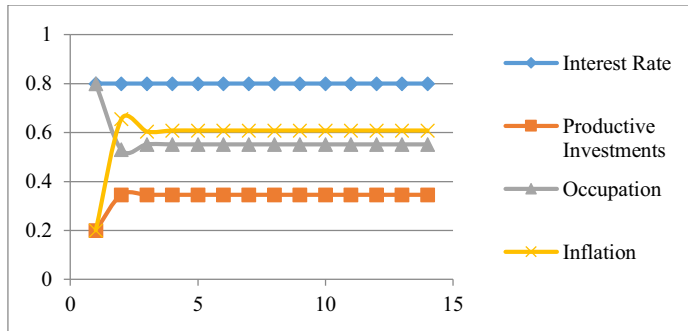
Figure 6. Simulation Result – Very high interest rate, with fixed output effect

Under higher interest rate assumption, the productive investments in the stable state are the lowest among the all scenarios assumed. Based on these three different simulation results, fruitful information about decision making can be derived.

4.3. Revisit the result – Why FCM modelling and simulation for impact analysis?

Policy impact evaluation assesses the changes that can be attributed to a particular intervention and change in policy. In contrast to outcome monitoring, which examines whether targets have been achieved, impact evaluation is structured to answer the question: how will outcomes change if the intervention will not be undertaken? This process involves counterfactual analysis; that is, "a comparison between what actually happened/will happen and what would have happened/will happen in the absence of intervention [2]. Counterfactual analysis enables evaluators to attribute the cause and effect between interventions and outcomes. The 'counterfactual' measures what would have happened to the beneficiaries in the absence of the intervention, and impact is estimated by comparing counterfactual outcomes to those observed under the intervention. From a perspective of impact evaluation, FCM modelling and simulation in Policy Compass platform can play an important role for policy impact evaluation. The simulation results of this scenario need to be revisited in this perspective (see table 2).

Table 2. Simulation Summary

	Interest Rate	Productive Investments	Occupation	Inflation
Low interest rate (as-is)	0.4	0.4206757	0.5627685	0.6106903
Very low interest rate	0.2	**0.4600851**	0.5685778	**0.6117947**
Very high interest rate	0.8	0.3452465	0.5516026	0.6085645

The policy maker may want to know that increasing interest rate will be effective to stimulate the productive investment. Although lower interest rate can induce productive investment, it may also cause unexpected and possibly related outcomes such as severe inflation (this scenario describes the situation with lower level of inflation). The policy maker has three options: keep the current interest rate, decrease the interest rate, or increase the interest rate. By analysing the simulation result above, we can develop an impact table for each of the decisions. Based on the results above,

we can confirm that decreasing the interest rate is the most effective decision among the possible decisions. If the level of interest rate is maintained, future investments will approach to around 0.42. However, if the interest rate is decreased, we can expect the highest future productive investments among all decisions, even with marginal increase of inflation. The worst decision is increasing the interest rate because it leads to the lower level of inflation with poor expected investments in comparison to the decision to keep the interest rate as it is. With the confirmation on the marginal impact of interest rate on inflation, we can choose the most effective decision for interest rate, which can result in maximum investment.

In this decision making, the situation 1 and 3 will not be chosen for optimal decision, but they can play a role as counterfactuals that can confirm the impact of the chosen decision. By comparing the possible decisions and their outcomes, we can make a more realistic decision for the given policy development situation.

5. Conclusion

Considering the increased demands on open data analytics for policy making process, Policy Compass can play a critical role in evaluating past policy impacts and preparing the blueprint for future policy development. As described in the prior sections, FCMs enable the user to model the complex causal relationship between the concepts relevant to the policy very intuitively. Not only the policy maker can evaluate the impact of policy, but also consider the use of open public data to expect the future impact of policy more easily. This innovative approach can also be a key contribution to FCM research field where most studies depend on human interventions for FCM modelling.

The value of FCMs as a policy modelling tool lies in evaluating the past policy impact. FCMs can provide a more intuitive and easier way to make use of the Europe's increasing amount of public sector open and structured data resources. By doing so, the Policy Compass will offer easy-to-use tools for both the lay public, as well as professional policy makers to improve the quality and transparency of the policy analysis and monitoring phases of the policy life-cycle.

References

[1] Blair, T. Labour Party Manifesto.(1997) http://www.labour-party.org.uk/manifestos/1997/1997-labour-manifesto.shtml
[2] Gertler, P.J. Impact evaluation in practice, Washington, DC: World Bank, 2010
[3] Khan,M.S., M. A. Quaddus and A. Intrapairot, "Application of a Fuzzy Cognitive Map for Analysing Data Ware Diffusion. Applied Informatics: Artifical Intelligence and Applications; Advances in Computer Applications.," in IASTED International Symposia, Innsbruck, Austria, 2001.
[4] Mago, Vijay K., et al. "Analyzing the impact of social factors on homelessness: a Fuzzy Cognitive Map approach." BMC medical informatics and decision making 13.1 (2013): 94.
[5] Panagiotis Kokkinakos, Ourania Markaki, Sotiris Koussouris, Kostas Ergazakis, Kostas Koutras, and Sarantis Pantelias, "Survey on policy impacts evaluation methods and tools and eParticipation and Social Platforms," Policy Compass (EU research project) Deliverable D1.1., (2014)
[6] Stylios, C. D., and Peter P. Groumpos. "Mathematical formulation of fuzzy cognitive maps." *Proceedings of the 7th Mediterranean Conference on Control and Automation*. 1999a.
[7] Stylios, C.D., and Peter P. Groumpos. "A Soft Computing Approach for Modelling the Supervisor of Manufacturing System," *Journal of Intelligent and Robotic Systems*, 26 (1999b) 389-403.
[8] Stylios, C. D., and Peter P. Groumpos. "Fuzzy cognitive maps in modeling supervisory control systems." *Journal of Intelligent and Fuzzy Systems* 8.1 (2000) 83-98.

This page intentionally left blank

General e-Government

This page intentionally left blank

Digitalize Building Permits Procedure: Canton of Valais, Switzerland

Georgia KOURAKOU[a,1] and Olivier GLASSEY[a]
[a] *Swiss Graduate School of Public Administration (IDHEAP),*
University of Lausanne, Quartier UNIL-Mouline, CH-1015 Lausanne
{georgia.kourakou, olivier.glassey.1}@unil.ch

Abstract. A growing interest in e-Government practices is currently expressed within public administrations. This paper describes the evolution of our study in the field of impacts of digitization on the management of administrative procedures. We present a full description of the methodology used, the assumptions and the analysis model on which our study is based. Due to its complexity and its interdepartmental characteristic, the building permit procedure is selected to be the representative administrative process of concern. As Swiss public administrations is the central subject of our research, a description of the relevant legal framework of the country is presented. Specifically, the case study of the Canton of Valais is described in detail. The experience gained from the semi-structured interviews conducted for this case study is exposed

Keywords. Digitization, building permits, e-Government, administrative procedure, optimization, time, complexity

1. Introduction

The recent trend of the last decades, in both personal and in organizational level, is the usage of new technologies. Private as well as public sector make reforming efforts towards modernization. In any case, internet and new technologies impose the transformation of public administrations. Public administrations try to take advantage of the electronic service delivery, as new technological solutions are promising. They invest in the provision of new opportunities offered by online and mobile applications [1].

Furthermore, optimization of administrative procedures was always one of the main targets of public administrations. Within the years, organizations applied several methods in order to achieve this goal. To mention that, when managing, administrations need to check if the adequate measures are applied [2]. Another important characteristic of public administration, usually blamed for, is the complicated, time-consuming nature of administrative procedures, in other words, the administrative burden and the delays in the provision of their services [3].

Trying to reach an interactive communication with their citizens-users, the public administrations rely on the optimal incorporation of e-Government tools and methods into their interior. E-Government implementation involves many elements of public administrations making the procedures even more complicated [4,5]. The complexity becomes worse by changes introduced during the implementation of e-services, due to

---

[1] Corresponding Author.

that usage of new electronic means [6]. Moreover, according to [7], internet can have a significant impact on governments, who will face really complex challenges while passing to a further interaction with their users.

The electronic delivery of services in the public sector could change them to more efficient and transparent ones [6]. Specifically, e-Government is considered to provide better quality of services [1] as they are responsive, efficient, and reliable with their users' needs, in terms of operation and responsiveness to the citizens [8]. In accordance with the second stage of the model of [9] concerning e-Government projects, administrations provide full access to electronic transactions and online databases for services, such as paying taxes, fees, etc.

As part of the improvement of the delivery of public services, the notion of "one-stop service", using information technologies, is an integrated service which can be offered by one provider, assuring interdepartmental coordination [10]. Also named as "one-stop portal", it is considered to be a cost-effective solution to the delivery of services [11], as users can carry out any level of government transaction contacting only one governmental point [12], contrary to their past need of contacting numerous different public units for the provision of one single service.

Finally, the effort to provide services electronically has already been described in several papers in the literature, in different ways by multidisciplinary researchers, which makes it difficult to create a knowledge basis [13]. The need of a complement study on the quality of government, including E-Government implementations, is also highlighted [14]. Moreover, according to [6] there is a lack of published research into the implementation of electronic services. The Organization for Economic Co-operation and Development (OECD) also mentions the need for international standards or indicators relevant to the impacts of e-Government on the facilitation of the decision-making process. Our study intends to fill the gap, regarding the impact of digitization of administrative procedures, trying to define a representative set of indicators.

2. Building Permits Procedure: Regulatory Provisions in Switzerland

Switzerland, also known as Swiss Confederation (official name of the Swiss federal state), has a federalist structure of three political levels: the Confederation, the Cantons and the Communes. Furthermore, Switzerland consists of 26 cantons, which have the status of federal states. Each Canton consists of Communes. Communes are the municipalities, being the basic level of political organization in Switzerland. All cantons apply almost the same policies. In addition to the tasks entrusted to them by their canton or by the Confederation, the communes have their own powers in several areas.

To begin with, at the federal level there is the Swiss Federal Law on Spatial Planning (LAT) [15]. According to article 1 (LAT) ''The Confederation, cantons and municipalities ensure the economical land use and the separation between building and non-building parts of the territory'' where at the same time ''they coordinate those activities that have an impact on the organization of the territory''. In the performance of their duties, they also support the management of the development of urbanization towards the inside of the built area, maintaining the quality of the appropriate habitat. Article 22 (LAT) stipulates that ''no construction or installation can be created or changed without authorization of the competent authority''. Equally, cantonal law assigns the exceptions inside the building zone (art.23, LAT). Canton laws also regulates

the procedures regarding the implementation, transformation or change of buildings and installations, defining the adequate deadlines as well. (art.25, LAT).

As a consequence, each canton arranges the building permit procedure independently. This is the reason why in the Swiss territory there are many different ways of approaching and implementing the issue of building permits. In some cases, there are cantons which centralize the procedure and don't provide flexibility to the inferior administrative levels. In other cases, municipalities have great eligibility for issuing building permits, thus reducing the workload of cantonal administration.

Besides, in the frame of optimization of cyber services, national sub-strategies with a thematic focus have been adopted within the framework of the eGovernment Strategy in Switzerland. This can be achieved by the implementation of individual prioritized projects so that they can be coordinated to a nationwide extent. In the Catalogue of Prioritized Projects of the Swiss Confederation [16] the undertaken projects can be identified in a coordinated manner within the framework of the eGovernment Strategy in Switzerland. As a result some specific public services can be provided primarily, due to the fact that they provide a worthy cost-benefit ratio for the target groups and the administration when provided electronically.

Regarding issuing building permits, a priority service project entitled A1.06 concerns the ''Application for construction permit'', led by the Swiss Conference of Directors of Public Works, Planning, and Environmental Protection (BPUK) [16].

In accordance to A1.06 project, the applicant can submit application and the necessary documentation for a building permit to the competent authority electronically. The applicant can track, at any time, the approval process and will also receive the building permit electronically after the procedure is concluded. Compared to the old paper-based approval procedure, internal administrative procedures are simplified, which means that the effort is significantly reduced. Additionally, in administrative terms, even partial automation of the procedure guarantees an optimization, which increases quality gains and finally efficiency for both applicants and authorities. High rationalization potential is generated, though, major obstacles to fully electronically processing building permits are still to be overcome, such as for example the need for blueprints on paper.

3. Research Methodology & Analysis Model

The objective of our research is to follow, to study and analyze the impact of the implementation of e-Government practices. In particular, the implementation and the dual effect that it has on both the function of public administrations (internal effect) and the interaction with its users (external effect). In order to achieve credible conclusions we use a combination of qualitative as well as quantitative methodology tools.

In this purpose and after thorough investigation of the literature, we focus on the three most appropriate dimensions in order to evaluate the digitization of administrative services and to measure the impact of electronic provisions of public administrations. The dimensions under consideration are administrative tasks, time and roles.

For each of the previously mentioned dimensions we have created a relevant research question, which will enable us to examine and deepen our understanding of the consequences that the use of Information and Communication Technologies (ICT) has to public administrations. Furthermore, each research question is associated with a specific theoretical assumption.

The principal research question corresponding to the dimension of administrative tasks is ''How does digitization transform administrative procedures are''. Based on the assumption that ''Digitization in the public sector drastically alters administrative procedure'', we expect that digitization will lead to less complicated procedures. It is also estimated that, especially for interdepartmental procedures, a digital approach will ease task completion process.

Regarding the aspect of time, we decided to study ''what the impacts of digitization on the time required to implement and provide an administrative service''. Our assumption is that ''Digitization in the public sector will accelerate administrative procedures''. As time is always a crucial factor, we suppose that digitization can considerably eliminate the processing time as well as the service delivery time. Digitalizing the process facilitates the communication between users and administration, eventually accelerating the procedure.

In relation to the dimension of role we will explore ''How digitization modifies roles and skills of civil servants''. We presume that labour conditions, employment needs and staff skill requirements are highly affected by a potential automation and dematerialization of the process. Modifications could apply to multiple levels, ranging from the human resource management to the redefinition of professional entities in an e-government environment. Furthermore, through the big task elimination civil servants will be allowed to be involved in more productive function of public administration.

In this respect, we have formulated an analysis model which includes the previously mentioned Hypothesis. Based on the three dimensions of interest and after an extensive exploration into relevant literature, we identified several appropriate indicators which will allow us to measure the qualitative and quantitative impact on the management of administrative procedures.

For each dimension of concern we detected a list of pertinent indicators. These indicators, presented in Table 1, will allow us to track and quantify the impact of digitization within the framework of public administration.

Table 1. Analysis model

Dimensions	Indicators
Administrative Tasks	• Volume of services provided by electronic means • Extent of the use of services offered by the traditional means (paper) • Number of activities involving several departments to provide a service • Request of documents; Reduction of volume of document requests • Existence of aid or technical support of a service (call center, help desk, information desk, etc.)
Time	• Time required to request a service • Time required to provide a service
Roles	• Number of people involved in providing a service • Need to recruit new specialized employees • Need for trained staff to adapt to new tools of e-services

The main field of this analysis concerns the Swiss public administrations. Specifically it involves research and comparative case study between three different Swiss Cantons, more specifically the Cantons of Geneva, of Neuchâtel and of Valais.

In an effort to facilitate our attempts, we selected a particular, process, the building permit procedure. It involves several complicated tasks and numerous stakeholders and multilevel communication. Currently, it is regarded as a time consuming procedure,

known for its complexity but at the same time convenient as several Swiss Cantons are introducing relative online procedures.

Up until now, we focus on field study, analyzing pertinent documentation and gathering further information. Following our methodology, we have distributed an online questionnaire survey [17] and in the context of our case studies he have also conducted a number of semi-structured interviews, presented in detail in following section.

4. Case Study: The Canton of Valais

On the way to our research path, the canton of Valais was chosen to be the preliminary field to explore. A canton with a special geographical and geological position, plays an important role as a traditional touristic point, having a significant number of building permit applications for both owner occupancy as well as for commercial space and hospitality industry or businesses establishments.

As mentioned above in the regulatory provisions' part, each canton, having an independent political structure, is responsible for developing its own law regulations as far as it concerns the building permit process. As required by the Swiss Federal Law on Spatial Planning, the canton of Valais sets this procedure by the Building Act of 8 February 1996. The competent authorities providing building permits in matters of the article 2 of this latter are ''the municipal council for projects located within the building zones, hamlets areas and maintaining areas of rural housing'' and ''the cantonal commission for the construction of buildings and facilities located outside building zones.''

Under the circumstances of the cantonal law of Valais mentioned above, the municipalities (communes) have a very great margin of manoeuvre (flexibility), as they are responsible for all the residential, central, industrial areas and the buildings of public facilities as well. On the other hand, according the cantonal law on constructions the Cantonal Committee for Constructions (C.C.C.) is also responsible for the building projects to which the municipality is an applicant or party.

The typical current situation, demands that the applicants send their building permit application folder to the municipality in paper form. After checking the folder, the municipality uses its own file management software for constructions and then sends it to the Canton in a paper form folder. Subsequently Canton stores over again the same information in its own database. Currently there is no citizen-oriented portal for applying building permits.

Equally important, so that its work can be facilitated, the C.C.C. has under its supervision a service which contribute to the proper function of the building permit procedure. This is the Cantonal Secretariat of Constructions (S.C.C.) which, at the same time, discharges the duties of the Police of Constructions (Po.Co.).

The S.C.C., for a pre-considering building permit application transmitted by the municipalities, requires the opinion-notice (préavis) and the decisions of the competent cantonal authorities. Then, within 30 days from the receipt of a complete application, it communicates to the municipalities the views of the authorities consulted. Furthermore, for a building permit competent to the cantonal level, S.C.C. prepares the files submitted and provides the secretariat for the C.C.C. Finally, as far as it concerns projects within the jurisdiction of the Confederation, S.C.C. collects the positions of municipal and cantonal authorities and sends them to the competent authority.

In this context, the links with the municipalities are particularly strong. Building permit applications are actually received and processed by the municipality which, if necessary, forwards them to the canton. The communication between partners is by mail (with a lot of paper) and the registration of the folder is performed both in the communal and cantonal level. As the whole procedure is carried out by the traditional paper-based channel[1], there exist several problems, delays, waste of time and for that reason we consider that there is still room for improvement.

According to administrative documentation, an analysis of the communal and cantonal building authorization process revealed the need of developing a unique process involving all the stakeholders (applicants, municipalities, cantons). The need for an electronic platform for building permits was studied as well. Additionally, the development of a standardized process, especially the organizational and the technical prerequisites, the stages, the deadlines and the necessary resources, were underlined as essentials.

In the frame of our study, the canton of Valais was the first district to inquire into. Having already taken into account the respective legislation, we collected several administrative documents relating to the building permit process provided by the relevant administrative authorities.

As already mentioned, the procedure of building permits in Valais's canton is mainly based on the traditional paper channel. In fact, our field research indicates that the main parts of this administrative procedure are far from being completely dematerialized. Nevertheless, there exists an internal task management system (SAP) which also involves some data relevant to the building permit procedure.

Additionally, applicant can only get online information relative to their application for the issue of a building permit. This can be achieved through a website where the applicant, having the identification of the application, can check the state of progress of the investigation of his folder, but they lack the right to get information about the positive or negative opinion of the different services.

Relying on these laws, on the administrative documentation and on our research model, we proceeded to the next stage of our study, which is nothing more than meeting with the main participants in the process. The evolution, the details and the initial results of the interviews will be presented in the following section.

5. Interviews

As mentioned above, the crucial component of this study is a series of semi-structured interviews with the main key actors and stakeholders. Beyond the analysis of documentation provided by the administrations and the field observations, we deemed appropriate to conduct one hour-long interviews. For each canton it is scheduled that roughly ten people be interviewed, participating in the procedure of issuing building permits.

On the basis of an open-ended questionnaire, as used for the semi-structured interviews, we are flexible to modify the content according to interviewee's profile. The interview takes the form of a discussion which is likely to be amended depending on the information provided through the answers [18].

Moreover, in order to select the suitable individuals among those who take part in this procedure, we based on our analysis model, taking into consideration the size, the type of the administration, as well as the people's professions. Under those circumstances,

we try to guarantee the weighty representation of participants in the building permit procedure.

Heretofore, as far as it concerns the canton of Valais, a number of people were selected within the main administrative departments participating in the procedure of concern, such as the Cantonal Secretariat of Constructions (SeCC) and Construction Inspection Services, Department of transport and environment (DTEE). To mention indicatively, some of the professional qualities of the interviewees are head of the department, computer coordinator and administrative collaborator. Equally important were some interviews with parties out of the public sector who are also involved in the process of issuing building permits.

The preliminary results show that the building permit procedure requires the cooperation of numerous administrative departments. For instance, for a small building permit project, the opinion-notice (préavis) of at least five different services is necessary, whereas for a large building permit project it could be necessary to obtain the opinion-notice (préavis) of up to twenty different services. This reflects the complexity as a result of the transversality of the procedure (Figure 1).

Furthermore, this characteristic of complexity reinforces the results of a previous online elementary questionnaire survey conducted in 2012 for the aim of the same research. More specifically, this latter questionnaire survey indicated that the more complex a service is, the less electronic means are used. This is also proven by the fact that in the moment issuing building permits in the canton of Valais is still being carried out on a paper-based procedure.

Figure 1

Such type of collaboration not only increases the degree of difficulty of the procedure but it also extends the length and the duration of the procedure. As a consequence of the necessity of the interdepartmental cooperation and of the complexity of the procedure's nature (ex. environmental issues) [19], the authorities find it difficult to meet the delivery deadlines.

Another interesting point revealed from the interviews conducted in the Canton of Valais, is that the enforced laws could also be an inhibiting factor for such administrative procedures. The lack of deadlines which do not put pressure to the working parts; or the lack of not clearly predefined responsibilities, for instance, could result in long delays. The need to correct the existing application form and subsequently to simplify-optimize it was also underlined. This is over again a jurisdictive problem to be solved.

Finally, there is frequently a delay, caused by the applicants, due to incomplete applications, curbing the progress of the issuing of building permit folder. The different services of concern cannot provide their opinion-notice (préavis) if the application is not entirely complete, a time-consuming factor, which also lies in the applicant's responsibility; whether they are the homeowner-citizen or an authorized professional [18].

6. Future Work

This ongoing research is an initial effort to conceptualize the impact of digitization on the management of administrative procedures. The theoretical framework and the research model used for the purposes of this study are still to be tested in order to tackle the question of how digitization transforms administrative procedures, and more specifically in terms of time and roles.

Currently, a detailed report concerning the treatment of building permits in the canton of Valais is being prepared. As part of our future work will be the documentary analysis and the analysis of interview transcripts, using the qualitative data analysis software ATLAS.ti. In parallel, we will finalize the respective interviews in the canton of Geneva and Neuchâtel, which will also be transcript, coded and at the final stage analyzed by the same software.

The ultimate step of this research will be the combination of all collected data. The target is to be able to draw conclusions on the impact of digitization especially on the building permit procedure, and generally on all administrative procedures.

Endnotes

1. The term "Service Channel" defines the way in which a service is requested and delivered.
    We have distinguished two types of service channels: 'Traditional Service Channel,' which includes paper services, telephone and fax; Electronic Service Channel,' which includes services through e-mail, the Internet and social networks.

References

[1] M.Barth & D.Veit., Electronic service delivery in the public sector: Understanding the variance of citizens' resistance. *Proceedings of the Annual Hawaii International Conference on System Sciences*, (2011), 1-11.
[2] R.Heeks, Understanding and measuring egovernment: International benchmarking studies, *UNDESA workshop E-Participation and E-Government: Understanding the Present and Creating the Future,* Budapest, Hungary, (2006), 1–44.
[3] V. Homburg, Understanding E-government: Information systems in public administration, Routledge, New York, 2008.
[4] R.M.Peters, M. Janssen &T.M.V. Engers, Measuring e-government impact : Existing practices and shortcomings, *Proceedings of the Sixth International Conference on Electronic Commerce* (2004), 480-489.
[5] J.Rose,J.S.Persson, L.T.Heeager & Z.Irani, Managing e-government: Value positions and relationships. *Information Systems Journal* (2015).
[6] R.El-Haddadeh,V.Weerakkody, & S.Al-Shafi, The complexities of electronic services implementation and institutionalisation in the public sector. *Information and Management* **50(4)** (2013), 135–143.
[7] S.Marche & J.D.McNiven, E-government and e-governance:the future isn't what it used to be, *Canadian Journal of Administrative Sciences* **20(1)** (2003), 74-86.
[8] R.O'Leary, D.M.Van Slyke & K. Soonhee, *The future of public administration around the world,* Georgetown University Press, Washington, 2010.
[9] K. Layne & J. Lee, Developing fully functional E-government: A four stage model, *Government Information Quarterly*, vol. 18, no. 2, (Jun. 2001), 122–136.
[10]A. Tat-Kei Ho, Reinventing Local Governments and the E-Government Initiative, *Public Administration Review*, (2002), 434-444.
[11]T.Kohlborn, A.Korthaus,C.Peters & E.Fielt, A comparative study of governmental one-stop portals for public service delivery (Report), *International Journal of Intelligent Information Technologies* **9(3)** (2013), 1-19.
[12]C. G. Reddick, A two-stage model of e-government growth: Theories and empirical evidence for U.S. cities, *Government Information Quarterly*, vol. 21, no. 1, (Jan. 2004), 51–64.

[13] I.Lindgren & G.Jansson, Electronic services in the public sector: A conceptual framework. *Government Information Quarterly* **30(2)** (2013), 163–172.

[14] E:Wihlborg & M.S.Gustafsson, In Search of Quality of eGovernment, A Conceptual Discussion on Impartiality in eGovernment. *Proceedings of the 4th Conference for E-Democracy and Open Government, Danube University Krems, Austria: Edition Donau-Universität Krems* (2014), 85-98.

[15] Loi fédérale sur l'aménagement du territoire, LAT 700

[16] Confédération Suisse, eGoevernment: Catalogue of Prioritised Projects, (20th February 2015).

[17] G.Kourakou, O.Glassey, F.Evequoz, The Impact of Digitization on the Management of Administrative Procedures: The Case of Building Permits*, Electronic Government and Electronic Participation, Joint Proceedings of Ongoing Research, Posters, Workshops and Projects of IFIP eGOV and ePart 2014* (2014), 197-205.

[18] A.Bryman, *Social Research Methods*, Oxford University Press, New York, 2012.

[19] P.Knoepfel, S.Boisseaux, Expertise sur l'efficience du service valaisan de l'environnement en matière de préavis sur des projets de construction, *IDHEAP*, (Jan. 2013)

Fourth Generation of User-centered Design – Developing for E- government and Cross-sector Collaborations

Sofie Pilemalm [a], Ida Lindgren [b] and Elina Ramsell [c]

Department of Management and Engineering, Linköping university, Sweden
[a] *sofie.pilemalm@liu.se**;* [b] *ida.lindgren@liu.se;* [c] *elina.ramsell@liu.se*

Abstract. This study explores the need for user-centered design (UCD) approaches to adapt to recent societal trends of cross-sector collaborations, civil citizen involvement and e-government initiatives affecting the public sector. This is achieved by studying three cases involving such trends, taking place in the Swedish Emergency Response System. Using results from the cases, information systems development related challenges are identified and related to adaption needs for current UCD approaches. Suggestions of such adaption are provided and a number of inherent challenges for the fourth generation of UCD are discussed, including challenges concerning (a) balancing ideological versus practical needs; (b) resources; (c) lack of know-how; and (d) design techniques and tool challenges.

Keywords. e-government, cross-sector collaboration, user-centered design, emergency response systems

1. Introduction

Word-wide financial crises with cutbacks of capital and resources in the public sector has led to new societal strategies including a focus on multi-functionality where people employed can attain several roles and engage in *cross-sector collaborations* aimed at creating redundancy in public systems [1]. In terms of information systems (IS), several of these strategies have been made possible by the availability of the Internet and mobile technologies, also leading to an increased focus on *e-government* and the responsibility of the civil citizens to perform certain functions previously carried out by the public sector authorities [2]. The trends can be seen in different areas such as climate change, environmental protection, poverty, natural disasters, and e-health [1], [3]. A recent example is *emergency response systems (ERSs)*, e.g. the organizations, personnel, methods, equipment and ISs involved in carrying out rescue operations. Here, we see an increased focus on other societal sectors, civil citizen volunteers and non-profit organizations (NPOs), to collaborate with professional response organizations. This is achieved by asking them to act as first responders at the emergency site, or by providing assistance or information using social media channels.

Relating the trends to IS development and different *user-centered design (UCD) approaches*, early approaches of user participation had clear ideological connotations [4]. Since then, different approaches and especially Participatory Design (PD) has reflected societal and technological development in moving from (1)

designing technology for homogenous shop-floor worker groups; to (2) design for collaborative office teamwork; (3) to designing for inter-organizational collaboration in increasingly heterogeneous and complex work contexts [5]. This may be referred to as the *first, second and third generation of UCD approaches*. Meanwhile, with regard to contemporary societal trends, several researchers conclude that user involvement in e-government service design and research as to date has been mostly non-existent [6], [7]. Others report on attempts to include the users in e-government projects but describe related challenges and absence of success [8]. Recent research has shown that civil citizens/end-users are in general willing to participate but their ability to do so is limited, often due to limited resources set aside at lower organizational levels [9]. While different user participation schools such as PD, UCD and user innovation (UI) have been demonstrated to fit into the development of e-government and public e-services [9], previously identified UCD challenges in certain aspects have become more substantial. For instance, in e- government, many different organizations and sometimes different societal sectors are to collaborate with the civil citizens. In addition, the primary user group is often difficult to distinguish. It seems that the fourth generation of UCD needs to adapt to new societal trends and to again embrace the civil citizen focus and ideological/political connotations. One way to do this is to transform knowledge from IS development in these areas into practically feasible UCD approaches and methods.

1.1. Study aim and objectives

This study explores the need for UCD approaches to adapt to recent societal changes of cross-sector collaborations and e-government affecting the public sector. In specific, the research objectives include:

- to identify UCD related challenges, departing from three case studies reflecting the trends presented above;
- to identify UCD needs for adaption; and
- to provide suggestions for how the fourth generation of UCD can address these needs.

The study uses data from case studies undertaken in the Swedish ERS, and relating to IS supported collaborations of professional response organizations with other societal resources and civil volunteers. Since the collaborations reflect general trends in the public sector, the results should be transferable to a wider UCD, e-government and public sector audience.

2. Background

This section includes brief descriptions of cross-sector collaboration and e- government trends and an overview of the emergence of different UCD approaches. Finally, the study context in terms of ERSs is described.

2.1. The public sector and cross-sector collaboration trends

In the public sector, shortage and cut-backs in resources due to global financial crises have forced public organizations towards to seek help from other organizations and societal sectors to create redundancy in societal systems. Cross-sector collaboration has been defined as partnerships involving government, business, NPOs, communities, and/or the public sphere as a whole [1]. It often includes elements of multi-functionality where people switch between roles and assignments in the organizations involved. Several studies claim benefits for cross-sector collaboration such as enhancing efficiency and improving organizational accountability [e.g.10]. Others point out related challenges such as distrust, managerial complexity, cultural conflicts, power imbalance, risk of dependence, and lack of incentive for collaboration [e.g.11]. While ISs are fundamental components in developing cross-sector collaboration most related studies discuss pure technological aspects and interoperability issues [e.g.12]. As to UCD, while different UCD studies have included specific focus on inter-organizational design, they do not yet seem to have embraced cross-sector collaborations.

2.2. E-government

Closely related is the increased spread of *e-government*: the use of the ICT in public administration. Combined with organizational change and new skills it is supposed to improve citizens' opportunities to interact with government authorities, to increase authorities' efficiency, and to increase democracy through greater governmental transparency [13]. E-government initiatives often entail that civil citizens are supposed to take on certain tasks previously carried out by the authorities themselves, for example through the use of e-services. We also see more complex e-government initiatives including cross-sector collaboration. ICT seems to be a driver of this development, accelerating interoperability and interconnectedness in the public sector, enabling collaboration and information sharing across organizational boundaries, and allowing for interplay between public and private organizations, NPOs and civil citizens [14]. Studies have shown, however, that the full potential and use of the e-government is seldom realized [e.g.15]. One challenge is the citizens themselves since they constitute a heterogeneous stakeholder group that is difficult to reach [9]. It has been claimed that, thus far, user participation and systems development research appears to have had little influence on the e-government field [7].

2.3. Different generations of UCD approaches

UCD is no clear or specific design methodology. On the contrary, several approaches and generations have emerged and adapted to prevalent societal trends and available technologies. Early schools such as cooperative design and early PD had clear political/ideological connotations and stemmed from actively involving civil citizens and shop-floor workers in public planning and technology development. Focus was on industry, production, and small, homogenous work groups [4]. The second generation of UCD approaches including later versions of PD came to embrace office teamwork, and intra-organizational collaborations. Focus in part shifted from ideology to practicability and the notion that active user participation actually lead to better and more accepted systems. Third generation of UCD included successive movement

towards inter-organizational collaboration in increasingly heterogeneous user groups designing large, complex ISs [5]. It also included an increased focus on usability and interface design. Today, UCD refer to various approaches such as PD, Human-Centered Design and UI. They all use slightly different perspectives, focus, methods, terminology and degrees of involving the end-user; however, their shared focus is the end-user. While the IS community in general agrees on the benefits of UCD, there are also critics that have pointed its limitations and inherent difficulties, e.g. [5]:

- *Difficulties in involving the shop floor end-users* who often have less time to divide between IS development and ordinary work tasks than the organizational managerial level leading to uneven representation in design groups.
- *Conflicts* where end-users and managerial level have different views on the emerging system and where power relations affect the prioritization of system requirements.
- *Lack of formalization* where UCD is deemed loose and lacking in a coherent design methodology covering the later implementation phases, leading to a situation where an end-product is not reached.

The above contributes to the criticism that UCD is too resource- and time-consuming. Critics have claimed that in particular the cooperative and PD approaches are academic constructions and are not feasible in industrial contexts characterized by short-term demands for cost-efficiency [16]. Some difficulties also seem to have grown with the growth of each generation of UCD having to deal with an increasingly complex societal and technological development such as [5]:

- Increasingly large and heterogeneous user/stakeholder groups leading to difficulties in representing all users in the design group.
- Job rotations reflecting contemporary labor market trends rendering it more difficult to achieve stability in the design group.
- Lack of easy-to-grasp development tools increasing with more complex ISs requiring technical support tools e.g. for modelling and handling requirements.

2.4. Fourth generation of UCD: emerging needs of renewal

Contemporary trends of cross-sector collaboration and e-government have to deal with ever increasing complexity relating to inter-organizational challenges, and to double and at times conflicting aims of increasing internal and external efficiency [17]. In e-government, there a numerous studies that point out the need for UCD and also a few that try to apply it in this context. For instance, [18] points at related challenges and propose a corresponding future design agenda including a design plan, understanding user technology availability expertise and preference, needs assessment, engage users, evaluate e- government services and form community- based partnerships. Other studies point at inherent difficulties including [6, 17, 18]:

- Contradictory visions of the envisioned tasks by users and authorities and their differing commitments to related law and governments.

- Unclear user target groups causing inability to fulfil usability and relevance goals.
- Lack of time and motivation being further complicated by the fact that civil citizens are without organizational belonging or residing at a work place.
- Lack of methodologies and "know how".

Thus, the challenges for previous generations of UCD have become even more complex in the e-government context, implying that the fourth generation of UCD approaches must adapt accordingly.

2.5. Study context: emergency response systems

Emergency response systems (ERSs) can be defined as systems of organizations, technology, procedures, and rules aimed to save lives and minimize material and infrastructure damages in the immediate operation following an emergency. In Sweden and internationally, there are current efforts to compensate for the decreasing number of professional response resources available. One way is to use actors from other societal sectors or civil citizen volunteers in emergencies and let them collaborate with the professional response organizations. The expected outcome is more effective response and maintained public services, thereby saving more lives, reducing the number of causalities, mitigating the damage to the infrastructure and saving money. This study is based on data from three projects in the Swedish ERS, relating to three different forms of collaboration:

- Permanent co-location of professional response organizations and supportive actors
- Cooperative use of equipment and existing on-call societal resources
- Involving civil citizen volunteers acting as first responders

The Swedish ERS is part of the Swedish public services and the above initiatives reflect inter-organizational and cross-sector collaboration and emerging forms of e-government.

3. Methods

This section presents the research approach in the form of a *triple nested case study* and the empirical data collection methods applied.

3.1. Case studies

Case studies are the study of real world phenomena e.g. an individual, an organization, a setting, a process, a situation, an incident, or an IS. *Exploratory case* studies investigate and characterize phenomena in relatively new domains. Case study research can include multiple cases. Case studies can also be embedded, meaning that each case study contains more than one sub-unit of analysis [19]. This is a triple embedded case study where each of the new forms of collaboration initiatives in

emergency response has been studied from an exploratory perspective. The cases are *Co-location in Safety House, Co-use in Nyköping* and *the Enhanced Neighbors project*. For the purpose of this study, the sub-units *ISs for cross-sector collaboration, e-government and civil citizens* and *its implications for UCD* have been chosen for analysis.

3.2. Interviews and focus groups

Semi-structured interviews involve pre-formulated questions or themes that keep focus on the subject without strict adherence to them. Focus groups are group interviews where people are asked about their perceptions, opinions and experiences in a similar way but from a group perspective [20]. In the studied cases, semi- structured interviews were performed with the Safety House project manager, and with three representatives from the police, the fire and rescue services and the Swedish Defense services. In the Co-use project the project manager/fire chief, a social care worker and a caretaker/technician were interviewed. In the Enhanced Neighbors project interviews were performed with the project manager/fire chief, and with the SOS Alarm operator handling the alarm calls. In addition, three focus groups of different size were performed with volunteers in three villages. A telephone interview was also conducted (because of the geographical distance) with one volunteer from a fourth village. All interviewees were interviewed about their respective experience in the new collaboration forms regarding perceived strengths, problems and further needs (including focus on IS support). Interviews were both documented in memory notes and audiotaped for further transcription and analysis.

3.3. Future workshops

Future workshop is a design technique originally stemming from PD. It allows users to reflect upon their work situation and needs for improvements, in order to identify realistic and innovative solutions and IT support. They are usually divided into *the critique phase, the fantasy phase* and *the implementation phase*. The first phase focuses on the work situation, potential problems and n e e d s f o r improvement. Futuristic solutions to the identified needs are the focus in the second phase and are not restricted by technical or organizational constraints. In the third phase, the identified solutions are transformed into realistic, organizationally and technically feasible implementations [21]. A half-day future workshop was arranged in all three projects. In Safety house, actors from the police, the municipalities, the rescue services and the representative from the Swedish Defense participated. In the Co-use project, actors from the rescue services, the social division and the technical division participated. In the Enhanced Neighbors project, actors from the rescue services was represented by the project manager/fire chief, an officer working in the project and an information officer working with the homepage and social media. Eight volunteers from different villages also participated. The focus of the workshops was on the improvements regarding the respective collaboration forms in related response operations, including (IT supported) solutions to current problems. They were documented using memory notes and post it notes.

3.4. Data sample and analysis

The respondents constitute a strategic sample chosen together with the project leaders in each project. In Safety House and the Co-Use project all interview respondents also participated in the workshops. In Enhanced Neighbors two interview respondents from the rescue services also participated in the workshop as did some volunteers from one of the focus groups (Table 1). All cases were analyzed using the same data analysis approach. Data was sorted into meaningful categories and finally into central themes for each study, and reported in detail elsewhere [e.g. 22]. Subsequently cross-case comparisons were made to filter the themes of relevance for the study objectives. Different researchers have been involved in the cases but two researchers have been involved in all three cases, in the data analysis, and the comparison across cases.

Table 1. Summary of data collected by interviews, focus groups and future workshops

Case	Interviews (1-1,5 hours each)	Focus groups (app. 2 hours each)	Future Workshop (half day)	Roles represented in case
Safety House	4 respondents	-	8 respondents	*Prof. response organizations:* the police, the rescue services, the Swedish Defense
Co-use project	3 respondents	-	10 respondents	*Prof. response organizations:* the rescue services *Complementary societal resources:* the municipality social division, the technical division
Enhanced Neighbors	5 respondents + 1 telephone interview	6 + 4 +3 respondents	11 respondents	*Prof. response organizations:* the rescue services, SOS Alarm *Civil citizens:* volunteers

4. Results

The results from each case related to the overall themes are first presented. This is followed by analysis of the implications for UCD and concrete suggestions on how UCD can correspondingly adapt in the different design phases.

Case1: Co-location in Safety House Östersund. *Co-location* is employed in the city of Östersund, in a sparsely populated province in north Sweden and with about 45,000 inhabitants. In the 'Safety house' building the professional response organizations, authorities, the Swedish Defense and the Swedish Church (NPO) are jointly located and work together. The goal of the collaboration is to achieve a citizen-centered service and to improve alarm management in order to shorten the discharge time of response teams. Co-location also aims to help actors quickly gain improved

shared situation awareness, and to create a platform for shared information management and information dissemination to the public and media. Safety House can be seen as an example of trying to further develop existing joint assignments and inter-organizational collaborations among professional response organizations, public authorities and, to some extent, NPOs. As to e-government, the project is aiming towards improving dissemination of information to civil citizens in the county.

Identified challenges: Given the nature of the co-location principles for inter-organizational collaboration should be fairly well established. However, the interviews and workshop respondents reported that when gathering around an emergency, it is often obvious that the inter-organizational hierarchy - who are the decision makers and who holds which responsibilities in which tasks - is not clear enough. As to ISs, the joint map systems and systems for handling incoming errands are under development. Here, the intensified inter-organizational collaboration has resulted in increasing challenges of information confidentiality and was mentioned both by the police and rescue services as a major problem inhibiting the sharing of information. For instance, the police have a higher level of confidentiality than the rescue services. The exactness of this level is not clear to the personnel and the police representative claims that more information (e.g. about work methods, video and photos of incidents.) can be exchanged without breaking laws if the personnel are trained to express or use the material in a correct way. A further issue is what organization actually owns or will own the ISs developed for supporting the co-location collaborative processes. The rescue services respondents further pointed out that actively involving NPOs and more public authorities is a must if a citizen-centered service is to be implemented. At the same time, they deemed this as problematic as no routines or methods regarding the integration of these actors are available.

Case 2: Co-use in Nyköping. *Co-use* is taking place in Nyköping municipality, located in the middle of Sweden with a population of about 52,000. It involves the municipal rescue services, the social care unit, and caretakers employed by the municipality. The different actors share certain equipment such as vehicles and the social care and the rescue services are co-located in the fire station and share their alarm service/system. The social care unit has night patrols act on certain alarms previously handled by the rescue services. The co-use project is thus an example of cross-sector collaboration.

Identified challenges: The lack of structures concerning tasks, roles and responsibilities discussed above is even more prevalent in the Co-use project since different occupational groups and social sectors that have not worked together before move between organizational belongings in solving joint tasks. Some tasks have already been implemented while others are yet to be developed. Reported problems in the interviews and workshops include those related to lack of knowledge and understanding of each other's roles, and ethical and practical conflicts over what assignments to prioritize. As a respondent put it:

> *Should I, as a nurse, leave a lingering patient at the elderly care center to go with the rescue services on a cardiac arrest alarm providing CPR until the ambulance arrives, knowing that the chance of survival is only slight?. Who takes care of my patient?*

One respondent from the rescue services describes this as conflicts between "concern thinking" and "concept thinking" and relates it to a lack of consensus view between politicians and operative organizational levels. They are perceived to be a result of lack of clear assignments and routines, and insufficient knowledge as to who decides who does what and has what responsibility. As to ISs, joint systems include shared operational pictures and map systems. A system for sorting and delegating incoming errands and alarms to different actors was requested. A major obstacle identified was the inhibiting systems development and once again there were different views on information confidentiality. The division for social care viewed secrecy as a central and recurrent problem when collaborating with the rescue services, whereas the technical division expressed no such concerns. Legal issues related to employee rights and obligations were further deemed as a problem that needs to be addressed urgently. Issues that had not been fully solved at the start of the project and which remained unclear to the participants included employment agreements (e.g. who is the employer of a social worker when he or she takes on a task from, or together with, the rescue services?) and which employer has the responsibility for the joint work environment. Further, it was deemed unclear as to which actors in the municipality had the right to procure requirements on the new collaboration forms and related equipment and ISs. E-government initiatives have only recently emerged in the Co-use project. The respondents however pointed out that information and feedback to and from the civil citizens are central and should be provided using digital channels and e-services.

Case 3: Enhanced Neighbors project. Around the city of Sundsvall, with a population of 50,000 situated in a sparsely populated area in north of Sweden, 10 remote villages are participating in the project *Enhanced Neighbors*. These villages are situated at a distance from Sundsvall that makes it impossible for the rescue services to reach the village in less than 25 minutes. As part of the project, civil citizen volunteers in these villages receive basic training, e.g. in cardiopulmonary resuscitation (CPR), and are alerted by the SOS alarm center using SMS technology when an emergency occurs nearby. The volunteers usually reach the emergency site before the rescue services and can provide first aid and other supportive actions. They can also send relevant information to the rescue services vehicles. Participation is voluntary and the volunteers start their tasks while waiting for the rescue services but never replace them. This project is an example of how new technologies enable cooperation in new ways, involving civil citizens sharing tasks, technology, e-services and ISs with public organizations.

　　Identified challenges: The volunteers have no organizational belonging affecting how the development work had to be performed. The half day Future workshop was performed in the evening, to enable participation. Principles about the alarms the volunteers are sent on have been decided by the rescue services (drowning, fires, traffic accidents and cardiac arrests) and training in recurrent tasks that are related to these accidents have been provided. Even though the volunteers themselves expressed that their role was relatively clear it became evident out of the interviews and the workshop that several central issues remain undefined. This includes who has responsibility if something happens to the volunteers at the emergency site, or if they cause damage to victims, material or property. When talking to the rescue services about what current laws say on this matter the rescue services project manager said:

　　　　　　　I am not really sure, I hope their home insurance covers it

When the researchers later talked to a counselor, he was of the opinion that this is not the case and that insurances need to be provided by the municipalities. Also, the volunteers are not economically compensated for any expenses they may have incurred in relation to a rescue operation. For instance, the village needs to invest in defibrillators and reflective vests. The respondents in the focus groups were first fine with this. But some months later, in the workshop, they requested some economic support to be able to continue their engagement. The current design of defibrillators entails that a component must be exchanged after each use; resulting in a cost of approx. 200 Euro per use. Further, accidents occur comparatively seldom in sparsely populated areas meaning that many of the defined tasks have not yet been performed by the volunteers, except when in training. The majority of the volunteers have not yet responded to a real alarm and the volunteers expressed a need for further training during the workshop. As far as technology and ISs is concerned the volunteers access information provided by the Sundsvall rescue services and can communicate with them through simplistic e-services at the project website. The personal data act currently in force stipulates that no names of those victims involved in an accident can be provided in the SMS alarms. The respondents viewed this as a limitation and that it would be easier to find the right location if they had more information than merely the address. On the other hand, when they can identify a person by reading the address, this may pose ethical dilemmas. In one of the focus groups it became clear that most of the respondents were more willing to go on an alarm if they knew the victim. Some of the identified challenges will need to be directly addressed in a dynamic resource allocation I T - system which is under development. It will be used by all alarm centers/operators in Sweden, and volunteers will successively be integrated as a resource that can be sent on an alarm.

4.1. Analysis: implications for UCD

The cases represent different aspects of inter-organizational and cross-sector collaboration, and are in different stages concerning involvement of civil citizens in the development of shared ISs and work processes. All cases involve public organizations' use of IT to engage civil citizens and other non-public organizations in work that was formerly performed by the public organizations alone. As such, these cases can be inspected through an e-government lens. In addition, these cases can be related to implications for systems development and UCD. In the following, these implications are analyzed with a focus on the system development general phases (*organizational analysis, needs and requirement analysis, design, implementation* and *evaluation*) and how UCD usually addresses and could address these.

Organizational analysis: In all cases the project representation of the collaboration/system primary and secondary end-user are unclear. In Safety House, only a subset of the organizations participating in the co-location was represented in the initial collaborative/development work. The Co-location project has a somewhat more homogenous group of primary end-users, but has to consider an indistinct group of secondary stakeholders in terms of "cross-sector collaboration/system owners" with financial influences. In Enhanced Neighbors, the primary user group includes the civil citizens; but perceived unclearness or absence of laws, regulations, and principles, require that a much larger stakeholder group is identified and

consulted. A retrospective experience is that starting with organizational-, task- and needs analysis in the identified user contexts, was insufficient. This means that future UCD approaches will need to consider increasing heterogeneity of the stakeholder group, and that frameworks should embrace a more solid and comprehensive *stakeholder analysis* targeted to the respective collaboration where primary users, secondary stakeholders from various sectors and ways to reach them in the subsequent development process are clearly identified.

Further, the aspect of unclear role(s), tasks and responsibilities of the involved stakeholders was prevalent in all the cases. While the structures in established work organizations often can be reproduced in IS functions, the new cross-sector collaboration/e-government forms by large lack such structures. The cases also displayed a lack of knowledge of existing laws and agreements regulating the new collaboration forms and sometimes even the absence of laws and agreements. Information confidentiality issues seemed to be a major challenge, in specific where different societal sectors have different confidentiality levels. The stakeholder analysis thus needs to embrace corresponding early *policy and legal analyses*. These need to be performed by external experts if the identified users/stakeholders do not have an adequate knowledge to provide such information. Existing principles should be identified, interpreted and perhaps challenged not to undermine the new collaboration opportunities and corresponding IS solutions; but at the same time, it is crucial to stay within those principles that are not open to misinterpretation.

A re-current system development approach is to work in development teams consisting e.g. of system architects, programmers and, in UCD, end-users/domain experts, and (sometimes) usability experts. From the above it follows that for the cross-collaboration/e-government projects competencies from social sciences such as policy analysis and law are clearly needed. Judicial matters in particular are not straightforward tasks that system developers can solve by reading laws and agreements, as these are often subject to interpretations. This implies that the fourth generation of UCD needs to embrace more distinct *interdisciplinary* focus than is currently the case. Organizational analysis should embrace early identification of what competences are needed in the interdisciplinary design team and populating it accordingly.

Needs and requirements analysis: UCD approaches are typically iterative, where users are asked to participate in design groups or teams with regular meetings, specifically in the needs and requirements elicitation phases. This does not seem as a feasible option in design projects where stakeholders come from different social sectors with sometimes substantially different possibilities to participate regularly (and often without pay) in design projects. Also, in UCD some kind of task analysis is often used to capture tasks and needs of the end-users. In the studied cases some of the collaborative tasks are yet undefined, or have not yet been carried out. The roles, responsibilities, and decision rights are new and unclear with the users taking on certain occupational roles and tasks simultaneously. The focus thus in part had to move from (a) *describing* the tasks, to (b) *jointly identifying, negotiating, and formulating* them with users from different societal sectors and with civil citizens. Tasks and needs analysis was merged and worked on simultaneously since no clear distinction could be made. In Enhanced Neighbors scenarios were used in the workshop instead of actually experienced situations.

For tentative UCD future frameworks, this implies that alternative sources of data collection need to complement design groups in order to capture the perspectives of relevant stakeholders. When possible, interviews and focus groups may be used. If end-users do not have an organizational belonging, alternative solutions can include telephone interviews, web surveys and on-line interaction through social media channels. Selection of respondents can be made together with involved organizations and authorities and it is probably a pre-requisite for the selected users to have a n already expressed interest to participate in the design project. If tasks do not yet exist, or are in the developmental state, solutions could be to work with scenarios and early lo-fi prototypes and to investigate attitudes towards engagement in possible tasks. As far as the *prioritization* of requirements goes, the reported unclearness of the ownership of the IS may lead to potential conflicts as to which actor has the formal decision right to put requirements on the ensuing system and partake in its procurement. Integrating the perspective of civil citizens (who are not likely to own, be responsible for, or have decision-rights regarding the IS) is central. Legal/policy analyses results need to be integrated in the requirements specification.

Design/prototyping: In above all the Enhanced Neighbors project, all development work has needed to be performed in the late afternoon/evenings since this is the only time it was possible to gather the civil citizens. To retain this approach in iterative design does not seem feasible in the long run. This phase is thus probably the most suited to remote user/civil citizen participation since prototypes are concrete tools to work with and provide feedback on. Prototyping therefore seem a suitable approach for evaluating emerging design solutions. Suggestions include low-fidelity prototypes, again complemented by scenarios. If system developers cannot perform design evaluations in real life, real- time on-line contacts with the users providing comments on the system functions can be used, as can retrospective interviews and web surveys.

Implementation and evaluation: Speaking about *implementation* in terms of system architecture and construction, the end-users are generally not substantially involved. With regard to training on the new system and developing complementary work routines and methods, much the same approach as in the design phase can be used. If activities and tasks are not yet fully experienced among system users, in real life training sessions/joint exercises is probably needed before using the technology for collaboration in a real situation. This has been the case in the Enhanced Neighbors project where using the alternatives of SMS technology and the Swedish standard RAKEL communication system for emergency response was tested as alternatives in a traffic accident exercise performed jointly with SOS Alarm and the municipal rescue services. The *evaluation*, feedback on the system in use can be gathered using various methods, not the least through the information platforms developed by the authorities/organizations themselves.

5. Discussion

Development of IS and the use of UCD as part of e-government initiatives is transforming as a result of recent societal trends. We see a need for UCD approaches to

expand to include the cross-sector societal level, where even more responsibility is put in the hands of the citizens. In IS development, related challenges have included the difficulty to identify the primary user groups, contradictions in visions and views on tasks by authorities and civil citizens, lack of time, motivation and know-how [e.g. 7, 8, 13]. The majority of these can be noted in the studied cases, and even in the cross sector-collaborations. Moreover, the target group of e-government initiatives has previously been large groups of civil citizens or even the broad masses performing tasks out of own interest or because they are obliged to (e.g. performing the tax declaration on-line or needing an e-id to access the information provided by your child's school). The Enhanced Neighbors project illustrates how new forms of e-government initiatives emerge and are directed towards certain groups of citizens that are to actively support the public authorities by actually carrying out certain tasks determined for them and for their co-citizens, and being integrated in the authorities own ISs, further complicating the issue of IS development.

5.1. Fourth generation of UCD: challenges

In Safety House and the Co-location project, focus was on improving communication with civil citizens and developing citizen-centered services. Still, in their own development efforts, the perspectives of the citizens are lacking, confirming previous research findings [6, 7]. In Enhanced Neighbors, the researchers entered the project after some time and explicitly involved the volunteers in subsequent development work. Doing this revealed many basic needs that had been overseen, e.g., telephone numbers and personal contacts with SOS alarm and the rescue services. Involving the end-users, and avoiding pre-defined technical solutions, seems as urgent as ever. The difficulties that have been reported concerning UCD in the e-government context are not new; but seem to be intensified. To add to complexity, UCD approaches in general and PD, with its extensive focus on active user participation specifically, have been criticized for lacking in structure and a clear methodology, and for not being widely applied in their entirety outside the academic environment [16]. The fourth generation of UCD, when applied in the cross-sector collaborative/e-government contexts, thus faces a number of challenges that need to be handled if ISs in these areas are to be developed successfully:

The back to the roots versus effectiveness challenge: The fourth generation of UCD seems to demand an ever increased focus on ideological aspects, interdisciplinary work and focus on the early system design phases. At the same time, the growing complexity of the collaborative work and ISs demand an increased focus on effectiveness, planning and formalization of the development process, with clear and specified methods from stakeholder analysis to implementation if ready-to-use ISs are to be produced. This situation can be compared to the 'rigor versus relevance' discussion in the IS field, where there is a tension between doing what is ideologically and methodologically 'right' and what is practically relevant.

The challenge of resources: The above challenge leads to another. While cross-sector collaboration and e-government strategies have been developed to overcome lack of financial resources in the public sector, designing ISs to support these strategies clearly demands investing a large amount of resources. This challenge is probably the most problematic for the fourth generation UCD approaches to handle; yet necessary

if they should motivate their vindication. Ideological arguments of user satisfaction and that long-term efficiency may be attained but will probably need to be sustained.

Lack of know-how: Know-how and experience from successful cross-sector collaboration and e-government projects are largely missing. Designing IS support for civil citizens through active participation is intricate and the task probably becomes even more challenging with emerging types of e-government specifically directed towards certain groups of civil citizens which cannot easily be replaced with other user voices from the broad masses. Besides, the emerging nature of the collaborative systems embraces undefined stakeholders, unclear policies, responsibilities and legal issues, requiring know-how stemming from actual experience.

Design techniques and tools challenges. To enable active and rewarding user participation, users must be provided with suitable design tools and techniques. E-government and the availability of social media channels provide on-line alternative interactions with the end-users, if they cannot be involved in design groups or devote their time in real life. However, parts of human interaction, in depth information, and group dynamics when different stakeholder groups interact, are lost when using remote technology and guidance. This becomes more challenging when moving from generic public services to the emerging form of e-government described in this study. The fourth generation of UCD thus has to balance face-to-face and on-line interaction in design work, again related to the effectiveness and lack of know how challenges described above.

6. Conclusions and future work

In a society where cross-section collaboration and e-government is rapidly expanding, it seems necessary to focus on organizing the desired activity, e.g. the provision of safety, and not on the organizations themselves, moving towards a society built on multi-functionality and role flexibility. This in its turn immediately raises questions of who has the authority, legitimacy and profession and who has not, as illustrated by previous research [11] as well as in this study. The corresponding challenges concerning UCD lie in a number of challenges, which need to be resolved if UCD is to be applied, used and is able to deliver benefits to IS development in a networked society. A limitation of this study is that it only embraces the initial collaborative efforts and some early IS design in the different projects, which is reflected in the focus on early design phases in the suggestions provided for the fourth generation of UCD. Future research should focus on further development of more coherent and specified methodologies and frameworks, covering all IS development phases and on applying and testing them in entire UCD projects directed towards e-government and cross-sector collaboration. Long-term economic evaluations and cost benefit analyses are also central to be used as motivations for user participation, in the new contexts.

References

[1] Goldsmith, S., Eggers, W. D.: Governing by Network: the New Shape of the Public Sector. Brookings Institution, Washington, DC (2004)

[2] Verdegem, P., Verleye, G.: User-centered E-Government in Practice: a Comprehensive Model for Measuring User Satisfaction. Government Information Quarterly. 26, 3, (2009)

[3] Agranoff, R.: Managing within Networks: Adding Value to Public Organizations. Georgetown University Press , Washington, DC (2007)

[4] Ehn, P.: Scandinavian Design: on Participation and Skill. In: Schuler, D., Namioka, A. (eds.) Participatory Design. Principles and Practices. Lawrence Earlbaum, Hillsdale, NJ, 41-77 (1993)

[5] Pilemalm, S., Timpka, T.: Third Generation Participatory Design in Health Informatics - Making User Participation Applicable to Large-scale Information System Projects. International Journal of Biomedical Informatics. 41, 327-339 (2008)

[6] Kotamraju, N. P., van der Geest, T. M.: The Tension between User-centred Design and E-government Services. Behaviour & Information Technology. 31, 3, 261-273 (2013)

[7] Karlsson, F., Holgersson, J., Söderström, E., Hedström, K.: Exploring User Participation Approaches in Public E-service Development. Government InformationQuarterly, 29, 2, 158–168 (2012)

[8] Melin U., Axelsson K.: Managing E-service Development – a Comparative E-government Case Study. Transforming Government: People, Process and Policy, 3, 3, 248-270 (2009)

[9] Holgersson J., Karlsson, F.: Public E-service Development: Understanding Citizens' Conditions for Participation. Government Information Quarterly. 31, 3, 396–410 (2014)

[10] Brinkerhoff, J. M.: Government–nonprofit Partnership: a Defining Framework. Public Administration and Development, 22, 1, 19-30 (2002)

[11] Gazley, B., Brudney, J. L.: The Purpose (and Perils) of Government Non-profit Partnership. Nonprofit and Volun-tary Sector Quarterly. 36, 3, 389-415 (2007)

[12] Otjacques, B., Hitzelberger, P., Feltz. F.: Interoperability of E-Government Information Systems: Issues of Identification and Data Sharing. Journal of Management Information Systems. 23, 4, 29-51 (2007)

[13] Axelsson, K., Melin, U., Lindgren, I.: Public E-services for Agency Efficiency and Citizen Benefit – Findings from a Stakeholder Centred Analysis. Government Information Quarterly. 30, 1, 10-23 (2013)

[14] Lu, Z., Zhang, N., Meng, Q. Understanding Cross-Sector Collaboration. In E-government Development: Theoretical Consideration Based on Extended Bryson's Framework. PACIS Proceedings. 99, 1007-1019 (2010)

[15] Lindgren, I,. Jansson, G.: Electronic Services in the Public Sector: A Conceptual Framework. Government Information Quarterly, 30, 2, 163-172 (2013)

[16] van den Besselar, P.: Democratizing Technological Change: Limits to Steering. In: Henderson Chatfield, R., Kuhn, S., Muller, M. (eds) Proceedings of the Participatory Design Conference. CPSR, Pao Alto, CA, 1-10. (1998)

[17] Jaeger, P.T., Bertot, J.C.: Designing, Implementing, and Evaluating User-centered and Citizen-centered E-Government. International Journal of Electronic Government Research. 6, 2, 1-17 (2010)

[18] Jansen, A.: Strategies for User Involvement in eGovernment Projects: what can be Learned from the Scandinavian IS tradition? NordiCHI'06 Workshop on User involvement and Representation in E-Government Projects. Oslo (2006).

[19] Yin, R. K.: Applications of Case Study Research. Sage Publications, Newbury Park, CA, (2012)

[20] Myers, M.: Qualitative Research in Business Management. SAGE Publications, (2013)

[21] Kensing, F., Halskov Madsen, K.: Generating Visions: Future Workshops and Metaphorical Design, In: Greenbaum, J., Kyng, M. (eds.) Design at Work. Lawrence Earlbaum, Hillsdale, NJ, 155-168 (1991)

[22] Mojir, K. Y., Pilemalm, S. : Emerging communities of collaboration: co-location in emergency response systems in Sweden. Proceedings of the 11th International Conference of Information Systems for Crisis Response and Management (ISCRAM), PE, USA, 546-555 (2014).

Formative evaluation and user engagement: A model to ensure value from e-government

Jesper B. BERGER[a,1]
aInformation Systems, University of Siegen, Germany

Abstract. Governments are investing in e-government to enhance public sector efficiency. It has been argued by e-government scholars that citizens' demand for e-government does not meet expectations; hence governments might not achieve the expected benefits. This study investigates formative evaluation as a method to ensure e-government benefits realization. The case is ex-post evaluation of value from digital communication in a citizen service center at a Danish municipality and was conducted as Action Research. Barriers to adoption were revealed, addressed and eliminated during the formative evaluation process. Clear expectations from top management, assessments and disclosure of workers' behavior and commitment from managers proved pivotal in the benefits realization process. Further research into the internal e-government adoption processes and the impact from external factors is needed in order to understand more profoundly the challenges for realizing benefits from e-government.

Keywords: e-government, evaluation, adoption, local government, value, benefits, action research

1. Introduction

The challenges of realizing value from IS are widely recognized. Ward et al. [1] conducted a survey of perceived satisfaction with IS projects among 102 private and public organizations. Only 45% claimed success with more than half or more of their IS projects. Analysis of U.S. local governments surveys concluded that 'few governments reported any changes that are attributable to e-Government, especially changes involving cost impacts' [2]. Goldfinch [3] argues, that one should be pessimistic about value from e-government. There is a need to address the organizational changes and desired value along with implementing IT systems. The Danish Government and the local governments agreed on an ambitious e-government strategy [4]. During the five year period, 70+ public e-services will be mandated; 80% of communication with citizens and companies will be also be mandatory digital.

E-government is described as multivariate and complex, covering a wide range of areas, actors and applications [5]. It is commonly accepted that e-government research needs to be multidisciplinary to meet this challenge. From a comprehensive e-

---

[1] Fakultät III, Wirtschaftsinformatik, Kohlbettstraße 15, 57072 Siegen, Germany; E-mail: jesper.berger@uni-siegen.de

government literature review, Heeks and Bailur [6] do not find much evidence of a multidisciplinary approach. They find a predominance of information systems' research influence, no inheritance of critical views and only scarce studies based on solid empirical work offering practical recommendations. Yıldız [7] states that e-government evaluation 'only focus on the measurement of the availability and development of web sites and on line services'. It misses the organizational and cultural change that is necessary for e-government to succeed.

This study will apply a multidisciplinary approach through empirical work and direct contact with data to address the organizational and cultural change. A formative evaluation model of e-government adoption in local government in an Action Research approach was applied. **Research question:** How can formative evaluation impact value from e-government? Can this evaluation model reveal generic factors that support or hamper e-government value realization?

2. Related work

E-government, understood as the services, delivered to citizens etc. through the internet, can be measured in many different ways (as IS implementation, against anticipated objectives or in terms of efficiency), considering different phases of the e-government initiative (implementation or operation) and with different objectives (e.g. comparison, supporting decisions or understanding phenomena) [8]. Even though this e-government evaluation ontology is convenient, it leaves a gap between anticipated and actual e-government as it only defines one phase after implementation; 'operation'. I will add the technology adoption processes to close this gap.

Public sector has an obligation to deliver accountability, transparency, equality and reliability in the services towards citizens and businesses [9]. Concluding, that there is a lack of consistencies in terms of value metrics and stakeholders included in evaluation research, the author presents a unified multidimensional framework that covers all value dimensions and stakeholders [9]. Luna-Reyes et al. [8] suggest an evaluation model composed of not only output but also technological characteristics, organizational form, institutional arrangements and contextual variables. Luna-Reyes et al. add value dimensions that extend the e-government perspective. Both models build on a positivist philosophy that perceives value dimensions as objective, value free and measurable. Other scholars are opposed to a unified model view. Carbo and Williams [10] note the diverse subject matter areas that e-government is applied to and state that there is no one model for local government evaluation.

A vast majority of e-government evaluation studies concern only stakeholders outside of the practitioners setting, primarily citizens [9]. Evaluation of citizens' adoption of e-government applying behavioral models such as TAM, TRA or TPB [11-13] provide understanding of the citizen adoption process. Content evaluation of public websites and surveys of managers' perception of e-government value and barriers [14, 15] constitute other models of evaluating e-Government. Jones et al. [16] claim, with support from many researchers, that most organizations 'have no ICT evaluation processes in place' Yildiz [17] addresses the oversimplification of e-government and recommends to evaluate 'the processes that shape the management of e-Government'. Luna-Reyes et al. [8] state that 'we still know little about the impacts and results associated with e-government'. According to Ndou [18], 'one of the reasons why many e-government initiatives fail is related to the poor understanding of the e-government

concept, processes and functions'. Carbo and Williams [10] underpin that without appropriate evaluation models, e-government may be costly and include political, operational and technology risks. These studies provide useful knowledge in regards to comparison of e-government output but are of limited use for understanding e-government adoption. Ndou [18] acknowledges the employees as an important actor; she states that 'the relationships, interactions and transactions between government and employees in fact constitute another large e-government block, which requires a separate and very careful handling'.

A constructivist evaluation approach, opposed to the positivist approach, can provide a more extensive understanding of e-government [e.g. 16, 19]. Applying an action-based grounded theory approach [20] in collaboration with two local governments in the UK, they aimed at 'seeking to increase the understanding and knowledge of e-government evaluation' [16]. This was done in an interpretive and inductive process leading to an understanding of 'social and human aspects of e-government evaluation'. The studies led to important themes for evaluating e-government, namely decision making, evaluation methods, what and how to assess and how the practitioners perceive the evaluation process. Moreover, the dilemma of agency was stated to dominate e-government. A grounded approach was applied by Irani et al. [21] with a series of workshops in the UK with e-government practitioners. The workshops inductively revealed the practitioners' perception of challenges in e-government of technological, social and organizational themes, e.g. lack of interoperability, shared services, legal issues and inter-governmental coordination and collaboration.

The degree of adoption of a new technology is argued to be dependent on the information decision process of an individual thus on relative advantages, compatibility, complexity, trialability and observability [22]. Gallivan [23] argues that adoption can be understood as primary adoption at the organizational level and subsequent adoption at the individual level. He finds from an empirical study that managerial intervention and captive use facilitated the adoption process. From an empirical study Braun et al. [24] state the importance of also considering the contextual factors at the organization level, i.e. value management capabilities, integration into managerial processes and support from top management. Tyre and Orlikowski [25] saw, that technology use congeals after a short time and that further adoption requires interventions into the organization.

3. Method

This study was conducted as Action research (AR) to explore the implementation of Digital Post in a Danish municipality. AR is based on a certain action in a particular setting that creates a response where the social action can be connected to a causal model [26]. It has the ability to create knowledge about deficiencies in the practitioners' world 'that research of a more positivist nature fails to do' [27]. AR is an appropriate methodology to investigate IS in organizations with its 'explorative yet rigorous nature, grounding in principles and methods' [28]. AR is conducted in this study as a cyclical process with five phases, namely diagnosing, planning, action, evaluating and specifying learning [26].

Digital Post constitutes a major cornerstone in the Danish 2011-2015 e-government strategy [4]. The system is basically an e-mail system in which identified actors can communicate encrypted. The empirical setting constituted the Citizen Ser-

vice Centre department in the municipality of Assens (ACS). ACS implemented Digital Post in 2010. After two years they had a suspicion that Digital Post was not used by the staff and they had not seen any drop in postal costs. ACS had two managers, and the department of 40+ employees was divided into 8 teams. The teams worked within a range of different public sector administrative services.

Data collection covered postal costs, number of transactions, staff surveys, focus groups [29], interviews with staff, managers and head of department together with notes from researcher's observations and diary from the two managers to reflect on their own learning [30].

4. Results

The number of messages through the OM went from 8 in March to around 1000 in June and July, performed by 5 workers in March and nearly all workers through July. Diagnosis was done by analyzing transaction data. Diagnosis revealed a very low adoption rate of the OM after the initial technical implementation. This worked as the overall baseline. An evaluation report was elaborated on a monthly basis. The report was a recurrent evaluation of the adoption and included quantitative and qualitative evaluation on department, team and employee levels. The report also included a list of barriers to adoption, together with an action plan, stating responsibility for action towards barriers. The report stated advice to the managers to decide from.

The survey of office workers' e-government readiness and attitude showed skepticism and negativity towards digital post and a very low rate of workers' own use of digital post as a citizen. Managers observed a variety of 'bad excuses' (their expression) for not using digital post.

> When I'm around and ask about why mail is sent physically and not digitally, I can hear that there is opportunity for development both in attitude and in terms of skills. So there will be enough to deal with when you arrive (Karen, manager, e-mail, April 25, 2013).

On the basis of the survey, from the diagnosis phase, two focus groups were conducted, distributed on age and with different attitudes to e-government in each group. Both groups included workers with positive and negative attitudes and different skills. The focus groups were planned solely to contribute to the research purpose. However, the focus groups turned out to also release some of the insecurity and tension about digital post hence reduced the skepticism and negativity.

Knowledge about barriers could be extracted from the work practices where workers would use physical mail. In order to identify these we planned to have workers to register every physical mail with a type indication in one week.

Many of the barriers were external, of which the municipality only had very limited impact, if any. Of the internal barriers, the municipality had control, but had only limited capabilities regarding configuration of the various systems and the over-all internal interoperability.

The involvement of the workers and having them to reflect on their own behavior together with on-going follow-up from management turned out to be necessary to maintain momentum in the realization process. Moreover, the 'disturbance' of the researcher and my motivational and creative capabilities together with the ability to create good personal relationships with all levels was stated as pivotal, especially the ability to meet the workers with respect and curiosity. These competences are stated by Mumford [31] as a prerequisite for succeeding with action research. Removing the

first barriers within hours created a certain momentum. The high momentum and the agile decision making was stated by the head of department as very uncommon for the organization, thus it made a great impact on the engagement of the participants.

The specified learning, elicited from interviews with head of department and managers, focused on assessments, role of management and involvement of workers. Assessments were agreed upon as pivotal for the formative evaluation leading to eliminated barriers and subsequently enhanced value realization. Expectation clarity (in form of orders) both from head of department towards managers and managers towards workers had the department focus on the Digital Post value. The importance of the operational managers as being in charge of the change process (and not the IT department) was stated by all. The study revealed specific types of barriers to digital post, see Table 1.

Table 1. Types of barriers to digital communication

Barrier type	Explanation	Example
External work processes (mainly other public organizations)	Processes, based on physical documents, stamps, signatures etc. that involves the municipality, legislative issues and missing awareness from externals	Housing loans, where legal text must be on the back of the loan document (Ministry of housing), cannot be sent digitally.
External technical infrastructure issues	Subject matter systems not integrated with Digital Post and the complexity of the digital post system	The welfare aid system (vendor has monopoly), used by all municipalities. Configuration failure in the Digital Post system.
Internal technical infrastructure issues	Local systems interoperability and configuration issues	Case handling system was not configured correctly
Managerial decisions	Internal work processes involving physical documents or other communication channels	ACS sent physical tickets to elderly people to pay for therapy, workers promise costumers not to send digital letters

5. Discussion

According to survey, staff uses between 10 and 20 systems every day together with Digital Post. One type of barrier to adoption constitutes other systems being incompatible with Digital Post. Failure of interoperability has roots in the Weberian bureaucracy model and is recognized as a true barrier to e-government [32]. Trialability will stimulate adoption of new innovations [22]. Digital Post was designed without the opportunity for staff to experience use of the system without involving real citizens. Together with the technology illiteracy that may characterize many skilled staff, this will induce uncertainty, hence slower the adoption.

This case revealed several policy issues, including both failure from legal recognition of digital documents and privacy issues. Failure to acknowledge digital documents is obviously fatal to Digital Post. The trustworthiness of this e-government initiative is under strain when civil servants experience that other parts of public sector do not recognize digital documents. The most serious obstacle is the uncertainty about whether the civil servant may extract CPR from the Civil Registration System without committing a crime. The legal department of the Danish Digitization Agency confirms that civil servants may do this, but they have not gone public with this. The law abiding gene is very strong amongst administrative workers in the public sector, thus con-

firmation from the Danish Data Protection Agency would resolve this. Many researchers find legal barriers to e-government [18, 33, 34].

The second issue lies with configuration of systems. Why is a system configured to send the expensive A-mail for three years without anyone noticing it? The economic constraints of the public sector lead to mere technical implementations. Systems are typically implemented by vendors and the local IT department from a default IS configuration scheme. It is pivotal for e-government initiatives that the business manager is in charge of system configuration and has the necessary capabilities.

Capability of producing the necessary and sufficient business intelligence (BI) is necessary to be able to document value and inform adoption process. In this case there were insufficient capabilities to extract, manipulate and present the BI data. The BI data revealed a potential major breakdown caused by workers omitting Digital Post. This led to an intervention that stopped this. Furthermore, the case revealed lack of project management and change management capabilities at the management level. The researcher acted like the 'visionary change agent' that Chircu and Lee [35] states as one of six e-government adoption key success factors. This is a challenge to local governments with busy work schedules.

It is vital for public institutions to adopt an e-government initiative when funding is reduced according to anticipated use from the initiative. Digital Post has existed since 2010 but only in 2013 was the state funding reduced, which have had an effect on local governments' CEOs. The municipality had to cut costs or reduce welfare service elsewhere. The CEO ordered the use of Digital Post and this was not questioned by managers or workers. This supports the response to authority claim as a major impact on adoption [36], which is also stated from a multiply case study by Chircu and Lee [35]. Secondly, the study shows how important it is to integrate the current follow-up into the managerial practice as stated by Braun et al. [24].

The skilled administrative worker in the local government carries a long and strong tradition of physical writing and serving the citizen. The focus groups in this study proved that making the mail digital conflicts with these values in many ways. Rogers [22] asserts that adopting a new value system is a very slow process. Captive use is also stressed as an important adoption factor [23], which was proven in this case by the mandated use of CPR. The two vendors in the market of the OM both declare that they will not enforce this feature on customers. The case of the Digital Post and use of the OM is characterized by a high degree of ease of use. A

This was confirmed by many workers throughout the project. The greater the perceived relative advantage is, the faster the adoption [22].

Partnership and collaboration are important elements of the e-government development process [18]. Collaboration with system vendor was important regarding support for configuration knowledge and export/interpretation of data. Collaboration with other public institutions was confirmed in this study as important for the adoption process. Several of the adoption barriers originated from lack of recognition of digital letters from other public institutions. This siloization is recognized by Bannister and Connolly [32] as a major barrier to e-government.

Leadership is necessary before, during and after project implementation' [18]. This study explicitly 'instructed' managers, head of department and head of division to act and perform management and leadership. Especially articulating the expectations towards staff and clearly 'giving orders' of how to perform work tasks was proven vital. This was explicitly mentioned by managers when specifying learning, i.e. addressing head of department and head of division. Staff also referred to 'the order'

many times, i.e. addressing the managers and head of department. The disclosure of measurements and status reports made it apparent to others, e.g. if one manager had not made a follow up on specific barriers. This integration of how to enhance value from e-government into the managerial processes was found to be key in e-government [24]. In the study, the managers exerted management, which was reflected upon both by managers themselves and staff, as a change, that enhanced perceived work satisfaction. This effort mitigated the role of agency, which is said to be significant in e-government [3, 16].

6. Conclusions

In this study, I show how a formative evaluation model of e-government adoption can enhance value from a specific e-government case: Adoption of Digital Post in a municipal citizen service center. Applying the evaluation framework in an AR approach revealed various barriers, within technical infrastructure (lack of interoperability), legal issues (unclear regulation and uncertainty about privacy issues), lack of human capabilities (project management, data management, systems configuration), change management issues (conflicts in value systems, resistance to change, vague and unclear management) and collaboration issues (other governmental institutions being opposed to digital post). To overcome barriers within the municipality, the study introduced interventions as clear leadership and management, authority based decisions, disclosure of individuals' behavior and clear e-government strategy, combining measurements with on-going changes in systems and work practices. This was overall done in a mutual collaboration between researcher and managers, but foremost with respectful involvement of staff.

The adoption of Digital Post increased to saturation, but postal costs only dropped one third. A number of e-government adoption barriers remained as unsolved. These barriers originated primarily from external factors. Interoperability problems (systems that could not integrate to digital post), unclear legal issues and governmental institutions, that did not acknowledge digital letters constituted remaining barriers.

References

[1] Ward, J., S. De Hertogh, and S. Viaene. *Managing benefits from IS/IT investments: An empirical investigation into current practice*. in *System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on*. 2007. IEEE.
[2] Coursey, D. and F. Norris, Models of E-Government: Are They Correct? An Empirical Assessment. *Public Administration Review*. **68**(2008). 523-536.
[3] Goldfinch, S., Pessimism, computer failure, and information systems development in the public sector. *Public Administration Review*. **67**(2007). 917-929.
[4] The Danish Government, Danish Counties, and Local Government Denmark, *The Digital Path to Future Welfare*. 2011: Cph.
[5] Rose, J., et al., eds. *IT Management in Local Government: The DISIMIT Project*. 2012: Aalborg, Denmark.
[6] Heeks, R. and S. Bailur, Analyzing e-government research: Perspectives, philosophies, theories, methods, and practice. *Government Information Quarterly*. **24**(2007). 243-265.
[7] Yıldız, M., Big questions of e-government research. *Information Polity*. **17**(2012). 343-355.
[8] Luna-Reyes, L.F., J.R. Gil-Garcia, and G. Romero, Towards a multidimensional model for evaluating electronic government: Proposing a more comprehensive and integrative perspective. *Government Information Quarterly*. **29**(2012). 324-334.

[9]   Chircu, A.M., E-government evaluation: towards a multidimensional framework. *Electronic Government, an International Journal*. **5**(2008). 345-363.

[10] Carbo, T. and J.G. Williams, Models and Metrics for Evaluating Local Electronic Government Systems and Services. *Electronic Journal of e-Government*. **2**(2004). 95-104.

[11] Bélanger, F. and L. Carter, Trust and risk in e-government adoption. *The Journal of Strategic Information Systems*. **17**(2008). 165-176.

[12] Carter, L. and F. Bélanger, The utilization of e-government services: citizen trust, innovation and acceptance factors *. *Information Systems Journal*. **15**(2005). 5-25.

[13] Horst, M., M. Kuttschreuter, and J.M. Gutteling, Perceived usefulness, personal experiences, risk perception and trust as determinants of adoption of e-government services in The Netherlands. *Computers in Human Behavior*. **23**(2007). 1838-1852.

[14] Moon, M.J., The evolution of e-government among municipalities: Rhetoric or reality? *Public Administration Review*. **62**(2002). 424-433.

[15] Norris, D.F., *E-Government at the American Grassroots: Future Trajectory*, in *38th Hawaii International Conference on System Sciences (HICSS-38)*. 2005, Computer Societry Press: Island of Hawaii (Big Island). p. 125b (1-8).

[16] Jones, S., et al., *E-Government Evaluation: Reflections on Two Organisational Studies*, in *39th Hawaii International Conference on System Sciences (HICSS-39)*. 2006, Computer Societry Press: Kauai. p. 76a (1-8).

[17] Yildiz, M., E-government research: Reviewing the literature, limitations, and ways forward. *Government Information Quarterly*. **24**(2007). 646-665.

[18] Ndou, V., E-government for developing countries: opportunities and challenges. *The Electronic Journal of Information Systems in Developing Countries*. **18**(2004).

[19] Irani, Z., et al., Evaluating e-government: learning from the experiences of two UK local authorities. *Information Systems Journal*. **15**(2005). 61-82.

[20] Glaser, B.G., A.L. Strauss, and E. Strutzel, The discovery of grounded theory; strategies for qualitative research. *Nursing Research*. **17**(1968). 364.

[21] Irani, Z., T. Elliman, and P. Jackson, Electronic transformation of government in the UK: a research agenda. *European Journal of Information Systems*. **16**(2007). 327-335.

[22] Rogers, E.M., *Diffusion of innovations. Fifth edition*. 2003, New York: Free Press.

[23] Gallivan, M.J., Organizational adoption and assimilation of complex technological innovations: development and application of a new framework. *ACM Sigmis Database*. **32**(2001). 51-85.

[24] Braun, J., F. Ahlemann, and K. Mohan, Understanding Benefits Management Success: Results of a Field Study. (2010).

[25] Tyre, M.J. and W.J. Orlikowski, Windows of opportunity: Temporal patterns of technological adaptation in organizations. *Organization Science*. **5**(1994). 98-118.

[26] Baskerville, R. and A.T. Wood-Harper, Diversity in information systems action research methods. *European Journal of Information Systems*. **7**(1998). 90-107.

[27] Susman, G.I. and R.D. Evered, An assessment of the scientific merits of action research. *Administrative science quarterly*(1978). 582-603.

[28] Davison, R.M., M.G. Martinsons, and N. Kock, Principles of Canonical Action Research. *Information Systems Journal*. **14**(2004). 65-86.

[29] Johnston, R.J., et al., Contingent valuation focus groups: insights from ethnographic interview techniques. *Agricultural and Resource Economics Review*. **24**(1995). 56-69.

[30] Baskerville, R. and A. Wood-Harper, A Critical Perspective on Action Research as a Method for Information Systems Research. *Journal of Information Technology*. **11**(1996). 235-246.

[31] Mumford, E., Helping Organizations to Change. *Qualitative Research in International Settings: Issues and Trends*(2001). 46.

[32] Bannister, F. and R. Connolly, Forward to the past: Lessons for the future of e-government from the story so far. *Information Polity: The International Journal of Government & Democracy in the Information Age*. **17**(2012). 211-226.

[33] Gil-Garcia, J.R., *What Does Lead to Highly Functional State Websites? Exploring the Effects of Organizational, Institutional, and Contextual Factors*, in *2005 Information Resources Management Association International Conference*. 2005: San Diego, California. p. 1-25.

[34] Norris, F. and C. Reddick, Local E-Government in the United States: Transformation or Incremental Change? *Public Administration Review*. **73**(2012). 175-185.

[35] Chircu, A.M. and D.H.-D. Lee, E-government: key success factors for value discovery and realisation. *Electronic Government: An International Journal*. **2**(2005). 11-25.

[36] Cooper, R.B. and A. Bhattacherjee, Preliminary evidence for the effect of automatic responses to authority on information technology diffusion. *ACM SIGMIS Database*. **32**(2001). 36-50.

A Data Oriented Approach to Derive Public Administration Business Processes

Riccardo Cognini[a], Flavio Corradini[a], Fabrizio Fornari[a],
Andrea Polini[a] and Barbara Re[a]

[a]*Computer Science Division, School of Science and Technologies*
University of Camerino 62032 -- Camerino (MC), ITALY
{name.lastname@unicam.it}

Abstract. The delivery of services to citizens by Public Administrations requires to set up and coordinate complex Business Processes. Typically homogeneous Public Administrations, such as municipalities, have to provide the same services to all citizens. Nevertheless their concrete implementation, and the supporting Business Process model and data object models, can slightly differ from one Public Administration to the other due to organizational factors. If such variability is not explicitly represented and managed, each office will have to reflect on and analyse the requirements posed by the delivery of the service; then they will have to derive a specific process and data model. On the other hand the explicit modeling of variability can reduce the work to be done and permits to define general specifications from which specific model variants can be derived according to specific needs. In this paper we propose a novel approach, inspired by Feature Modeling techniques, for data object variability modeling that can be used to provide high level blueprints from which detailed Business Processes and data object specifications can be derived. Finally, a complex scenario has been applied to validate the approach with encouraging results.

Keywords. Business Process Variability, Data Object Variability, Feature Modeling

1. Introduction

Complex organizations have to support many different Business Processes (BPs) to provide services/products to customers and to reach organization objectives. Typically BPs can present similarities when applied to support similar services or to make similar products. The explicit management of such variability, both with respect to documents (following referred as data objects) and to the activities to carry on, can be extremely useful in order to reduce complexity. Therefore methods explicitly supporting the variability modeling and the management of data objects and of BPs are very much needed [1]; the delivery of services to citizens by Public Administrations (PAs) can be re-conducted to such a situation. The case of PA is particularly interesting in reference to the possibility of modeling and managing data and BP variability [2]. At a certain level of abstraction all PAs will share the same abstract data object and BPs. Nevertheless, when details have to be defined in order to concretely support the service delivery, the BP models start to differentiate in order to include organization dependent characteristics [3]. When variability is not explicitly managed, the result is that each PA works independently to design its BPs. This certainly causes a waste of

time and money since most of the BPs requirements are the same also for other PAs, therefore its explicit modeling will permit to save time and money to PAs, and at the same time common modeling will reduce interoperability risks.

Business Processes in the PA sector are regulated by the legislation for all PA delivering a service, nevertheless some degree of freedom is left regarding the concrete implementation of such supporting BPs. This means that some behavior is not completely generalizable since each single PA has specific needs, for instance in relation to the location, the size of the territory which it covers and other many possible factors.

According to the given scenario we present a novel approach to model public services from laws, starting with the generation of a shared data view in order to reduce the cost of the BP design step. The data object view allows to have an immediate idea of which are the data involved and provides a way to facilitate the development of Configurable Process Models (CPMs). A CPM consists in a single model that groups many BP variants sharing a common goal [4] then it provides rules defining how each single BP variant can be modeled and adapted to each single PA. In our case, we use a Feature Model (FM) for the data object modeling and the Business Process Feature Model notation to represent CPMs.

The paper is organized as follows. In Section 2 background material is provided, then the proposed approach is presented in the Section 3, and its application to a real case study is presented in Section 4. Finally, Section 5 presents related works and Section 6 closes the paper drawing some conclusions and opportunities for future works.

2. Background

2.1. Feature Modeling

Feature modeling is an approach emerged in the context of Software Product Lines to support the development of a variety of products from a common platform introducing the concept of a family of BPs [5].

A FM is a graphical model which makes use of a tree representation, with the root representing the general product to develop; it permits to express different relationships among the possible features that can be included in a specific variant of the product. In particular, in the first feature modeling approach proposed, named Feature-Oriented Domain Analysis (FODA), *mandatory*, *optional* or *alternative* constraints on features have been introduced [6]. A *Mandatory* feature represents a characteristic that each product variant must have. For instance considering the production of different mobile devices we could define a constraint requiring that any mobile device variant has to include a screen. An *Optional* feature is used to represent characteristics that a product can have, but a fully functional product can also be derived without including such a feature. For instance this could be the case of mechanisms supporting connection to 4G networks that could be included only in high-profile products. An *Alternative* feature represents characteristics that cannot be present together in a product. For instance a mobile device can have a standard screen or a touch screen, but not both. It is also possible to express relationships such as "*at least one feature in a set of features is needed in each product*"; this is done via *OR features* constraints. Additionally, *include relationship* constraints have been added to express that a feature selection implies the

selection of another feature that is on a different branch of the tree; *exclude relationship constraints*, instead, are used to express that a feature selection requires to discard another feature that is on a different branch of the tree.

Once a FM has been defined it is possible to derive, according to the constraints, a specific product defining a *configuration* that expresses explicit features selection. In Figure 1 we depict a simple scenario of a family of mobile phones. Each mobile phone has to provide a display that can be touch screen or a standard display. Each phone can also have Internet connection, if it is included the 3G connection has to be mandatory and, optionally, 4G connection can be available. The features in gray in the FM are selected and they represent the configuration to generate a mobile phone product variant with a touchscreen display and a 3G Internet connection.

Researchers have proven that basic FM models are too restrictive to represent all the relationships between features useful to characterize a family of products [7]. As a result the FM notation has been extended to permit the definition of feature cardinality, permitting to define how many features in a set are needed to have a working product [8]. It means that features can be arranged into feature groups, where each feature group has a group cardinality. A group cardinality is an interval of form $<m\text{-}n>$, where $m, n \in \mathbb{Z} \land 0 \leq m \leq n \leq k$ where k is the number of features in the group.
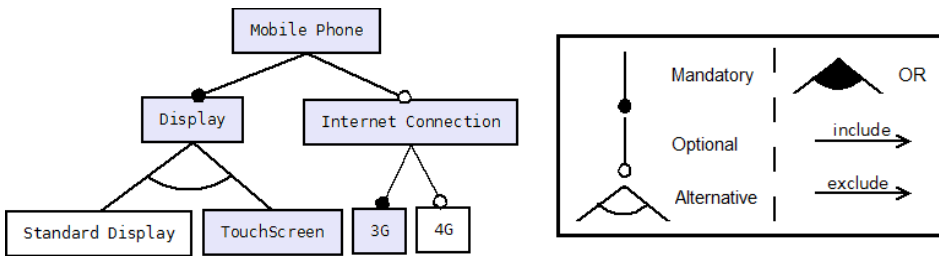

Figure 1. Feature Model Example and Configuration (Selected Features in Gray.)

2.2. Business Process Modeling

A BP is ``a collection of related and structured activities undertaken by one or more organizations in order to pursue some particular goal'' [9]. The accuracy of the BP modeling phase is critical for the success of an organization, in particular in scenarios in which it is necessary to adapt to changing requirements. In order to design a BP, different classes of languages have been investigated and defined.

In our work we refer to BPMN 2.0, an Object Management Group standard [10]. It is the most used language by domain experts due to its intuitive graphical notation. We have mainly used process diagrams, focusing on the point of view of system users. The BPMN 2.0 elements in Figure 2 are the core elements of the language and those we will use in our approach. In particular the following are concepts that can be modeled in BPMN 2.0, and their respective interpretation. Events are used to represent something that can happen. Activities are used to represent a generic work to perform within a BP. Gateways are used to manage the flow of BP both for parallel activities and choices. Data Objects permit to model documents, data, and other artifacts used and updated during the BP, in most of the cases activities take data objects as input, modify them, and give them back as output.
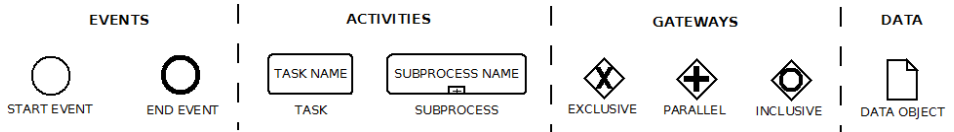
Figure 2. BPMN 2.0 Core Elements.

3. Approach

The proposed approach is organized in three main steps as shown in Figure 3. The main actors involved in the approach are: a regional *Competence Center* and a set of PAs. The *Competence Center* should include BP designers, domain experts and legislation experts; instead, the PAs are the organizations that have to deliver the services described in the referring laws. In particular the approach is organized as follows.

- The first step aims at defining a data object model that explicitly provides information about documents that are needed by PAs. This step includes knowledge acquisition through the study of laws of the involved PA services. It is carried on only once by the competence center.
- The second step aims at defining general models that constitute the basis for the definition of process variants for each specific deployment context. General models will be codified using CPMs notation, and they will be designed considering the data object model and the referring laws. This step should be carried on only once by the competence center and it will permit to derive models that will include activities that have to be carried on, the relations among them, and the data they possibly get in input or produce in output. The activities defined in the CPMs will be also linked to the data objects of the data object model.
- The last step concerns the derivation of the fully specified BP variants configuring the CPMs and the data object model. Each involved PA has to select the activities in the CPMs considering its organizational structure, then a set of BP variants is extracted (exactly one for each CPM).

Data Modeling via Feature Model. A data object model is a representation of data that may help a BP Designer during the design of a BP. In particular, with a data object model, a BP Designer can have an immediate view of the data objects that are *mandatory* or *optional* for a correct BP configuration. Then, having an overview of the data to include in the BP can facilitate the choice of which activities have to be included in the BP (e.g. activities that operate on those data). A data object model is generated from the first step of our approach and it refers to a high level representation of the data objects involved in the BP without going into details like their organization in tables, columns, or the physical means used to store them. In our case we use FM to represent the structure of data objects, whether they are composed by other data objects and to express variability in their composition. Moving from the root to the leaves means going from a data object to the parts that compose it.
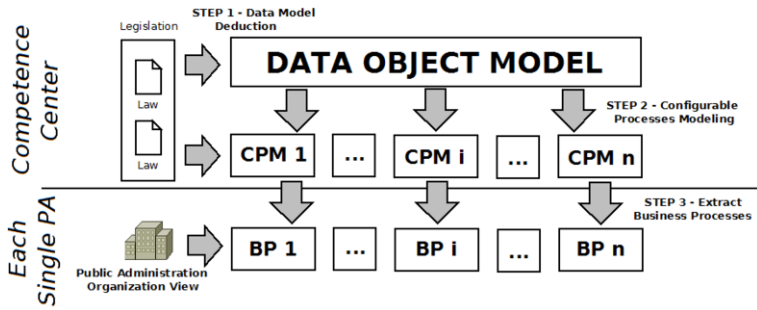
Figure 3. The Proposed Approach.

Configurable Processes Modeling via BPFM. A set of CPMs is the output of the step 2 of the approach, since many BPs can be described by a law. In this paper we use Business Process Feature Model (BPFM), a feature model extension to deal with variable BPs; its suitability to model families of BPs for PAs, has been proved in [11] [12]. A BPFM model is constituted by a tree of related activities. The root identifies the family of the BPs under analysis. Each internal (non-leaf) activity denotes a sub-process, and the external (leaf) activity represents a task. BPFM introduces different levels of detail in the BP family specification going up- down on the tree. BPFM also provides the possibility to define constraints between activities in two adjacent levels of the tree. Each constraint has only one father activity, and it has one or more child activities depending on the type of the constraint. Constraints are used to express if child activities have to be inserted in the BP variant and if they have to be included in each execution path of a BP variant.

Data Objects toward Configuration and Variant Derivation. In the third step each PA has to derive their BP variants from the data object model and the CPMs resulted in the steps 1 and 2. The PA has to consider its internal structure in order to select the data objects and the activities it needs in the BPs. The activities selection has to be done for each CPMs.

The selection of activities is generally done manually by a BP designer which has to select the activities to include in the BP. Since the Competence Center also provides a Data Object Model together with the CPM, the BP designer may be guided, in the choice of those activities, by the presence of the constraints imposed on the data objects.

4. SUAP Case Study

To describe our approach we focused on the SUAP case study. It involves more than 110 different BPs that support more than 150 data objects types in order to provide services to citizens. In particular, we applied the approach to the case of the Marche Region (Italy) that coordinates the implementation of the SUAP service for 239 municipalities or aggregations of municipalities, and 45 third parties administrations that can be involved as third parties. For the sake of space we focus in this paper to only one BP related to the *Standard request to start a Business Activity* for shops in a

fixed location. This is the BP in which an entrepreneur requests to the municipality, and to third parties organizations, the permission to start a business activity. The entrepreneur will have to wait the office decision before starting the activity.
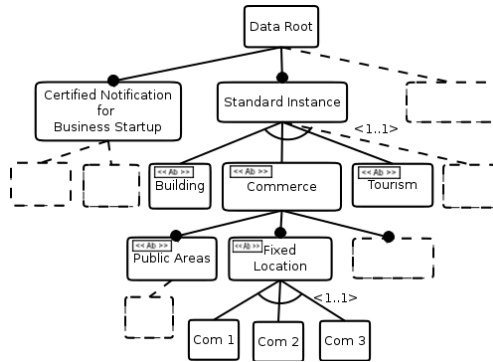
Figure 4. SUAP Data Object Model.

The Data Object Model of the SUAP case study is reported in Figure 4. It allows us to express that the Data Root must be composed by data objects such as: Certified Notification for Business Startup, Standard Instance, and other data objects that we do not represent for a matter of space. The use of cardinality in the feature model allows us to express that the data object instance must include only one part that composes it out of: Building, Commerce, Tourism and others. We use abstract data object to represent a categorization of data objects that go under a same topic; Commerce, in this case, groups together data objects such as: Public Areas, Fixed Location, and others. It is also mandatory that Commerce is composed by only one of those data objects. Finally, in our case study the Fixed Location data object must be composed by one out of the parts: Com 1, Com 2, Com 3. Representing data in this manner, gives to the BP designer the possibility to have a view of all the data objects that are necessary for configuring the business processes required by the law.

Figure 5. SUAP Configurable Process Variant.

After modeling the data view, the second step of our approach implies that all the CPMs are developed by a Competence Center. In this paper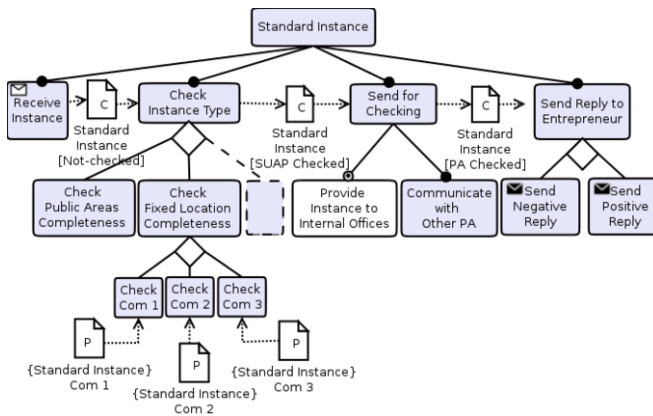 for a matter of space we just present the CPM of the *Standard request to start a Business Activity*, considering only the activities needed for the case of fixed location shop; we modeled it in BPFM notation. Since the paper is not focused on BPFM notation, then we just highlight the variability of the case. In particular, two possible BP variants can be generated from the BPFM model, one in the case of the SUAP office is provided by a single municipality and the other in the case the office is provided by an aggregation of municipalities. In the model it is represented by a *BPFM Special Case Constraint* linked to the activity *Provide Instance to Internal Offices*.

Performing the third step of our approach, the Competence Center has already designed the Data Object Model and the CPM; the BP designer of each single PA can then derive a BP variant defining a configuration that permits to express which data and activities to include. The activities in gray in the FM of Figure 5 represent a possible configuration, while Figure 6 illustrates the corresponding BP variant derived thanks to the mapping rules defined in BPFM.
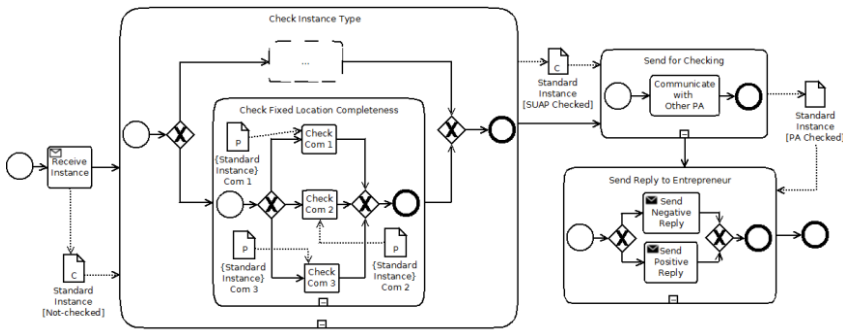
Figure 6. Business Process Variant.

5. Related Works

We found other works that attempt to provide approaches and tools to simplify the management of Business Processes in the Public Administration sector. In [2] the authors underline the importance of using Configurable Process Models for the Swedish Public Sector, providing also positive results of evaluation questionnaire and feedback from municipal officials.

Other works attempt to derive a Business Process model directly from the law. Particularly interesting is the approach used in [13] which focuses on the natural language that composes the law and the importance of having a visual representation such as a derived Business Process model. The authors derive a formal, and therefore verifiable, model directly from the code of law to proof the thesis that many laws directly implicate executable process sets. They use an existing graphical and textual illustration of the Swiss obligation law as basis and they were able, with the help of the Semantic Process Language [14], to translate the illustration into an executable Module Net.

In our approach we do not refer to a direct translation of the law text into a model, but we rely on the presence of a competence center that does this work manually.

6. Conclusions and Further Work

In this paper we presented an three steps approach to derive public services BPs from legislation. In the first step a data object model is extracted from legislation by a competence center, then in the second step the competence center models a set of BP families highlighting the variability between the BP variants. In the last step each single PA extracts its own BP variant from the families with the help of the data object model. We applied the approach to the Marche Region SUAP scenario in which 239 municipalities and 45 third parties administrations are involved in the execution of more than 110 BPs. This first experiment made with the proposed approach, provided encouraging results and permitted to model quite easily a complex scenario, and to derive the corresponding BPs.

In the future we plan to apply the approach to other scenarios in the Italian Public Administration, including a check of the compliance between the data model and the CPMs in order to ensure the correctness of the models and of the relations modeled by the competence center.

References

[1] Reichert, M. and Weber, B.: Enabling flexibility in process-aware information systems: challenges, methods, technologies. Springer Science & Business Media (2012)

[2] Lönn, C.M., Uppström, E., Wohed, P., Juell-Skielse, G.: Configurable process models for the Swedish public sector. Advanced Information Systems Engineering. Springer (2012) 190-205

[3] Erkoçak, E., Açıkalın, S¸ .N.: Complexity theory in PA and metagovernance. In: Chaos, Complexity and Leadership 2013. Springer (2015) 73-94

[4] van der Aalst, W.M., Dreiling, A., Gottschalk, F., Rosemann, M., Jansen-Vullers, M.H.: Configurable process models as a basis for reference modeling. In: Business Process Management Workshops, Springer (2006) 512–518.

[5] Pohl, K., B¨ockle, G., Linden, F.J.v.d.: Software Product Line Engineering: Foundations, Principles and Techniques. Springer-Verlag (2005)

[6] Kang, K.C., Cohen, S.G., Hess, J.A., Novak, W.E., Peterson, A.S.: Feature-oriented domain analysis (FODA) feasibility study. Technical report, DTIC (1990)

[7] Capilla, R., Bosch, J., Kang, K.C.: Systems and Software Variability Management. Springer (2013)

[8] Czarnecki, K., Helsen, S., Eisenecker, U.: Staged configuration using feature models. In: Software Product Lines. Springer (2004) 266–283

[9] Lindsay, A., Downs, D., Lunn, K.: Business process - attempts to find a definition. Information and Software Technology 45 (2003) 1015–1019

[10] OMG: (Business Process Model And Notation - BPMN)

[11] Cognini, R., Corradini, F., Polini, A., Re, B.: Modelling process intensive scenarios for the smart city. In: EGov - 13th IFIP WG 8.5 International Conference, EGOV 2014, Dublin, Ireland, September 1-3, 2014. Proceedings. (2014) 147–158

[12] Cognini, R., Corradini, F., Polini, A., Re, B.: Using data-object flow relations to derive control flow variants in configurable business processes. In Fournier, F., Mendling, J., eds.: BPM 2014 Workshops, LNBIP 202. (2014) 210 – 221

[13] Olbrich, S., Simon, C.: Process modelling towards e-government–visualisation and semantic modeling of legal regulations as executable process sets. Electronic Journal of e-Government 6(1) (2008) 43–54

[14] Simon, C.: Integration of planning and production processes. Mathmod 2006, Special Session: Petrinets: Current Research Topics and their Application in Traffic Safety and Automation Engineering (2006)

Towards a "Balanced" Historical Assessment of E-government Research

Lemuria CARTER[a,1]
[a]*Virginia Commonwealth University*

Abstract. This paper utilizes the Balanced Scorecard to provide a forward looking historical assessment of e-government research. The Balanced Scorecard provides an interesting and strategic lens through which to view e-government research. This paper identifies prominent e-government articles to highlight the financial, customer, internal process and organizational capacity elements of e-government research. This paper is not a traditional literature review; rather, it is a call for a more strategic approach to e-government research. As the field matures, the community should identify our mission and vision for e-government research and target our efforts accordingly.

Keywords. E-government, Balanced Scorecard, Literature Review

1. Introduction

Researchers in public administration have been exploring the role of technology in government for years. In 1926, Leonard White, a historian who specialized in public administration and later won the Pulitzer Prize for history, discussed the transformation of office equipment in the public sector. Before "e-government" emerged, researchers and practitioners identified ways to harness technological advancements for government services (Beard 1931).

This paper provides a forward looking historical assessment of e-government research over the last two decades. Although the discussion of information communication technologies (ICT) in government is centuries old (Bain 1937; White 1926), the focus of this paper is on the evolution of the "e-government" phenomenon. The remainder of this paper provides an overview of the evolution of the term e-government, a summary of the balanced scorecard, a review of e-government articles that are highly cited and published in the "Basket of Eight" IS journals, and recommendations for future research based on a "balanced" categorization of the e-government articles.

2. E-government

Researchers and practitioners are constantly conducting studies to assess government interaction with its constituents. Caudle et al. (1991) provides the first national survey of public managers' ratings of IS issues. The survey was administered in 1988 to high-

---

[1] Lemuria Carter.

level public managers in federal, state, and county agencies. A few years later, the United States National Performance Review coined the term 'e-Government' in 1993 (Alasem, 2009). Ten years later, Chadwick and May (2003) provide a compressive definition. They state "the principal features of e-government managerialism can be summarized as follows: a concern with the "efficient" delivery of government information to citizens and other groups of "users"; the use of ICTs to improve flows of information within and around government; a recognition of the importance of "service delivery" to "customers"; the view that speeding up information provision is, by itself, "opening up" government; a general absence of user resource issues, such as ability to receive and interpret information; and "control" and presentational professionalism (often termed "spin") as defining logics." Chadwick and May (2003) explore the genesis of the move towards "e-government" in the United States, Britain, and the European Union. They posit that "the democratic potential of the Internet has been marginalized as a result of the ways in which government use of such technology has been framed since the early 1990s. An executive-driven, "managerial" model of interaction has dominance at the expense of "consultative" and "participatory" possibilities (p. 271)." In the two decades since the term "e-government" was introduced, we have seen explosive growth in e-government research fueled by targeted efforts –special issues, RFPs, etc. - to increase the body of knowledge in this area.

Johnson and Ward (1972) describe "a citizen information system. Such a system (panel in type, neighborhood in orientation, public in character) is an important link in extending the notion of citizen participation in both policy and programs in a rapidly changing society." They raise the question "what kind of citizens' participation vehicles emerge in a society which is diverse, complex, technologically oriented and constantly and rapidly changing?

3. The Balanced Scorecard

The balanced scorecard is a "strategic planning and management system that is used extensively in business and industry, government, and nonprofit organizations worldwide to align business activities to the vision and strategy of the organization, improve internal and external communications, and monitor organization performance against strategic goals Anonymous (2014)." It was created by Kaplan and Norton (1992) to highlight the importance of both traditional financial metrics and strategic non-financial indicators. Hence the proposed performance measurement framework offers managers and executives a more 'balanced' view of organizational performance. The balanced scorecard includes four measurement categories: financial, customer, internal process and organizational capacity

According to Gartner Group, more than fifty percent of large firms in the United States firms utilize the balanced scorecard (BSC). An international study by Bain & Co ranked the balanced scorecard fifth on its top ten most widely used management tools around the world (Anonymous 2014). One of the major components of the BSC is the strategy map. "Strategy maps are communication tools used to tell a story of how value is created for the organization. They show a logical, step-by-step connection between strategic objectives (shown as ovals on the map) in the form of a cause-and-effect chain (Anonymous 2014)."

According to Van Grembergen (2005) the balanced scorecard can be applied to information technology. In this study, I use the balanced scorecard to evaluate e-

government research. This framework is well suited to this historical assessment. The Financial element refers to cost and time savings realized by e-government initiatives. The Customer is the constituent (citizen or business) that benefits from an e-government service. The Internal Process measure captures studies that evaluate e-government development, interoperability, etc. Finally, the organizational capacity refers to initiatives and studies that seek to advance e-government knowledge and tools.

4. Methodology

To identify the relevant articles published in the "Basket," I started with the articles published in Belanger and Carter (2012)'s review of e-government and the "Basket of Eight." Then, I searched for the term "e-government" in each individual basket journal via the Business Source Complete and ABI/INFORM Complete databases. Belanger and Carter (2012)'s review of the Basket included 30 articles. My search resulted in twenty-nine additional articles, for a total of fifty-nine articles from the "Basket of Eight." To identify the top 10 most highly cited articles, I searched for the term "e-government" via the Web of Science online platform. Future studies could use a more expansive search process (including electronic government, digital government, IT in the public sector, etc.). However, the purpose of this paper isn't to provide a comprehensive overview of the literature. The goal is to start identifying high-level trends in e-government research. The purpose is to provide an overview of e-government research as it relates to a useful strategic management tool: the balanced scorecard.

In the private sector, the balanced scorecard is used to help companies link their mission and vision to financial and non-financial indicators of success. I posit that this scorecard can be used to evaluate diverse facets of e-government research and help the research community identify content areas that need more attention. Some studies could fit into more than one quadrant. I placed each study in the quadrant that is most relevant.

5. Selected Findings

The top 10 most highly cited e-government articles are presented below in table 1.

Table 1. Balanced Scorecard Strategy for the Top 10 Most Cited E-government Articles

Rank	Author	Year	J	Balanced Scorecard Strategy
1.	(Layne and Lee)	2001	GIQ	Organizational Capacity
2.	(Moon)	2002	PAR	Financial
3.	(West)	2004	PAR	Organizational Capacity
4.	(and Belanger)	2005	ISJ	Customer
5.	(Ho)	2002	PAR	Internal Process
6.	(Medjahed et al.)	2003	VLDB	Internal Process
7.	(Heeks and Bailur)	2007	GIQ	Organizational Capacity
8.	(Welch et al.)	2005	JPART	Customer
9.	(Yildiz)	2007	GIQ	Organizational Capacity
10.	(Norris and Moon)	2005	PAR	Organizational Capacity

J=Journal; WOS = Web of Science; GS=Google Scholar
*All citations are as of October 2014

The Balanced Scorecard Mapping of the 10 most highly cited e-government articles includes one Financial article, two Customer articles, 2 Internal Process articles and five Organizational Capacity articles. Half of the top 10 most highly cited e-government articles address organizational capacity. The research community has focused on improving the e-government knowledge and skills and improving the tools and utilization of technological advancements to support ICT in the public sector. This focus on organizational capacity is logical given e-governments relative infancy. The most cited papers were published between from 7 - 13 years ago. Hence, this list is a reflection of the content that was initially very important to the field. The content covered in these articles has been an integral part of e-government growth and development. As organizational capacity increases, e-government research will also evolve and emphasize other areas of the balanced scorecard such as the customer (i.e. citizen) and internal processes. Currently, the financial benefits (e.g. lower costs) are frequently touted by proponents of e-government but seldom appear in research models.

Regarding the fifty-nine e-government articles published in the "Basket of Eight," zero are Financial, twenty eight focus on the Customer, fifteen address Internal Process and sixteen explore Organizational Capacity.

Perhaps this lack of focus on the financial benefits of e-government by the research community is due to 1) our focus on the societal benefits (e.g., increased citizen participation) and 2) the public sector provides not-for-profit services. Currently, only one of the top ten most highly cited e-government papers address the Financial quadrant of the Balanced Scorecard strategy map. Arveson (2014) posits that the fundamental metric for government performance is not financial; its mission effectiveness. He states "at any given time, some departmental missions may be more important than others for the needs of the country. The selection of the departmental mission priorities is an ongoing strategic planning responsibility (np)." However, given the recent financial challenges and budget cuts experienced in many agencies, the financial benefits of e-government are increasingly more important to practitioners.

6. Moving Forward

Bain (1937) states "technology creates and destroys groups; it modifies those that survive; these groups are the fundamental societal realities with which government must deal; more accurately, they are the very stuff of the political institution. They must become an integral, functional part of political organization. Men make machines, but they also are made by machines."

Watsell and White (2010) call for more research on public services. Persson and Goldkuhl (2010) posit that e-government is a synthesis between traditional bureaucracy and new public management. In light of the Balanced Scorecard Mappings and the extent literature, I recommend the following avenues for future research in each area.

6.1. Financial

- Explore the value of e-government.
- Disseminate financial benefits and cost-savings to practitioners.

6.2. Customer

- Determine what constitutes e-government success and/or failure. Identify metrics for success for diverse initiatives, levels of government and cultural norms.
- Identify the best ways to interact with a variety of stakeholders in a variety of contexts.
- Promote transparency via open government.
- Enhance the number and capabilities of smart cities.
- Utilize social media to increase citizen participation.
- Minimize the digital divide, especially as it relates to an aging population.

6.3. Internal Process

- Utilize emerging trends in data analytics to identify trends, prevent fraud, and meet citizen needs.
- Ensure sensitive data is secure and explore the role of information systems security in the public sector.
- Develop systems that enable ubiquitous computing.
- Develop interoperable applications.

6.4. Organizational Capacity

- Identify how "e" processes compare to manual processes? Does the "e" matter? What does it change? How does "e" change what matters? How can we avoid reinventing the wheel? We should learn from existing knowledge about manual processes. Don't assume that the earlier method was ALL bad. We should start by acknowledging what we know and then move forward.

- Leverage our knowledge to improve society. Unlike other disciplines (e.g. philosophy), understanding the phenomenon is not enough; we should also improve conditions (knowing vs. leveraging what we know).
- Account for culture and context when developing e-government policy.
- Understand how e-government links to other communities.
- Promote interdisciplinary partnerships. There is a need for more partnerships between IS and Public Administration, Computer Science, Psychology, and Agricultural Economics to promote diverse concepts such as, big data, IS security, green IS, social responsibility and the triple bottom line.

The aforementioned recommendations highlight a few of the possible avenues for future research. This paper is intended to contribute to the discussion on the past, present and future of e-government research. There are a few limitations to this study. Given that most of the highly cited e-government articles are 10+ years old, to identify current gaps in the literature researchers should supplement this commentary with a review of 2014 and 2015 e-government articles. In addition to reviewing journals, future research could incorporate findings from diverse sources including books, proceedings and dissertations. This paper does not provide an e-government literature. Instead, it explores a strategic mapping tool to provide a high-level overview of highly-cited e-government research.

References

[1] A. Alasem, An Overview of e-Government Metadata Standards and Initiatives based on Dublin Core. *Electronic Journal of e-Government*, **7** (2009), 1-10.

[2] Anonymous, The Balanced Scorecard Institute: Strategy Management Group, http://balancedscorecard.org/Resources/About-the-Balanced-Scorecard (2014).

[3] P. Arveson, Translating Performance Metrics from the Private to the Public Sector, http://balancedscorecard.org/Resources/Performance-Measures-KPIs/Translating-Metrics (2014).

[4] R. Bain, Technology and state government, *American Sociological Review* **2**(1937), 860-874.

[5] F. Bélanger, L. Carter, Digitizing Government Interactions with Constituents: An Historical Review of E-Government Research in Information Systems. *Journal of the Association for Information Systems* **13** (2012), 363-394.

[6] W. Beard, Technology and Political Boundaries. *The American Political Science Review* **25** (1931), 557-572.

[7] L. Carter, F. Bélanger, The Utilization of E-government Services: Citizen Trust, Innovation and Acceptance Factors, *Information Systems Journal* **1** (2005), 2-25

[8] S. Caudle. L, W. Gorr, L, E. Newcomer , Key information Systems Management Issues for the Public Sector, *MIS Quarterly* **2** (1991), 171-188.

[9] A. Chadwick, C. May, Interaction between States and Citizens in the Age of the Internet: "e-Government in the United States, Britain, and the European Union, *Governance-An International Journal of Policy and Administration* **2** (2003), 271-300.

[10] R. Heeks, S. Bailur, Analyzing e-Government Research: Perspectives, Philosophies, Theories, Methods, and Practice, *Government Information quarterly* **24** (2007), 243-265.

[11] Ho, A.T.K, Reinventing Local Governments and the e-Government Initiative, *Public Administration Review* **4** (2002), 434-444.

[12] N. Johnson, E. Ward, Citizen Information Systems: Using Technology to Extend the Dialogue Between Citizens and Their Government, *Management Science* **19** (1972), 21-34.

[13] R.Kaplan, S. Norton, D.P, The Balanced Scorecard: Measures that Drive Performance, *Harvard Business Review* **January-February** (1992), 71-79.

[14] K. Layne, J. Lee, Developing Fully Functional E-government: A Four Stage Model, *Government information quarterly* **18** (2001), 122-136.

[15] B. Medjahed, A. Bouguettaya, A. Elmagarmid, Composing Web Services on the Semantic Web, *The International Journal on Very Large Databases*, **12** (2003), 333-351.

[16] J. Moon, The Evolution of E-Government among Municipalities: Rhetoric or Reality?, *Public administration review* **62** (2002), 424-433.

[17] F. Norris, J. Moon, Advancing e-Government at the Grassroots: Tortoise or Hare? *Public Administration Review* **1**, (2005), 64-75.

[18] A. Persson, G. Goldkuhl, Government Value Paradigms—Bureaucracy, New Public Management, and E-Government, *Communications of the Association for Information Systems* **27** (2010), 45-62.

[19] W. Van Grembergen, S. De Haes, Measuring and Improving IT Governance through the Balanced Scorecard, *Information Systems Control Journal* **2** (2005), 35-42.

[20] D. Wastell, S. White, Facts, Myths and Thought-Style and a Rallying Cry for Civic Engagement. *Journal of Strategic Information Systems*, **19** (2010), 307-318.

[21] E.W. Welch, C.C. Hinnant, M.J Moon, Linking Citizen Satisfaction with e-Government and Trust in Government. *Journal of Public Administration Research and Theory,* **15** (2005) 371-391.

[22] D.M. West, E-government and the transformation of service delivery and citizen attitudes, *Public Administration Review*, **64** (2004) 15-27.

[23] L. D. White, *Public Administration*, New York, 1926.

[24] M. Yildiz, E-government Research: Reviewing the Literature, Limitations, and Ways Forward. *Government Information Quarterly*, **24** (2007) 646-665.

Information Systems Evolution Efficiency – Differences between the Public and the Private sector

John KROGSTIE, Tor Kristian VELD
IDI, NTNU, Trondheim Norway

Abstract. This paper provides a comparison between public and private organizations on key figures relative to the information systems support activities. It has been claimed that public sector has a less satisfactory conduct of information system support than private sector. In earlier work we have tried to find support for such claims without getting significant results. Whereas these earlier investigations only had data from around 70 organizations, in this article we present selected data from an extended survey investigations performed among more than 250 Norwegian organizations on how they conduct information systems development and maintenance. This investigation has earlier been compared with similar investigations of this sort. A major finding from the previous comparisons is that even if we witness large changes in the underlying implementation technology and approaches used, a number of aspects such as the overall percentage of time used for maintaining and evolving systems in production compared to time used for development is remarkably stable. When we compare public and private organizations in the last survey on the other hand, we find statistically significant differences supporting the claims of better usage of resources for information systems evolution in private than in public sector.

Keywords. e-government, information systems development and maintenance, public vs. private sector

1. Introduction

Heeks [1] claims that "Most eGovernment systems fail." Also many other authors claim more IT-oriented problems in the public than the private sector. In [2] it is reported that "Public projects had an average effort overrun of 67%, as opposed to the 21% average in private projects. This observed difference appears to be caused by systematic differences between private and public organizations found at 1) the political level, 2) the organizational level, and 3) and the individual level". A number of reports indicate that this is an international trend [3] in OECD countries, USA and in the UK [4]. A reason these failures get so well-known though might be because they are in the public sector, thus information about success and failure is also public information. As reported e.g. in [5] failure is happening both in public and private sector, and usually, the failures are only partial; most systems get delivered and are used in some way. Since most work on IT is not on developing new systems, but on maintenance, operation and user-support [6], an alternative way to compare IT efficiency than only looking at the development efficiency, is looking upon how work is distributed in the IT-organization, and the amount of work being done on value-adding activities. Application systems are valuable when they provide information in a manner that enables people to meet their objectives more effectively [7]. An application system is part of an encompassing organizational system, which in turn is

part of a broader business environment. This environment of change that an organization must address implies that the supporting information systems also must be easily adaptable. As stated already in [8], it is one of the essential difficulties with application systems that they are under a constant pressure of change. Given the intrinsic evolutionary nature of the sources of system needs as described above, it should come as no surprise that information system must evolve as well [7].

The goal of both development activities and maintenance activities is to keep the overall information system support of the organization relevant to the organization, meaning that it supports the fulfilment of organizational goals. A lot of the activities usually labelled 'maintenance', are in this light value-adding activates, enabling the users of the systems to do new task. On the other hand, a large proportion of the 'new' systems being developed are so-called replacement systems, mostly replacing the existing systems without adding much to what end-users can do with the overall application systems portfolio of the organization [9]. Based on this argumentation we have earlier developed the concept application portfolio upkeep as a high-level measure to evaluate important aspects of to what extent an organization is able to evolve their application system portfolio efficiently. How application portfolio upkeep is different from maintenance is described further below.

In this paper, we present results from a survey-investigation performed in Norwegian organizations in this area during the end of 2013/early 2014. We have earlier compared the overall results with similar investigations done in 2008, 2003, 1998 and 1993 [10], finding a stable overall pattern of distribution of work from the last four investigations. We will in this paper look more closely at the results from the last investigation, comparing figures from public and private organizations on central variables. Norway has quite a number of companies which are defined as private, but yet having substantial public ownership, the state being a major shareholder. Also, a lot of previous public organizations have recently been transformed to private companies or state owned limited companies or other kinds of organizations with varying degrees of freedom being run more according to private business principles than what was usual some decades ago. On the other hand we find certain important traits among public organizations e.g. that they all have to abide to the same non-optimal rule of procurement found in the EU legal framework for public procurement and development of IT-solutions when external companies are involved in developing the requirements to a system, they are not allowed to be involved in the implementation of the system [11]. Public sector organizations also have a different safety net. E.g. a municipality will not cease to exist due to bankruptcy. This makes the dichotomy between private and public companies meaningful. Thus our core research question is: *Is information systems development support conducted in a less optimal way in the public sector, compared to how it is done in the private sector in Norway.*

We will first give definitions of some of the main terms used within information systems evolution. We describe the research method, including a number of more detailed hypotheses spawned from the field detailing the above research question, before the main results from our investigation are presented. Then a closer investigation on the differences between private and public sector respondents are presented. The last section summarizes our results and presents ideas for further work. The structure of the paper is quite close to [12] where a similar investigation was made, but including data from the last investigation.

2. Definition of Core Concepts

Maintenance has traditionally been divided into three types: corrective, adaptive and perfective [13] based on e.g. [14]. This vocabulary is well established both in theory and practice, and we here use the IEEE terms with some clarifications and further division also anchored in the literature:

Maintenance is defined as the process of modifying a software system or component after initial delivery to production.

- Corrective maintenance is work done to correct faults in hardware and software.
- Adaptive maintenance is work done to make the computer program usable in a changed technical environment.
- Perfective maintenance is work done to improve the performance, maintainability, or other attributes of a computer program. Perfective maintenance has been divided into enhancive maintenance [15] and non-functional perfective maintenance. Enhancive maintenance involves changes and additions to the functionality offered to the users by the system. Non-functional perfective maintenance implies improvements to the quality features of the information system and other features being important for the developer and maintainer of the system, such as modifiability. Non-functional perfective maintenance thus includes what is termed preventive maintenance, but also such things as improving the performance of the system.

In addition to the traditional temporal distinction between development and maintenance, we have introduced the concepts application portfolio evolution and application portfolio upkeep (originally termed functional development and functional maintenance when introduced [16]. At the same time concepts such as devtenance (cf. devops) were introduced).

1. Application portfolio upkeep: Work made to keep the functional coverage of the information system portfolio of the organization at the current level. This includes: Corrective maintenance, adaptive maintenance, non-functional perfective maintenance and development of replacement systems
2. Application portfolio evolution: Development or maintenance where changes in the application increase the functional coverage of the total application systems portfolio of the organization. This includes enhancive maintenance and development of new systems that cover areas, which are not covered earlier by other systems in the organizations

We note that some writers provide more detailed overview of maintenance tasks [6, 17]. Jones [6] has in total 21 categories: This includes user-support as a part of maintenance (a view shared with e.g. Dekleva [18]); an area usually looked upon as belonging to 'other work' in other overviews.

3. Research Method

In connection to this work, we have performed two related surveys. One is our main replication study covering a large number of topics matching the ones we have

investigated 4 times earlier. The other is in connection to the yearly 'IT i praksis' [19] investigation done by Rambøll in early 2014, where we have included the questions relative to work distribution from our replication study, to compare this with e.g. the benefit of IT and alignment of IT and business strategy. We present these investigations individually below.

Our replication survey was implemented in the SurveyMonkey web-tool and invitations were distributed by e-mail to 388 Norwegian organizations. The organizations were randomly selected from the list of member organizations of DnD (The Norwegian Computer Society - NCS) (NCS has currently around 1000 member organizations primarily in the private sector) and OSDF - the public sector IT-forum, to have also respondents from the public sector. 'IT i praksis' was sent out to more than 500 organizations, equally divided between the public and the private sector, and we made sure to avoid overlap between the respondents.

The form in our replication study contained 41 questions including demographic data. The contents of the form were based on previous investigations within this area; especially those described in [20-26]. The full survey form and data is available on request. In this paper, we present result combining the data from the replication study with the IT i Praksis study.

On the replication study 87 responses were returned, giving a response rate of 22%. Out of these only 68 responses could be used for the analysis. The additional responses were not complete, and in particular did not include responses to the questions relative to distribution of work. Out of 533 distributed survey forms in 'IT i praksis', 272 responses (i.e. 51%) where returned, although only 208 provided responses to the questions matching our main replication study. When we put the results together, we had an overall response-rate of 39%.

3.1. Previous Investigations

We have earlier compared some of the results of the last investigation with the results of similar investigations [10]. A number of later investigations on the distributions of work have been done, but they typically focus on the distribution of maintenance tasks only [27-29], many only looking on the situation in one organization. Potential threats to validity of such surveys are discussed in [10].

3.2. Hypothesis

To detail the main research question presented in the first section the following hypotheses were formulated to investigate the development of the different measures for distribution of work between private and public sector. Since we are looking for differences, we have formulated the hypothesis as if private and public sector are equal (to potentially refute this).

H1: There is no difference between the breakdown of maintenance work (in corrective, adaptive, enhancive and perfective maintenance) in public and private organizations. Rationale: Whereas perfective (in particular enhancive maintenance) provides more value than other types of maintenance, it is interesting to look into this breakdown. Investigations reporting on the distribution of time among maintenance tasks [30, 31] report very different numbers. On the other hand these investigations vary greatly. Whereas some look on single systems of numerous organizations or the whole portfolio of numerous organizations, other look only at one or a few (important)

applications in one organization. Since this distribution naturally will differ according to where the system is in the lifecycle (development, evolution, servicing, phase-out, closed [32]), this difference between the maintenance work on individual systems should be expected. When averaging across a large number of application portfolios on the other hand, we have found a more stable distribution.

H2: There is no difference between the percentage of time used for development in private and public sector Rationale: When comparing the percentage of time used for development activities in organizations earlier, we have found this to be decreasing, but not so much between the three last investigations. Thus is interesting to see if this is equal also between private and public sector.

H3: There is no difference between the percentage of time used for maintenance in private and public sector Rationale: When comparing the percentage of time used for maintenance activities in organizations earlier, we have found this to be stable on around 40 percent of the overall time in investigations both in the seventies, eighties, and nineties in both USA and Norway. It is interesting to see if this is different between private and public sector.

H4: There is no difference between the distribution of work among maintenance and development between private and public sector when disregarding other work than development and maintenance. Rationale: Since the amount of other work than development and maintenance is taking up more time now than 15-20 years ago, we found it beneficial also in the surveys in 1993, 1998, 2003 and 2008 to look at the proportion between development and maintenance time only. The proportion of time used for maintenance has earlier shown to be stable on around 60% (i.e., 40% for development) in all investigations, across countries. When a larger percentage of maintenance is claimed, this often includes, e.g. user support [6].

H5: There is no difference between the distribution of application portfolio upkeep in private and public sector. Rationale: These numbers were on the same level in 2008 and 2003 as in 1998, and it interesting to see if it would be equal also across private and public sector. A high percentage on application portfolio upkeep would in particular signal poor IT support practice cf. the discussion in the introduction.

4. Results

Work on application systems was in the survey divided into the six categories presented in section 2. The same categories were also used in 1993, 1998 and 2003 and 2008. We also asked for the time used for user-support and for systems operations which took up the additional time for the work in the IT departments. For these figures we have numbers both from the main replication study and the IT i praksis- study, and we present the aggregated numbers from these studies below. Note that in the surveys, we do not ask for numbers of our specific figures on application portfolio evolution and upkeep, but calculate them from figures of the more well-established types of maintenance and development.

40.7% of the total work among the responding organizations is maintenance activities, and 16.6% is development activities. When disregarding other work than development and maintenance of application systems, the percentages are as follows: maintenance activities: 73%, development activities: 27 %. This is a bit more skewed towards maintenance than in the previous investigations, back to the level reported in 1998. 65% of development and maintenance work was application portfolio upkeep,

and 35% was application portfolio evolution. This is almost the same as in 2008, 2003 and 1998, which in turn was significantly different from the situation in 1993.

We have also investigated the breakdown of maintenance activities from our investigations where we look upon the complete portfolio of the responding organizations. Most interesting for comparison with other surveys is looking at corrective, adaptive, and perfective maintenance, which appears to be much more stable than the numbers reported from others. We do note though that the enhancive maintenance part of perfective maintenance appears to have stabilized on a lower level than we found 20 years ago.

In the light of this stability of figures, we have divided the population to test H1-H5 comparing private and public organizations. Before looking for significant relationships, the variables used in the comparisons were tested for normality. Data for all variables except application portfolio evolution cannot be investigated as if they were normally distributed. For these either the Shapiro-Wilks (S-W Sign) and/or the Kolmogorov-Smirnov (Lilliefors-Sign.) significance levels are less than 0.05. On some variables (e.g. application portfolio upkeep) we could use the assumption of normal distribution in the tests below, using t-tests, whereas for the others we use a non-parametric test (Mann-Whitney).

We tested H1-H5 by comparing the numbers from private and public sector as summarized in Table 1. We list the number of cases, the mean and the standard deviation for all relevant figures to test the eight hypotheses (for H1, there are four test, for the difference in corrective, adaptive, perfective and enhancive maintenance respectively). Δ is the absolute difference in the mean between private and public sector, and p is the probability for erroneously rejecting the equality of means. All hypotheses except H1a (corrective maintenance) were rejected.

Table 1. Test of hypothesis

	Sector	N	Mean	SD	Δ	P
Corrective maintenance, percentage of all work (vs. H1a)	Private	106	24.1	14.9	-2.5	.207
	Public	163	26.6	17.1		
Adaptive maintenance, percentage of all work (vs. H1b)	Private	106	22.5	13.6	**-2.8**	**.033**
	Public	163	25.2	11.9		
Perfective maintenance, percentage of all work (vs. H1c)	Private	106	53.5	18.8	**5.3**	**.005**
	Public	163	48.2	17.1		
Enhancive maintenance, percentage of all work (vs. H1d)	Private	106	36,5	16,9	**10.3**	**.000**
	Public	163	26,2	16.1		
Maintenance, percentage of all work (vs. H2)	Private	107	44.4	16.7	**5.9**	**.003**
	Public	167	38.5	15.6		
Development, percentage of all work (vs. H3)	Private	107	20.4	15.4	**6.2**	**.000**
	Public	167	14.2	12.8		
Maintenance, disregarding other work (vs. H4)	Private	107	69.5	18.6	**-6.2**	**.012**
	Public	164	75.7	19.0		
Application portfolio upkeep (vs. H5)	Private	107	60.7	14.9	**-7.5**	**.000**
	Public	164	68.2	16.7		

5. Discussion and Conclusion

Overall percentage of time used for evolving systems in production compared to time used for development is remarkably stable over time. In earlier investigations the

differences found between private and public sector was not significant [12]. On the other hand, as we see in the current investigation with data from around 275 organizations (compared to around 70 responses in earlier investigations) we find significant differences between public and private organizations, generally pointing to that private organizations are able to use more of their time on value adding IT-activities. In addition to this difference the efficiency of the time used for development and maintenance tasks are not captured in these investigations, i.e. the amount of new functionality provided through the development of new systems or enhancive maintenance, cost overruns etc.

Several of our results have spurred new areas that could be interesting to follow up on in further investigations, and we have in addition to the survey performed several detailed case studies in different public sector IT-departments.

On the short-term, we will collaborate with Rambøll on 'IT i praksis' also in 2015 to get additional data points being able to confirm or refute the pattern found in this investigation. A long-term plan is to do a similar investigation in 2018 following up the 5 year cycle of investigations.

References

[1]   Heeks, R.: Implementing and Managing eGovernment, Sage (2006)
[2]   Moløkken-Østvold, K.: Effort and schedule estimation of software development projects. , PhD thesis, University of Oslo, Norway (2004)
[3]   Flak, L.S., Dertz, W., Jansen, A., Krogstie, J., Ølnes, S., Spjelkavik, I.: What is the Value of eGovernment - and How can we actually realize it. Transforming Government: People, Process and Policy 3 (3):220-226 (2009)
[4]   PASC: Government and IT – "a recipe for rip-offs": time for a new approach", Volume I, 28 July (2011) http://www.publications.parliament.uk/pa/cm201012/cmselect/ cmpubadm /715/715i.pdf (Accessed March 15, 2012).
[5]   Budzier, A.,Flyvbjerg, B.: Double Whammy – How ICT projects are Fooled by Randomness and Screwed by Political Intent. Said Business School Working paper available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2238057 (2011) (Accessed March 15, 2015)
[6]   Jones, C.: The Economics of Software Maintenance in the Twenty First Century http://www.compaid.com/caiinternet/ezine/capersjones-maintenance.pdf Last accessed March 2012 (2006)
[7]   Boehm, B., Sullivan, K.: Software economics: status and prospect. Information and Software Technology 41(14) 937-946 (1999)
[8]   Brooks, F. P.: No silver bullet. IEEE Computer 20 (4) 10-19 (1987)
[9]   Davidsen, M.K.: A Study of Development and Maintenance in Norway. Master Thesis, NTNU, Trondheim, Norway (2009)
[10]   Veld, T.K, Krogstie, J.: Work Distribution in Information Systems Development, Maintenance and Evolution: A Replication Study Proceedings NOKOBIT (2014)
[11]   Følstad, A., Jørgensen, H., Krogstie, J.: User Involvement in e-Government Development Projects. Paper presented at the NordiChi'2004 Tampere, Finland (2004)
[12]   Krogstie, J. Comparing Private and Public Sector on Information Systems Development and Maintenance Efficiency. In proceeding of eGov 2012 Lecture Notes in Computer Science, 7443, pp. 222-233 (2012)
[13]   IEEE Standard Glossary of Software Engineering Terminology (1991)

[14] Swanson, E.B.: The dimensions of maintenance Proceedings of the Second International Conference on Software Engineering, San Francisco, USA, August, pp. 492-497 (1976)

[15] Chapin, N.: Software Maintenance Types – A Fresh View. Proceedings of the International Conference on Software Maintenance (ICSM'2000) pp 247-252 (2000)

[16] Krogstie, J.: On the distinction between functional development and functional maintenance, Journal of Software Maintenance, 7 pp. 383-403 (1995)

[17] Chapin, N., Hale, J., Khan, K., Ramil, J. Tan, W-T.: Types of Software Evolution and Software Maintenance Journal of Software Maintenance Vol 13. 3-30 (2001)

[18] Dekleva, S. M.: Software Maintenance: 1990 Status. Journal of Software Maintenance 4 pp 233-247 (1992)

[19] Rambøll: IT i praksis 2014 Rambøll Management Consulting (2014)

[20] Davidsen, M. K., Krogstie, J.: Information Systems Evolution over the Last 15 Years. In proccedings CAiSE-2010 Conference on Advanced Information Systems Engineering, Hammamet, Tunis (2010)

[21] Holgeid, K.K., Krogstie, J. and Sjøberg, D. I. K.: A study of development and maintenance in Norway: Assessing the efficiency of information systems support using functional maintenance. Information and Software Technology. 42: p. 687-700. (2000)

[22] Krogstie, J., Sølvberg, A.: Software maintenance in Norway: A survey investigation, in: H. Muller and M. Georges (Eds.) Proceedings of the International Conference on Software Maintenance (ICSM'94), IEEE Computer Society Press, Victoria, Canada, pp. 304-313. (1994)

[23] Krogstie, J. Work distribution, Methodology, and Technology for ISD and Maintenance: Trends over the last 30 years. In proceedings of the IFIP WCC TC8, Santiago, Chile. (2006)

[24] Lientz, B.P., Swanson, E.B.: Software Maintenance Management, Addison Wesley (1980)

[25] Nosek, J.T., Palvia, P.: Software maintenance management: Changes in the last decade, Journal of Software Maintenance, Vol. 2 157-174 (1990)

[26] Swanson, E.B., Beath, C. M.: Maintaining Information Systems in Organizations, Wiley Series in Information Systems, John Wiley & Sons (1989)

[27] Gupta, A., Slyngstad, O. P., Conradi, R., Mohagheghi, P., Rønneberg, H., Landre, E.: An Empirical Study of Software Changes in Statoil ASA - Origin, Priority Level and Relation to Component Size. In Proceedings of ICSEA (October 29 - November 03). IEEE Computer Society, Washington, DC, 12 (2006)

[28] Lee, M.-G., Jefferson, T.L.: An Empirical Study of Software Maintenance of a Web-Based Java Application, Proceedings ICSM'05 (2005)

[29] Schach, S. R., Jin, B., Yu, L., Heller, G. Z., Offutt, J.: Determining the Distribution of Maintenance Categories: Survey versus Measurement. Empirical Software. Engineering Vol. 8, 4, 351-365 (2003)

[30] Benestad, H.C., Anda, B. C. D., Arisholm, E.: Understanding software mainten-ance and evolution by analyzing individual changes: A literature review, Journal of Software Maintenance and Evolution: Research and Practice (2009)

[31] Gupta, A.: The Profile of Software Changes in Reused vs. Non-reused Industrial Software Systems Doctoral Thesis 2009:90 NTNU (2009)

[32] Rajlich, V.T., Bennett, K.H.: A staged Model for the Software Life Cycle, IEEE Computer, Vol. 33 66-71 (2000)

Metadata management requirements and existing solutions in EU Institutions and Member States

Makx DEKKERS[1], Nikolaos LOUTAS[2], Stijn GOEDERTIER[2],
Athanasios KARALOPOULOS[3], Vassilios PERISTERAS[3], Suzanne WIGARD[3]
[1]*AMI Consult, Barcelona, Spain*
makx@makxdekkers.com
[2]*PwC EU Services, Brussels, Belgium*
{firstname.lastname}@be.pwc.com
[3]*Directorate-General Informatics, European Commission, Brussels, Belgium*
{firstname.lastname}@ec.europa.eu

Abstract. EU Member States and Institutions deal in daily basis with a considerably big amount of data. It is of high priority for them to focus on applying different solutions for better managing and governing their metadata to achieve better handling of their data. We conducted an in-depth analysis of selected cases via structured interviews. The input was used to identify a number of observations of good practices put forward by the selected case studies and a number of recommendations.

Keywords. Structural metadata, management, governance, data standard, ADMS, reference data, data model

1. Introduction

This paper presents the findings of the study on metadata management requirements, governance and existing solutions in EU Institutions and Member States which was commissioned by the Interoperability Solutions for European Public Administrations (ISA)[1] Programme of the European Commission, in the context of its Action 1.1 on improving semantic interoperability in European e-Government systems[2].

The prime objective of this work was to document what is available in the Member States and the EU Institutions with regards to policies and practices related to the management of structural metadata, i.e. data models and reference data.

Tailoring TOGAF's definition[3] of governance, we define structural metadata governance as comprising well-defined roles and responsibilities, cohesive policies and principles, and decision-making processes that define, govern and regulate the lifecycle of metadata. We define metadata management as the good practice of adopting policies, processes, and systems to plan, perform, evaluate, and improve the use and reuse of data models and reference data.

---

1 http://ec.europa.eu/isa/
2 http://ec.europa.eu/isa/actions/01-trusted-information-exchange/1-1action_en.htm
3 http://www.opengroup.org/togaf/

2. Methodology

In the context of this work, we performed an in-depth analysis of selected cases conducted via structured interviews following a detailed analysis framework.

2.1. Selection and analysis of the case studies

The cases for the detailed analysis were selected upon predefined criteria such as inter-organisational focus, regional diversity, information accessibility, and maturity of the solution. The following cases were selected:

- EU - The Statistical Office of the European Union - Eurostat, which process and publish comparable statistical information at European level.
- EU - Joint Research Centre - INSPIRE, which enables the sharing of environmental spatial information among public sector organisations.
- EU - Inter-institutional Metadata Management Committee (IMMC) [4], Publications Office Metadata Registry (MDR), which registers and maintains metadata elements, named authority lists, schemas, etc.
- DE - Koordinierungsstelle für IT-Standards (KoSIT), which coordinates the development and operation of IT standards for data exchange in the public administration in Germany.
- ES - Centre for Semantic Interoperability (CISE), which carries out projects related to information and communication technologies (ICTs) in accordance with the strategic priorities in Spain.
- LT - Lithuanian Spatial Information Portal (LSIP)), which makes conditions for the provision of centralized spatial data.
- NL - Knowledge and Exploitation Centre Official Government Publications (KOOP), which manages the metadata standard for information published by the Dutch government.
- UK - Local Government Inform (LG Inform/LG Inform Plus), which provides a benchmarking data service of Local Government Association for councils in the UK.

2.2. Case study analysis framework

2.2.1. Structural metadata governance

The structural metadata governance dimension of each case was analysed as follows:

- **Goals:** What are the long-term goals of cross-border/sector governance, management and reuse of common structural metadata?
- **Governance structure:** What are the roles responsible for structural metadata governance and what is their mandate?
- **Policy domain**: Is structural metadata governance restricted to specific policy domains? If so, then in which policy domains?
- **Enforcement policy:** Where relevant, the applicable legal context must be completed.
- Sharing: which governance model is used to encourage/ensure sharing of structural metadata?

- Reuse: which governance model is used to encourage/ensure reuse of structural metadata?
- **Authoritative source:** Is there an authoritative source, e.g. a repository or a file server, on which the structural metadata is housed?
- **Licensing framework:** Under which licensing framework are structural metadata shared and reused?
- **Quality controls:** What is the quality management process for structural metadata?
- **Metadata Schema:**
- Vocabulary: Is there a common vocabulary in place, such as the Asset Description Metadata Schema - ADMS[4], for documenting metadata?
- Identifiers scheme: Are there common guidelines and design patterns for creating identifiers for the metadata?
- Schema documentation: Is there schema documentation in place, which explains commonly agreed definitions about the meaning of the data?
- Multilingualism: Are different languages supported for the metadata properties and/or values?

2.2.2. Structural metadata management

The metadata management dimension of each case study is analysed according to the following aspects:

- **Design and development:** which entails the processes of agreeing on the syntax and the semantics, and encoding the structural metadata in different formats.
- **Update:** which entails the processes of updating the structural metadata, and deleting/deprecating outdated versions.
- Update frequency: What is the average update frequency of structural metadata?
- Change management process: Is there a defined procedure in place?
- Version control: Is there a version control system in place?
- **Harmonisation:** which covers the processes that have to be put in place for:
- Creating mappings between related structural metadata sets.
- Assessing alternative structural metadata sets.
- **Documentation:** which covers the processes that facilitate the sharing and reuse of structural metadata:
- Publication of structural metadata documented according to a common vocabulary on an authoritative source which is accessible and supports search capabilities for both humans and machines.
- Retrieval of structural metadata either by humans (e.g. downloading a file) or by machines (e.g. by consuming the metadata via an API).
- Supported formats of publication tool(s): Human-readable only (H); Machine-readable only (M); Human and Machine readable (HM).
- **Standards**: are there standards used for structural metadata management, such as ISO 11179 Metadata Registry standard [0], ISO 25964 Thesauri and interoperability [[2]], the Simple Knowledge Organisation System – SKOS [[3]] and ADMS.

---

[4] https://joinup.ec.europa.eu/asset/adms/description

- **Costs:** the number of FTEs occupied with structural metadata governance and management in particular organisation.
- **Benefits:** The key benefits of structural metadata governance and management in a particular case. We also checked whether the benefits offset the costs.

3. Analysis of the findings

3.1. Structural metadata governance

The most common goals of structural metadata governance and management are:

- To achieve and ensure interoperability between existing applications and systems (e.g. KoSIT, Eurostat, INSPIRE, IMMC/MDR, etc.);
- To provide an authoritative source from where structural metadata could be made available, stored and managed, e.g. see the cases of Eurostat and LGI).

The most common challenges faced are the following:
- The goals and the importance of metadata management are often not clear to the management of the organisation and there is a lack of support from the leader-ship.
- It is important to keep track of new extensions meaning i.e. well thought version-ing process and manage links of the common models i.e. that updates or new ex-tensions would be properly aligned with the existing model. For example INSPIRE and KOOP apply this in practice for their common models.
- The alignment of structural metadata managed locally and structural metadata managed in the central authority. There is an intention to manage structural metadata while at the same time managing links to local extensions and local models e.g. the Netherlands KOOP case.
We observed a generic four level governance structure.

- Legal level establishing the legal environment, e.g. involving parliament passing laws as with INSPIRE and CISE (not always there);
- Strategic level on which an Authority sets the high-level goals (often Ministries or bodies like IMSC);
- Functional level where decisions are taken on (changes to) models and vocabularies;, such as OWMS User Group, IMMC);
- Operational level where the agreed changes are implemented such as MRT, JRC, KOOP, CISE etc.

We saw two ways of stakeholder involvement, either the person/organisation that wants to participate needs to ask for an invitation/permission by the authority responsible for metadata governance, or the governance body invites the people/organisations that they believe that should/have a stake in participating.

The decision making process can be distributed through three different inter-organisational levels with certain responsibilities:

- Steering body or so called 'authority' that decides on vision and strategy.
- Organisational tasks are assigned to the institution reporting to the Steering

body;
- Collection of change requests and implementation of decisions is usually assigned to communities and working groups (i.e. OWMS community in case of KOOP).

The enforcement policy depends on the domain and the country. The way it is implemented varies depending on the purpose and the context of use of the metadata (e.g. provision of statistics, geographic information).

There is trade-off between stability and flexibility when it comes to choosing the enforcement policy for structural metadata governance. Having enforcement on the basis of legal requirement caters for stability since the use of the structural metadata is obligatory and changes can only be made by amending the legal environment. On the other hand, not having strict legal requirements gives more flexibility and provides conditions for faster adaptation to changing requirements to the upcoming needs.

There is an authoritative source for structural metadata in all eight case studies. In most cases the authoritative source is a tool (registry) that is accessible online where structural metadata is publicly available.

In all case studies licensing is not clearly identified on authoritative sources of information. To summarise, we can state that EU Institutions seem to be more aware of licensing than the Member States.

The quality control process is usually manual and ad-hoc (e.g. KOOP, CISE). In-house developed scripts or tools are being used (e.g. INSPIRE validator) and in most cases the process itself is not clearly defined.

A set of descriptors are used for the structural metadata in each case study reviewed. In almost all cases ADMS is either being considered internally for the metadata descriptions (e.g. KOOP), or ADMS exports are available, e.g. in the case of the Publications Office and Eurostat. All of the cases have documentation of their metadata in place. INSPIRE has guidelines and xml dictionary, KoSIT generates schema documentation directly from UML model.

Documentation of all pan-European initiatives is grounded on an authoritative language i.e. English, and labels are translated into different languages   European initiatives do support other languages (in terms of tools functionality).

3.2. Structural metadata management

The analysis shows that metadata management process is more standardised in EU Institutions than in organisations of Member States. We recommend the use of metadata management standards that give advantages of cost effectiveness and reusability.

Most of the organisations studied, i.e. Eurostat, IMMC, LSIP and LG Inform, have established processes for mappings between related metadata sets and for assessing alternatives, yet they are commonly used only for reference data and not for other sources such as vocabularies. Our analysis suggests that XSD[5] could be used for the publication of metadata schemas and the transmission of protocol schemas, while SKOS[6] files could be used to manage value and code lists.

The findings of the study show that there are usually few FTEs responsible for metadata management and it is emphasised that benefits notably offset the costs.

---

[5]   http://www.w3.org/TR/xmlschema11-1/
[6]   http://www.w3.org/2004/02/skos/

In all of the reviewed cases human- and machine-readable formats are available for retrieval, yet human- and machine-readable distributions of the structural metadata are published separately. The possibility to reuse the tool depends not only on its functionality, but also on the type of licence and owner/ vendor. Commercial and open-source products are easily reusable, while in-house solutions require efforts for adaptation.

In terms of tools used for metadata management and governance there are three groups to mention:

- General purpose commercial tools that can be easily reused (given that the organisation can afford them);
- General purpose open source tools developed by third parties that can be easily reused;
- Open source built in-house tools that would require adoption and configuration to cater for new needs and requirements.

3.3. Identified good practices

We observed a number of good practices put forward by the selected case studies:

- In a good governance structure, the roles concerning legislation, strategy, functionality and operations are clearly distinguished and assigned to designated bodies.
- The involvement of direct stakeholders in the metadata governance process ensures that the interests of the stakeholders are taken into account which maximises buy-in and take-up.
- Voluntary sharing and re-use works best if stakeholders are aware of the advantages of collaboration and of the benefits for interoperability.
- Application of a standard for metadata management creates a well-structured management environment based on existing good practice. We defined two families of standards:
  - Standards for metadata management, such as ISO/IEC 11179 and ISO 19135.
  - Standards for documentation and representation, such as ADMS and SKOS.
- Good change management processes are based in stability where possible without sacrificing flexibility where needed and take into account an alignment between the life cycles of structural metadata development and software development.
- Changes in structural metadata are well planned and tracked, preserving backward compatibility as much as possible; in cases where disruptive changes are unavoidable, these changes should be planned and communicated well in advance.
- Structural metadata is managed in formats that are appropriate for the type of use. Metadata describing the structural metadata is expressed or exported using ADMS.
- Standard reference data is used wherever appropriate; if locally defined reference data is used, this is mapped to standard reference data to enable interoperability.

- Structural metadata is distributed in machine-readable formats that can be processed by the tools available by the re-users.
- Content negotiation is used to manage and provide different types of formats from the same URI.

3.4. Recommendations

In this section, we form recommendations in order to better organise day-to-day operations related to structural metadata governance and management:

- Legislation should be formulated on a sufficiently high level and should not specify details like the values in a code list or the elements of a data model; these details should be specified as part of the implementation and made available from an authoritative source to which the legislation can refer. In case of enforcement by law, it is recommended to mention a link in the law referring to the authoritative source of structural metadata, rather than citing the structural metadata itself in the law.
- The structural metadata management processes should be documented. Having stable and transparent processes is expected to improve the efficiency of the process itself, and to set a common ground for operating metadata management, taking decisions, and resolving conflicts. In order to foster reuse, the documentation of both the management process and the structural metadata itself should be made available in different languages.
- Owners of structural metadata should be made aware of the importance of clear licensing arrangements that specify unambiguously under which conditions the metadata can be reused. A clear license is considered as an enabler for sharing and reusing structural metadata. If no licence is in place, users will hesitate to use the metadata, since it is effectively not clear to them if they can do so and under which conditions. Open licences for structural are recommended with protection against misrepresentation. Our findings show that, especially in the case studies of the EU Institutions, EUPL and the EU legal notice are usually applied, but it is not always explicitly publicly visible.
- In order to gain management support and encourage the sharing and reuse of metadata, the direct stakeholders should be made aware of the expected benefits. Sharing and reuse of metadata contributes to the realisation of the expected benefits, particularly the ones related to interoperability. Especially in cases when enforcement is not mandatory, stakeholders should be made aware of the advantages of sharing or reusing of metadata, and should be persuaded to do so.
- Management processes and publication frequencies should be different for changes to data models on one hand, and reference data on the other hand. Changes in structural metadata should be justified and well-planned. A change request should be issued only when the metadata cannot support an application scenario, because of a new requirement. It is necessary to make a distinction among processes and publication frequencies for changes in data models, which are usually more stable and changes there may also impact the operation of production systems. The structural metadata lifecycle – and its change management process - should be aligned with the software development lifecycle.
- Structural metadata should have persistent unique identifiers. It is

recommended that such identifiers come in the form of HTTP URIs. The ISA Programme as well as W3C have created good practices and guidelines for the design and management of well-formed, persistent URIs, e.g. see ISA's 10 Rules for Persistent URIs [5].

The toolset that will be deployed by an organisation to support metadata management and governance should, as a minimum, cater for the following functionalities:

- Provision of an authoritative source for the structural metadata, which allows for storing the metadata and its documentation, and offers search, visualisation and browsing functionalities;
- Issue tracking and ticket management to manage metadata changes;
- Version control to support change and release management;
- Maintenance and linking to support metadata harmonisation.

4. Conclusion

The activities around metadata governance and management appear to be in an early phase: in most cases investigated, governance structures have been established relatively recently, and some of the management processes and procedures are still in the process of being defined and implemented. In many of the cases considered, the organisations responsible for the implementation of these processes and procedures expressed a vivid interest in sharing their experiences and learning from others.

It is the intention of this study that the identified good practices can serve as guiding principles by organisations that are in the process of defining structures, procedures and workflows for governance and management.

In a more general context, given the importance that is attributed by public administrations around Europe and beyond on the governance and management of structural metadata (as an enabler to interoperable information exchange and cross-border/-sector public services), we expect that this topic will rank higher on the policy agendas of different countries. In this vein, works like this one, will play a significant role in ensuring that public administrations can learn and benefit from international good practices and recommendations, avoid re-inventing the wheel, and receive guidance in reusing metadata governance and management practices and tools in their own contexts.

References

[1] ISO (International Organization for Standardization). "ISO 11179-3:2013-- Information technology -- Metadata registries (MDR) -- Part 3: Registry metamodel and basic attributes"
[2] ISO (International Organization for Standardization), "ISO 25964-2:2013-- Information and documentation -- Thesauri and interoperability with other vocabularies -- Part 2: Interoperability with other vocabularies"
[3] Bechhofer, S., & Miles, A. (2009). SKOS Simple Knowledge Organization System Reference. W3C.
[4] Publications Office of the European Union. (2011). Proposal for metadata governance on interinstitutional level.
[5] ISA Programme of the European Commission. (2012), Study on persistent URIs, with identification of best practices and recommendations on the topic for the MSs and the EC.

Innovation in eGovernment Information Infrastructures

Arild Jansen

Norwegian Center for Computers and Law, Arildj@jus.uio.no

Abstract. Although many governments are developing digital Government Information Infrastructures (eGovII), it seems to be little understanding of eGovII governance approaches other than those applied for traditional information systems. There is an increasing focus on how such eGovII may support innovations, both in developing public services, and when private enterprises uses public data in new ways. The paper aims at exploring what types of governance approaches for eGovII that may be most appropriate for supporting innovations. Our empirical base is the Norwegian eGovII including 3 common components as the Norwegian Population Register, the business reporting portal AltInn and the Land Registry and Cadastre.

Keywords Information infrastructure, innovation, e-Government, IT Governance.

1. Introduction

The Norwegian government, like governments in many other countries, is facing great challenges in their efforts to overcome the obstacles created by the highly fragmented public sector, resulting in a silo-organization of its information systems. As a response, they aim at implementing a more coherent Government information infrastructure (eGovII), including an ICT architecture with core functions and services that can enable information exchange and interaction between public agencies. However, such efforts imply technical, as well as organizational and not least legal challenges. We will argue that a major challenge is to build an open and flexible eGovII that also can stimulate innovations, as innovation is a key priority both at a national as at European. Innovations are essential for provide better services for its users and stimulate growth in private sector in general.

However, technological and organizational innovations in public sector have to be supported by changes in the legal system. Accordingly, we have to cope with another challenges associated with their long-established structures of governing that may be hard to change. These are traditionally based on the Webern governance model which involves hierarchy, authority, command and control, and uniformity are primarily based on meeting the requirements of equality, openness, accountancy etc., resembling a rational, centralized governance form. Infrastructures can be characterized by its technical complexity, their many actors involved, they span more organizations and they need to support diverging users' perspective [1]. Their developments involve stakeholders outside the hierarchical control of any single entity. The room for finding a common ground may be constrained, not least due to path-dependence linked to the installed base. Thus, II governance in general and eGovII governance in particular seems to present a rather different type of complexity which is contrasting the management of traditional structures. This paper is motivated by the need to understand

the missing coherence between the characteristics of eGovII and the governance models traditionally being used, by asking this overall question:

What governance approaches for eGovII can be best suited to stimulate innovation?

To assist our analysis, we have developed three governance models; i) a rational, centralized model, ii) a decentralized model, and iii) a network model. Our empirical base is the Norwegian eGovII including common components as the Norwegian Population Register, the business reporting portal AltInn and the Land Registry and Cadastre. We believe that a comparative analysis of such infrastructure elements can provide insight into eGovII governance.

The structure of the paper is as follows: section 2 outline the theoretical basis: the conceptualization of eGovII and innovation theory. Furthermore, three different governance models are developed. In section 3, we present our empirical cases, followed by an analysis of the governance model and lastly the concluding discussions.

2. Theoretical Foundation

2.1. An Information Infrastructure perspective

We understand information infrastructure as networks of distributed yet more-or-less interlinked and interoperable information systems. Hanseth and Lyytinen[2] provide a more precise definition: *"Information Infrastructure (II) is a shared, open and unbounded, heterogeneous, and evolving socio-technical system consisting of a set of IT capabilities and their user, operations, and design communities."* As a consequence of the dispersed and distributed ownership, lack of centralized control is a fundamental attribute of information infrastructure [1]. Different actors shape, maintain, and extend information infrastructure "in modular increments, not all at once or globally" [3].

The importance of eGovIIs is more visible as its increasing number of central registers and other core components are shared in order to provide online government services. eGovIIs must be evolving, flexible and scalable in order to meet new requirements. Of particular importance is the *installed base* including non-technical elements as organizational, economic and legal factors, illustrating their network effects determining the further development of the infrastructures [2]. The evolution of IIs in general and eGovII in particular is path-dependent due this "living legacy" of technical solutions, interconnected practices and regulations that are institutionalized in the organization [4]. The legal framework is an essential part of the installed of an eGovII as it governs the execution of public authority. New laws and regulations are enacted without considering interoperability issues, implying that the installed base is neither static nor controlled by a single authority, including local ICT-solutions in use at various levels of the public sector.

2.2. Information Infrastructures and innovations

The understanding of the installed base of an II is essential for its governance, in order to handle the existing collection of nonstandard legacy systems. However, these legacy systems also create barriers for innovations [5, 6] argues that the governance of II requires another approaches than traditional top down IS management styles, based on mechanisms as cultivation and drifting. Cultivation can be understood as nurturing "natural" bottom-up initiated processer, which is necessary in order to stimulate innovations, e.g. to create new services including changes in socio-technical practices. Tuomi [7] argues that "innovation emerges in a complex iterative process where

communication, learning and social interaction play important roles". Furthermore, he argues that "the adoption of new innovation requires learning and developing skills".

Innovation in technology use is often linked to generativity. According to Zittrain [8], generativity denotes a technology's overall capacity to produce unprompted changes driven by large, varied, and uncoordinated audiences. Zittrain claims that generativity is useful in analyzing technologies on the basis of its five criteria: (1) how extensively a system or technology leverages a set of possible tasks; (2) how well it can be adapted to a range of tasks; (3) how easily new contributors can master it; (4) how accessible it is to those ready and able to build on it; and (5) how transferable any changes are to others, including non-experts. Andersen and Jansen [9] shows how innovations in IIs can take place at an *infrastructure*, at an *application* or an *organization* level.

2.3. eGovII and IT-Governance

IT governance refers to the 'patterns of authority for key IT activities" [10 p. 261]. In analogy, eGovII governance can be viewed as the authority patterns to direct eGovII development and use. Sambamurthy & Zmud [11] identify three primary modes of IT governance; centralized, decentralized and federal mode. Building eGovII is different from traditional software development project, as eGovII encompasses multiple organizations, they changes continuality and the direction of the development is uncertain. We thus argue that governing an infrastructure require other approaches than traditional IS management styles. To guide our analysis, we have constructed three models; a rational, a decentralized and a network (distributed) model that aim at capturing the basic and essential variations in governance approaches related to decision-making, alignment and communication [12].

The *rational* model promotes the view that the three mechanisms of IT-governance; decision-making, alignment and communication must be seen as well planned, top-down processes [13]. Accordingly, the decision-making structures are hierarchically organized in that the most important decisions are made by a few top executives, the implementation of decisions is managed from one central point of control, and communication follows the hierarchical pattern of authority.

The *decentralized* model assumes that public bureaucracies have long-lasting institutional features, which may have evolved differently in distinct organizations. Decision-making, alignment processes and communication are therefore seen as relatively stable manifestations of the norms, values and practices that gives the bureaucracy its distinctive mode of operation [13]. This implies that IT-governance is not done in one particular way (as with other models).

Table I. Major features of three governance models (partly based on [12]

Governance models	Central governance (The Rational Model)	Decentralized IS governance (Decentralized Model)	Distributed governance (The Network Model)
Management focus & type	Rational, top-down. Management by objectives.	Maintain traditional institutional form.	Reduce hierarchy, stimulate cooperation. Quality mnt.
Decision-Making	Hierarchical (decisions made at the top)	Mostly decentralized	Participatory (decisions involve lower levels)
No_of. stakeholders	Few	Varying , few to many	Many, professional oriented, specialist.

The *network* model assumes that decision-making and alignment processes are achieved through participation, consultation and involvement [12]. Communication

follows the same horizontal pattern of interaction as the collaboration of internal and external constituencies are sought. This means that IT-governance is not so much about designing "grand plans", but by adopting a "softer touch". Also, traditions are replaced by a much greater emphasis on change and flexibility. Table 1 lists their main features.

2.4. Research approach

Our research theme is "*What governance approach for eGovII can be best suited to stimulate innovation?*" To help our analysis, we will address the following questions:

i) What governance model(s) is most adequate for eGovII components?

ii) What opportunities for innovations can be identified in different eGovII components and their governance?

This research is based on an exploratory approach, as the aim is to identify factors in eGovIIs that seem more relevant for innovation than others, and not to test specific hypothesis. A case study based research methodology is particularly well-suited to IS research, since the object of the discipline is the study of IS in organizations, and the '*interest is shifted to organizational rather than technical issues'* [14]. Our three cases are; the Norwegian Population Register *(NPR),* the business reporting portal *Altinn*, and the Land Registry and *Cadastre,* which all are basic elements of the Norwegian eGovernment ICT architecture.

The empirical data was collected from the ministries budget documents and relevant white papers and government reports, and from the corresponding assignment letters of selected subordinate agencies. Furthermore, we have interviewed key civil servants. The assessment of the different agencies governance approaches, i.e. which model of governance that can be said to characterize the cases, is based on a) an evaluation of the policy objectives, means and instruments that are specified in the assignment letters and (b) of how our informants (in interviews) describe the way they control and supervise the performance of the infrastructures. The evaluation of the innovation potential of the three cases is based on analysis of how various public agencies and private businesses are using services and data from these components, furthermore to what extent there have been created new services by third parties.

3. The Norwegian ICT architecture

3.1. ICT governance policy in Norwegian public sector

Norway is a rather sector-oriented and decentralized, but unitary state where the municipalities have autonomy within the national legal framework. One implication is that Norwegian reform processes might be more segmented and sector-oriented than in other countries [15]. One Minister coordinates public sector reform, however without overruling the other ministries.

Overall governance structure: The historical- institutional "climate" in Norway is characterized by a strong statist tradition, homogeneity in norms, mutual trust between political and administrative leaders, incremental changes and the balancing of many considerations, which indicates that reforms will be implemented slowly [15]. During 1990, the principles of Management by Objectives and Return (MBOR) were gradually implemented. Today, the "formal" Norwegian governance approach is based on important elements of the rational model; top-down decision structure, mainly vertical communication patterns, centrally controlled project management principles, etc. However, a recent study of IT governance approaches in the individual ministries in

Norway has revealed a far more diversified picture, indicating that management practices corresponding with all the three models outlined above can be identified [16]. At a more detailed level, it was found that the Ministry of Finance, being responsible for the Norwegian Population Register (NPR), has adopted a rational approach. The Ministry of Trade and Industry, having the primary responsibility for the business Altinn, however practices a somewhat more decentralized, partly network-oriented IT governance approach. The Ministry of Environment, being responsible for Cadastre, has also adopted a mix of decentralized and network-oriented approach [17].

ICT-architecture: In 2009, ICT architecture for the state government was established, including a set of central registers and core functional components, among the Altinn platform, and 3 main registers, including NPR and Cadastre [18]. The guiding principle for this framework is "comply or explain", implying that new systems has to accommodate the architecture unless strong grounds justify another alternatives.

3.2. National Population Register (NPR)

NPR includes information about everyone living in Norway. Such information is gathered by local tax offices and entered into the NPR, and is overseen by the Directorate of Taxes. Information from the NPR, e.g. names, addresses, citizenship, identification numbers etc. is only accessible to authorized public sector offices. However, private stakeholders may apply for access to information from the NPR for legal purposes. NPR is owned and operated by the Tax directorate, but are now being used by a number of state agencies and all municipalities, which are heavily dependent on such data in their daily operations, not least in the provision of egovernment services. However, neither the information quality nor the business model has proven adequate for serving an increasing number of users and usage patterns created by e-government services [19]. The register does not include all relevant data about citizens, and its on-line accessibility is difficult and costly, implying that we see few new applications utilizing its data. Most municipalities have established local copies of NPR, disintegrating the quality of the master NPR.

3.3. Altinn

Altinn is a portal for public reporting, and it is an integral part of the ICT-architecture. More than 120 electronic forms and services from a large number of Norwegian government agencies are currently available through Altinn. The responsibilities for these forms reside in the individual agencies that manage the corresponding regulations. Since Altinn was launched in December 2003, more than 23 million electronic forms have been submitted through the system. Altinn is operated by The Brønnøysund Register Centre, which is a government body under the Norwegian Ministry of Trade and Industry, and also operates a register of all Business Enterprises in Norway. A key element is the *metadata register,* which shall support interoperability by assisting the exchange of data between different systems. Altinn has been developed through a bottom up approach, originally initiated through cooperation between three agencies, but has gradually been extended to support the reporting of a large variety of data. Altinn has now been expanded into a more general portal, and supports a large part of the communication between the public sector, businesses and also citizens. It has supported the development of a number of applications throughout the public sector, which illustrates its potential for supporting innovations[20],[21]. But there are still accessibility challenges related to nonstandard, proprietary solutions.

3.4. Land Registry and Cadastre

Norwegian properties are registered in the land register, which is the official register of legal rights and obligations associated with fixed property and housing cooperatives. The land register lists ownership and encumbrances such as mortgages, leasing rights, and furthermore borders, areas, buildings and addresses etc. The Norwegian Mapping Authority is the judicial registration authority. While the earlier version was organized in a rather centralistic way, the current model has a distributed governance structure, where each municipality is responsible for updating own data. The data quality is considered to be in general satisfactory. The register acts now as a hub for a number of public and private registers, and supports the development of new services. Accessibility to the infrastructure layer is adequate, building on international standards, while higher level may include non-stand interfaces.[22]

We may summarize the characteristics of our three cases in table II:

Table II Main characteristics of three ICT architecture components

	NPR:	**Altinn:**	**Cadastre:**
General description	Centralized governance. Many users, but limited influence.	Partly decentralized, many stakeholders, elements of a network approach.	Combining central overall management and distributed maintenance (network oriented)
Service level	Inflexible, poor data quality, inadequate accessibility	In general adequate functionality. some accessibility problems	Satisfactory service level and data quality. Local (municipal) responsibility quality assurance.
Dev strategy/ funding	Top-down strategy Ordinary budget.	Overall bottom-up, experimental initiative. Specific projects funds	Mix of bottom-up and top-down Partly ordinary budget and partly specific project fund
Access/ payment model.	Access through a private operator transaction cost	Transactions-based cost sharing by service providers (users).	Free access for public sector. Others: cost-based access.
Innovation supported	Very few – if any	An increasing - in various areas	Many, open access to basic data through standardized interface

4. Analysis and Discussion

We see that the NPR is centralized, being managed by one major stakeholder and resembling a rational governance model. It was mainly built for collecting taxes along with statistical purposes, and not for being part of a government-wide infrastructure across the government. Accordingly, it does not meet the needs from neither state agencies nor the municipalities in adequate ways.

Contrary, the governance of Altinn is (partly) decentralized and coordinated by its many stakeholders, including several ministries, indicating that also a network oriented governance approach is taken. A new version is being implemented which will gradually replace the current version, where there have been some challenges, both related to implementation and governance, as e.g. alignment with processes in user agencies. We find that Altinn management practice to some extent corresponds to typical characteristics of information infrastructure governance strategies; partly bottom-up and evolutionary development approach and being evolving and more flexible. Similarly, the Cadastre has both elements of a decentralized and a network-oriented governance structure, in that all municipalities are responsible for updating their individual part of the register, but in a coordinated manner. There has been established a network of practitioner to coordinate running operation and maintenance.

The Cadastre is being used by a large number of eGovernment systems and has supported the creation of a number of innovative services, though mainly at infrastructure level. The accessibility to both Altinn and Cadastre is perceived as mainly adequate, although there are obstacles related to proprietary solutions and non-standard interfaces. We may summarize our findings in table III:

Table III Governance models in the 3 ICT components

	NPR	**Altinn**	**Cadastre**
Governance type and structure	Well-defined authority rational, bureaucratic governance by MBOR.	Mix of hierarchical and decentralized governance Stimulate cooperation.	Mostly networked, but also rational; Stimulate cooperation and interaction.
Decision making / communication	Hierarchical and centralized structure.	Multi-stakeholder, mix of vertical /networked structure.	Partly horizontal/networked decision making structure, also elements of hierarchy.
Legitimate stakeholders	Few being influential	Many stakeholders having influence	Many, dominating by professional interest

Our research question 1 is: «*What governance approaches is most adequate for eGovII components”?*

To answer this question, we must see whether the governance support openness, accessibility, flexibility etc. In addition, we have analysed whether the installed base amount to specific obstacles for developing new (innovative) application. We found that NPR is neither open, evolving nor flexible, while the two others largely conform to the characteristics of IIs. On the other hand, the installed base of NPR today is not seen as a significant problem; as few other systems are integrated with NPR, thus illustrating that NPR do not act as an infrastructure. Table IV shows how the different architecture components and their governance meet these requirements:

Table IV: characteristics of ICT architecture components

II characteristics	NPR:	Altinn:	Cadastre:
Open	No/restricted	Largely yes	Yes
Shared	No	Significant	Extensive
Evolving	No (not yet)	Yes	Yes
Flexible &modular	No	Partly	Yes, mostly
Appropriate standardization	No/inadequate	Partly, but also proprietary solutions	Partly; at the infrastructure level, not on higher levels
Barriers towards innovation			
Installed base as problem	Potentially, many local copies	Some; potentially there may be more bindings.	Some: non-standard interfaces, local copies

We see that the current centralistic way of managing NPR proves inadequate as an infrastructure governance approach, which is in line with other studies [1, 5, 6] On the contrary, both Altinn and Cadastre do to a large extent comply with these requirements and have proved more successful even though they do not solve all problems. It is however important to note that the current governance model for Cadastre is the result

of a long and rather problematic development process, in which the previous centralized management approach proved inadequate [22], [23].

Research question 2: *What opportunities for innovations can be identified in different eGovII components and their governance?*

The point of departure for our analysis is the five functional requirements as listed in section 2.2, see also [8]. We find that NPR does not meet any of these requirements sufficiently, due to its current centralized architecture and design, along with its inadequate technical solution and pricing structure. Altinn, on the other side, has a rather flexible and modular design, it has been attractive to many users (service providers in both the state sector and in the municipalities), and it is perceived at easy accessible. We also find that Cadastre meets these requirements sufficiently. In table V we summarize our findings:

Table V: Evaluation of ICT architecture components as generative technologies

Cases Function	NPR:	Altinn:	Cadastre:
Leverage tasks	Limited	Many, but not all	Yes
Adapted to a range of tasks	Only a few tasks	Many, but not for all	Mostly yes
Easily to master	Some stakeholders	Yes, in overall	Yes (lower levels)
Accessible	No	Yes	Yes

We can conclude that the current version of NPR do not easily foster innovations, which we believe, at least partly, can be attributed to their management structure along with its inadequate technical solution. On the other hand, both Altinn and Cadastre have proven adequate for supporting innovative activities. A difference between these two is that while Cadastre is partly built on accepted standards at the infrastructure level, Altinn has had to define its own standards, thus having to balance between flexibility and rigidity. The different stakeholders have quite different interests in its future development; see e.g. [20, 21]. If Altinn is being extended to support a large number of different online "forms" in non-stand ways, it will build an increasingly complex and unmanageable installed base. Thus, the challenge for Altinn is to separate the generic infrastructure components from more specific support functions. Also, improvements in Cadastre at higher level interfaces are requested. We can illustrate such mechanisms in table VI.

Table VI: ICT architecture governance and innovation mechanisms at different levels

Levels	Cases	NPR	Altinn	Cadastre
Organization	E.g. allow flexible business model	No	Mostly	Partly yes, e.g. based on open data
Application	E.g. adequate basic services	Very few	Partly	Many generic services
Infrastructure	Access to II components	Limited	Partly	Yes, standard interfaces

Our analysis do clearly illustrate that the building and operation of eGovIIs involve complex processes having many uncertainties and in which many stakeholders are involved. We see from table IV that the three cases have rather different dominating

governance structure. While the NPR so far has been managed principally in a rational, centralized way that has proved inadequate, we find that the Altinn and Cadastre management structures (partly) resembles a network governance structure that allow for innovations at various levels.

5. Conclusion

What are the best principles of governance that also foster for innovations? We found that both Altinn and Cadastre do have some elements of central control. But at the same time, there are in both cases significant decentralized or networked management mechanisms which essential for providing flexibility and local adoption. By taking a decentralized approach, it is possible to avoid directions from the central level that are constraining local development. By taking a networked approach, one can allow for cooperation and collaboration through mutual negotiations and adaptions, based on a common infrastructure that is flexible and being able to cope with local differences. A main finding is thus the need to balance centralized and decentralized or networked governance, as there is a need to have at least some elements of a rational, central governance approach to constrain options and to set directions.

We thus held that there is a priori no single governance model that can handle all parts of eGovIIs in an adequate ways. We claim that building eGovIIs should also have focus on local conditions and be adapted to the specific characteristics of each component or layer in the infrastructure. We believe that each layer in an infrastructure may be managed differently in terms of central or local control, alignment processes etc. Lower layers that include technical standards and operations need to be nearly fully standardized and more centrally governed (as we find in the case of Internet), whereas in higher layers, more diversity is clearly necessary. Thus, the lower layers of the infrastructure may have a more rational, instrumental approach (guided by enterprise architecture principles); whereas the higher layers must grow through more networked oriented management structures.

Thus, our understanding is in line with [24] conclusion that *"it is fruitful to regard information infrastructure as an ICT-based organizational form", "the lack of theories about IT artefacts, the ways in which they emerge and evolve over time, and how they become interdependent with socio-economic contexts and practices, are key unresolved issues for our field and ones that will become even more problematic in these dynamic and innovative times"*.

One (of many) remaining question is how much functionality should be located in the lower layers, and thus involving many local parties which complicates the governance structure. The innovative success of Internet is much attributed to in minimal standards and flexibility regarding its layered and modular structure. Higher layers may have to manage complexity, or as in the case of Internet; keep most of the complexity outside the infrastructure. Thus, we should investigate whether new, more flexible models for management and governance may be more suitable, where we suggest that one should design more distributed architectures including minimal kernels which are linked to a number of distributed components, being operated and managed in coordinated ways, adapted to their individual characteristics.

References

1.   Hanseth, O., C. Ciborra, and K. Braa, *The Control Devolution. ERP and the Side Effects of Globalization.* The Data base for advances in information systems. Special issue: Critical Analysis of ERP systems, 2001. **32**(4): p. 34-46.
2.   Hanseth, O. and K. Lyytinen *Design theory for dynamic complexity in information infrastructures: The case of building internet.* Journal of Information Technology., 2010 **25**(1): p. 1-19.
3.   Star, S.L. and K. Ruhleder, *Steps toward an Ecology of Infrastructure: Design and Access for LargeInformation Spaces.* Information Systems Research, 1996. **7**(1): p. 111-133.
4.   Horness, E., A. Jansen, and O. Langeland, eds. *How to Develop an Open and Flexible Information Infrastructure for the Public Sector?*   . Lecture Notes in Computer Science Springer 2010. 2010, Springer Verlag.
5.   Ciborra, C.U., *A Critical review of the literature on the Management of Corporate Information Infrastructures*, in *From Control to Drift: The Dynamics of Corporate Information Infrastructures*, C.U. Ciborra, Editor. 2000, Oxford University Press.
6.   Ciborra, C.U., *The labyrinths of information. Challenging the wisdom of systems*. 2002, Oxford University Press.
7.   Tuomi, I., *Networks of Innovation – Change and meaning in the Age of the Interne*. 2002: Oxford University Press
8.   Zittrain, J., *Th eGenerative internet.* Harward Law review 2006. **119**(May).
9.   Andersen, S.T. and A. Jansen, *Installed base as a facilitator for user-driven innovation.* International Journal of Telemedicine and Applications. , 2012.
10.   Sambamurthy, V. and R.W. Zmud, *Arrangements for Information Technology Governance: A Theory of Multiple Contingencies.* MIS Quarterly, 1999. **23**(2): p. 261-290.
11.   Sambamurthy, V. and R. Zmud, W. , *Arrangements for Information Technology Governance. A Theory of Multiple Contingencies.* MIS Quarterly, 1999. **23**(2): p. 261-290.
12.   Peters, B.G., *The Future over Governing*. 2001: University Press of Kansas, Kansas.
13.   Christensen, T. and P. Lægreid, eds. *New Public Management. The Transformation of Ideas and Practice*. 2001, Aldershot: Ashgate.
14.   Benbasat, I., D.K. Goldstein, and M. Mead, *The Case Research Strategy in Studies of Information Systems.* MIS Quarterly, 1987. **11**(3): p. 369-386.
15.   Christensen, T. and P. Lægreid, *A Transformative Perspective on Administrative Reforms* in *New Public Management . Th etransformation of ideas and practise* T. Christensen and P. Lægreid, Editors. 2000, Ashgate Publishing Company: Hampshire
16.   Jansen, A. and T. Tranvik, *The State of IT Governance: Patterns of Variation at the Central Government Level in Norway*, in *Electronic Government - 10th IFIP WG 8.5 International Conference, EGOV 2011.*, M. Janssen, et al., Editors. 2011, Springer 2011: Delft, The Netherlands. p. 159-172.
17.   Jansen, A. and I. Berg-Jacobsen, *Styring av den elektroniske forvaltning i Norge - en tilstandsrapport*. (ICT governance in the Norwegian Public Sector. Complex ed. N.R.C.f.C.a. Law. Vol. 1 /2011. 2011, Oslo, Norway: Unipub.
18.   DIFI, *Veikart for nasjonale felleskomponenter*. (Roadmap for core II components, 2014. Working paper,
19.   Olderbakk, H., *Exchange of personal information in public sector. Report,Ministry of Fianance,* 2007.
20.   AltInn, *Innovasjon i Altinn. (Innovation in Altinn)* 2013 URL: https://altinnett.brreg.no/no/Nyheter/Innovasjon-i-Altinn/
21.   Brekk, L.P., *Samarbeid for Digital firvaltning.*. 2014. Presentation 1.12.2014 URL: https://altinnett.brreg.no/PageFiles/7816/3%20Brekk_Altinndagen_2014_endelig.pdf
22.   Cadastre, *Applications based on Matrikkelen in Norway.* 2014.. URL: http://kartverket.no/Documents/Matrikkel/adresse/Nordisk%20adressesamarbeid/TheMatrikkelSystem_Norway.pdf
23.   Kihle, U.. *Matrikkelen Felleskomponent . Innhold, styring og organisering* 2014.. Presentation . URL: http://www.uio.no/studier/emner/jus/afin/FINF4001/h14/heldagsseminar-23.10.-foredrag/matrikkelen_unndkihle_sem23102014.pdf
24.   Bygstad, B., "Information Infastructure as Organization: A Critical Realist View", *Information Infastructure as Organization: A Critical Realist View.* ICIS 2008 Proceedings., 2008.

This page intentionally left blank

e-Government Services

This page intentionally left blank

245

Standardization of Service Descriptions, Process Models and Forms in Public Administrations: Results from a Survey in Germany

Hendrik SCHOLTA [a,1], Dian BALTA [b], Petra WOLF [b], Jörg BECKER [a] and
Helmut KRCMAR [b]

[a] *University of Muenster – ERCIS, Münster, Germany*
[b] *Fortiss – An-Institut der TU München, Munich, Germany*

Abstract. Public administrations face a high need for service orientation, process efficiency as well as digitalization of documents. Thus, service descriptions, process models and forms are relevant artifacts of administrative procedures. Due to their interrelation, an integrated standardization regarding these information artifacts is promising. Our overall research objective is to analyze initiatives that aim at a standardization regarding service descriptions, process models and forms in German public administrations. This article provides a first step of ongoing research to reach this aim by investigating initiatives that deal with forms. It builds upon interviews with public managers of 15 IT standardization initiatives and describes those concerning criteria such as supported standardization features, targeted federal levels and number of application domains. Given the lack of standardization found in practice, this article raises the need for an instrument that enables the integrated standardization of service descriptions, process models and forms in public administrations.

Keywords: Standardization, Reference Model, Process, Form, Public Administration

1. Introduction

Public administrations are confronted with an increased demand for digitalization [1], process efficiency [2] and service quality [3]. Digitalization can be achieved by substituting paper-based documents and forms with electronic equivalents throughout administrative procedures. Business process management and process modeling has gained more relevance in public administrations in order to construct efficient sequences of activities and enable the (semi-)automatic processing of electronic documents. Transparent service descriptions for external stakeholders such as citizens and companies can lead to an increased service orientation due to higher user conven-

---

[1] Corresponding author, e-mail: hendrik.scholta@ercis.uni-muenster.de.

ience. Hence, service descriptions, process models and forms are important information artifacts of administrative procedures [4].

A standardization of these information artifacts can lead to several benefits for public administrations, citizens and companies. Due to a harmonization of service descriptions and forms, a more uniform appearance of public administrations to citizens and companies can be achieved. A standardization of processes can result in higher efficiency if unnecessary activities are removed. Companies can benefit from accelerated data exchanges due to unified interfaces through standardized forms. Additionally, certainty in legal interpretation can be achieved in public administrations due to guidelines on the amount of requested information on forms and its processing.

In this article, standardization refers to the provision of reference models to public administrations. "Reference models are generic conceptual models that formalise recommended practices for a certain domain" [5, p. 595]. Hence, a reference model may represent a standardized artifact, i.e. recommend a design of the artifact. The reference model can be used in order to implement the artifact. Hence, we restrict standardization to the development and provision of reference models and do not focus on the implementation of the standardized artifacts.

An integrated standardization – i.e. a standardization of the three information artifacts of administrative procedures in a holistic manner - offers potential for synergies due to the interrelation of the artifacts. First, service descriptions contain information relevant for external stakeholders, e.g. the actions a citizen or company has to perform as well as the inputted and outputted forms. Hence, service descriptions provide valuable information for the standardization of processes and forms. Second, process models offer information on internal activities which includes details on data that is requested on forms in order to be able to deliver the services. Third, the order of processed information in process models may even suggest an order of fields on forms. Thus, process information is useful during the standardization of forms. Hence, in this context the term 'integrated' does not mean a necessarily simultaneous standardization of the different artifacts but refers to the usage of other artifacts' information in the standardization process. Moreover, higher consistency and convenience for public administrations are achieved if there is an integrated platform as single point of truth to assess reference models for the three information artifacts on administrative procedures.

By considering initiatives that primary deal with forms as a first step in our ongoing research, the aim of this paper is to provide first insights on the following research question: *To what extent do initiatives exist that aim at an integrated standardization of service descriptions, process models and forms in German public administrations?* Initiatives cover short-term projects as well as durable institutions which may have just been started or well-established.

This article is structured as follows. Section 2 presents the research background. In section 3, the research design is described. Afterwards, the results are provided in section 4 and discussed in section 5. Finally, a conclusion is drawn and an outlook on future work is given in section 6.

2. Research Background

The standardization of services, processes and forms in public administrations has been considered in scientific literature. The concept of *shared service centers* (SSC) in public administrations [6–8] describes the consolidation of activities at an organization to deliver common services to an internal group of partners. The notion of SSC is relevant in the context of this article since a SSC requires a standardization of processes since each partner does not carry out processes individually but the execution of processes is bundled at a SSC. Additionally, a standardization of forms occurs since forms may be processed by standardized processes in the context of SSCs.

Besides, a standardization of forms is conducted when *interoperability* is established in public administrations [9–11] in order to enable the information exchange between IT systems. Using data exchange standards, electronic information can be transmitted automatically. As forms are views on data [12], a standardization of electronic forms is conducted by interoperability concepts and data exchange standards.

The concept of *reference models* has been studied in the context of public administrations mainly with the focus on reference process models [13–15]. Reference models are relevant for standardization since they may represent a standardized artifact, e.g. a standardized process. Hence, reference models provide valuable information for standardization purposes. For instance, a standardized process can be represented as reference process model. In order to share knowledge on reference process models, dedicated libraries [16, 17] exist which aim at a collaborative business process management among public administrations.

3. Research Design

In order to address the research question, an empirical study of the phenomenon of standardized service descriptions, process models and forms was performed. The overall research approach can be split into two steps: (1) The identification of standardization initiatives and (2) the in-depth evaluation of these initiatives.

In the first step, primary data was gathered through a cross-sectional survey in order to identify standardization initiatives in German public administrations. We developed a questionnaire with a fixed set of open questions. Since we aim at a description of the status quo regarding the phenomenon, the addressees were selected as a representative cross section of federal and state administrations. Each addressee – employed in an IT management position – was questioned to describe the initiatives towards standardization of services, processes and forms in the corresponding organization. Additionally to the primary data gathered through questionnaires, we conducted a literature review [18] in order to discover the landscape of standardization initiatives as presented in German public meetings, conferences and congresses. The sampling of data resulted in a set of 40 potentially relevant initiatives with focus on standardization of service descriptions, process models or forms.

In the second step, additional secondary data was obtained through the projects' web sites and publications. Complementary primary data was collected by interview-

ing public administration employees engaged in the 40 initiatives. Upon interviewees' request for preparation, we sent the scheme for analysis that should allow us to describe the status of standardized services, processes and forms in German public administrations. Finally, we were able to conduct a total of 40 structured in-depth interviews in person or per telephone, with duration between 15 and 45 minutes per interview. The results of the interviews were used to complement the information on the initiatives obtained during the first step. After the analysis of the interviews, we considered several potential initiatives as not relevant. Eventually, we identified 30 relevant initiatives in German public administration towards the standardization of services, processes and forms. As a first step of our ongoing research, this paper focuses on 15 initiatives which integrate the standardization of forms into their scope.

Using the data from the two steps described above, the 15 initiatives were evaluated using the criteria presented in table 1. Whereas the first seven criteria are generally applicable, the following seven criteria are artifact-specific for forms. The criteria have been derived in the course of a workshop series with domain experts, method experts, practitioners from the private and public sector and researchers.

Table 1. Criteria for Evaluating the Standardization Initiatives

Criterion	Rationale
1. Artifact	As described above, service descriptions (S), process models (P) and forms (F) are important artifacts of administrative procedures.
2. Classification Scheme	A classification scheme is important in order to pre-structure existing elements which are standardized and identify common segments. Additionally, it supports the detection of suitable standardized service descriptions, process models and forms for a specific case.
3. Structural Standardization	This criterion covers the actual standardization of the artifacts' structures.
4. Organizational Standardization	This criterion indicates whether the initiative aims at a standardization of the procedures to develop, maintain and apply the standardized artifacts.
5. Technical Standardization	In order to enable machines to further process the standardized artifacts, a standardized technical specification of the information artifacts is required.
6. Federal Level	This criterion is important since an initiative is more comprehensive the more federal levels it covers.
7. Number of Application Domains	This criterion affects the extent of initiatives since initiatives with more application areas (e.g. "Environment" and "Education") are usable in more scenarios.
8. Fields	Fields are used on forms to enter the requested information, e.g. the street or postcode. They have metadata such as their labelling and datatype. A standardization of fields enables their reuse on different forms and harmonizes their occurrences.
9. Groups of Fields	Similarly to fields, the standardization of groups of fields is useful since fields are normally grouped to logical entities. For instance, the group 'address' consists of the atomic fields 'street', 'street number', 'postcode' and 'city'.
10. Entire Forms	The standardization of forms as entire entities is relevant since it affects their metadata such as the name and their contents, i.e. the fields and groups of fields.

Criterion	Rationale
11. Editorial Processes	The definition of a process of editing is useful in order to incorporate all stakeholders in the standardization process according to their interests as well as roles and ensure the appropriate creation, maintenance and application of the standardized artifacts.
12. Quality Criteria	The provision of a set of quality criteria is relevant since they should be considered during the creation and maintenance process and aim at a correct, consistent and complete creation of information.
13. Dependency Rules	Dependency rules are useful to describe relationships between form elements. For instance, a form should ensure that a starting date is always scheduled before an end date when a citizen completes a form.
14. Layout	A standardization of layout characteristics is useful since it affects the positions of the forms' elements, i.e. their order.

4. Results

This section presents the first results of the study. Table 2 contains the results regarding supported standardization features. Each of the investigated initiatives deals with forms but may support more than one standardization feature. The rows contain figures for each possible combination of services, processes and forms that at least covers forms. The last row presents aggregated figures for all form-related initiatives.

In general, there are 15 initiatives which deal with forms. There are few initiatives which integrate forms with services and processes for standardization purposes. Seven initiatives focus on more than one artifact. On the contrary, eight initiatives consider forms individually. One initiative aims at a standardization of all three artifacts of administrative procedures.

Table 2. Supported Standardization Features (Criteria 2 to 5)

Artifact	Classification Scheme	Structural Standardization	Organizational Standardization	Technical Standardization	Total
S, P, F	1	0	0	0	1
S, F	0	1	0	1	2
P, F	2	3	3	1	4
F	1	3	1	4	8
Total	4	7	4	6	15

Considering the supported standardization features, the results reveal that a classification scheme and organizational standardization are delivered by the fewest number of initiatives. In general, two initiatives provide three or more standardization features (figure is not given in the table). They deal with forms and processes and one of them covers all four standardization features. On the contrary, the initiative supporting all three artifacts of administrative procedures only provides a classification scheme.

The results concerning the practical application at different federal levels are displayed in table 3. The initiatives may support more than one level. The results reveal an almost equal distribution on the three federal levels. Besides these levels, the initia-

tives are used in other institutions. In total, eight initiatives address all federal levels, two of them are additionally applied in other organizations (figures are not given in the table). Four out of those eight initiatives deal only with forms, two consider forms and processes and further two forms and services.

Table 4 provides results for the criterion "Number of Application Domains". Four initiatives are dedicated to one specific domain and three initiatives are generic approaches applicable to a wide range of areas. Each of these generic approaches aims at standardizing only forms individually, or the combination of forms and processes or forms and services, respectively.

Table 3. Federal Level (Criterion 6)

Artifact	Fed. Government	Fed. States	Local Authorities	Other Institutions	None / No Indication	Total
S, P, F	0	0	0	0	1	1
S, F	2	2	2	0	0	2
P, F	2	3	4	2	0	4
F	6	5	5	1	1	8
Total	10	10	11	3	2	15

Table 4. Number of Application Domains (Criterion 7)

Artifact	1	4	>8	No Indication	Total
S, P, F	0	0	0	1	1
S, F	0	0	1	1	2
P, F	2	0	1	1	4
F	2	1	1	4	8
Total	4	1	3	7	15

Table 5 provides insights on the initiatives' comprehensiveness with regard to the artifact-specific criteria for forms. Whereas two-thirds of the initiatives support a standardization of entire forms, layout and editorial processes are barely considered. The highest quantity of supported standardization aspects is five out of seven (figure is not given in the table). Three of the initiatives with more than three aspects of standardization cover processes and forms while two initiatives deal with forms individually.

Table 5. Comprehensiveness of Standardization (Criteria 8 to 14)

	S, P, F	S, F	P, F	F	Total
Entire Forms	0	2	4	4	10
Groups of Fields	0	0	3	4	7
Fields	0	1	3	4	8
Editorial Processes	0	0	0	1	1
Quality Criteria	0	0	1	2	3
Dependency Rules	0	0	4	2	6
Layout	0	0	1	0	1
No Indication	1	0	0	1	2
Total	1	2	4	8	15

5. Discussion

Our findings suggest that initiatives that aim at an integrated standardization of service descriptions, process models and forms in German public administrations exist to a limited extent, although its existence is of crucial importance.

At first, the integration of the three artifacts of administrative procedures as well as corresponding standardization features is insufficiently present in practice. As table 2 reveals, there is one initiative that aims at an integration of services, processes and forms, but it has limited standardization features. Similarly, there is only one initiative covering all standardization features and supports two information artifacts of administrative procedures. Secondly, an integrative initiative for all information artifacts which supports all federal levels and a wide range of application domains is missing. Several initiatives address all federal levels (cf. table 3) or target multiple application areas (cf. table 4). However, the combination of all federal levels and multiple application areas with service descriptions, process models and forms is missing. Thirdly, even the extent regarding the artifact-specific standardization features of forms is limited (cf. table 5). There is not a single initiative that addresses all aspects of comprehensiveness. Additionally, two standardization aspects – editorial processes and layout – are only provided by one initiative.

In summary, we are able to identify an existing gap in standardization initiatives of service descriptions, process models and forms in German public administrations. Hence, three is a need for an instrument that enables the integrated standardization and covers the criteria presented in table 1.

6. Conclusion and Outlook

This study develops an understanding regarding the current extent of existence of initiatives that aim at an integrated standardization of service descriptions, process models and forms in German public administrations. The findings reveal that an integrated initiative does not exist.

Since this article presents ongoing research, it is especially subject to limitations and potential for further research. The main focus of our upcoming activities is to integrate the remaining 15 initiatives which do not cover forms into the scope of investigation. Although this will not yield an integrated approach, their analysis will provide insights on the integration of services and processes and the comprehensiveness regarding each of these artifacts individually. For this purpose, the set of criteria is to be extended, especially for services and processes. Moreover, we recognize the focus on German public administration initiatives as limitation which may be expanded to other countries in the future.

Despite these limitations, we still believe that our findings are already of value for both theory and practice. Implications for theory include the recognized gap of a detailed approach of the integrated standardization of service descriptions, process models and forms as well as guidelines for its operationalization and application in practice. Researchers have the methods such as business process modeling and the con-

cepts of reference models as well as the domain knowledge to design such an approach. Additionally, the set of criteria to describe standardization initiatives is a contribution itself. Implications for practice include the integrated view on standardization and the existing potentials for integration given the current heterogeneity of initiatives. As described above, the development and implementation of an integrated standardization approach can yield significant synergies such as process efficiency, certainty in legal interpretation, customer convenience and facilitated data exchange.

References

[1] Dunleavy, P., Margetts, H., Bastow, S., Tinkler, J.: New Public Management Is Dead - Long Live Digital-Era Governance. J. Public Adm. Res. Theory. 16, 467–494 (2005).

[2] Houy, C., Fettke, P., Loos, P.: Empirical research in business process management – analysis of an emerging field of research. Bus. Process Manag. J. 16, 619–661 (2010).

[3] Teicher, J., Hughes, O., Dow, N.: E-government: a new route to public sector quality. Manag. Serv. Qual. An Int. J. 12, 384–393 (2002).

[4] Ahrend, N., Wolf, P., Räckers, M., Dentschev, A., Jurisch, M.: Federal Information Management – Context and Effects. Proceedings of the 44. Jahrestagung der Gesellschaft für Informatik. pp. 1499–1511. Stuttgart (2014).

[5] Rosemann, M.: Application Reference Models and Building Blocks for Management and Control. In: Bernus, P., Nemes, L., and Schmidt, G. (eds.) Handbook on Enterprise Architecture. pp. 595–615. Springer-Verlag, Berlin, Heidelberg (2003).

[6] Janssen, M., Joha, A.: Motives for establishing shared service centers in public administrations. Int. J. Inf. Manage. 26, 102–115 (2006).

[7] Grant, G., McKnight, S., Uruthirapathy, A., Brown, A.: Designing governance for shared services organizations in the public service. Gov. Inf. Q. 24, 522–538 (2007).

[8] Janssen, M., Wagenaar, R.: An Analysis of a Shared Services Centre in E-government. Proceedings of the 37th Hawaii International Conference on System Sciences. Big Island (2004).

[9] Scholl, H.J., Klischewski, R.: E-Government Integration and Interoperability: Framing the Research Agenda. Int. J. Public Adm. 30, 889–920 (2007).

[10] Guijarro, L.: Interoperability frameworks and enterprise architectures in e-government initiatives in Europe and the United States. Gov. Inf. Q. 24, 89–101 (2007).

[11] Gottschalk, P.: Maturity levels for interoperability in digital government. Gov. Inf. Q. 26, 75–81 (2009).

[12] van der Aalst, W.M.P., Weske, M., Grünbauer, D.: Case handling: a new paradigm for business process support. Data Knowl. Eng. 53, 129–162 (2005).

[13] Karow, M., Pfeiffer, D., Räckers, M.: Empirical-Based Construction of Reference Models in Public Administrations. In: Bichler, M., Hess, T., Krcmar, H., Lechner, U., Matthes, F., Picot, A., Speitkamp, B., and Wolf, P. (eds.) Proceedings of the Multikonferenz Wirtschaftsinformatik 2008. pp. 1613–1624. Munich (2008).

[14] Algermissen, L., Delfmann, P., Niehaves, B.: Experiences in Process-Oriented Reorganisation through Reference Modelling in Public Administrations - The Case Study Regio@KomM. Proceedings of the Thirteenth European Conference on Information Systems. Regensburg (2005).

[15] Hinkelmann, K., Thönssen, B., Probst, F.: Referenzmodellierung für E-Government-Services. Wirtschaftsinformatik. 47, 356–366 (2005).

[16] Eid-Sabbagh, R.-H., Kunze, M., Weske, M.: An Open Process Model Library. In: Daniel, F., Barkaoui, K., and Dustdar, S. (eds.) Proceedings of the BPM 2011 International Workshops. pp. 26–38. Clermont-Ferrand (2011).

[17] Algermissen, L., Instinsky, M., Schwall, J.: BPM as a Strategic Tool for Administrative Modernization: The IMPROVE Approach. In: Becker, J. and Matzner, M. (eds.) Proceedings and Report of the PropelleR 2012 Workshop. pp. 51–56. Moscow (2012).

[18] Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Q. 26, 13–23 (2002).

A Framework for Evaluating Online Services and E-Participation Tools: UN Methodology Application to Russian Regions

Lyudmila BERSHADSKAYA (VIDIASOVA), Andrei CHUGUNOV, Dmitrii
TRUTNEV [a], Gulnara ABDRAKHMANOVA [b]
[a] *ITMO University, Russia*
[b] *Higher School of Economics, Russia*

Abstract. The article discusses the experience of UN international methodology for assessing e-government development at federal and regional levels in Russia. United Nations procedure for e-government development measuring is widely used in the world and allows comparison of 191 countries. Focusing on international standards, the authors conducted survey of federal and regional executive authorities' web-sites. The article describes methodological approaches to web monitoring in accordance with UN methodology, web-monitoring results and recommendations to international methods' application in assessment of regional specificity of e-government and e-participation in Russia. The paper touches the limitations of international rankings and provides supply-side assessment of e-government implementation.

Keywords. e-government, E-Government Development Index, EGDI, e-participation, evaluation, international index, web-monitoring

1. Introduction

On the world stage it has been a tendency to expand the openness of government, online public services development, as well as e-participation tools for citizens and business engagement in decision making. In this regard, it is urgent to develop valid methods to measure the progress in this area. Various international organizations (UN, ITU, WB, OECD etc.) create indices for cross-country comparisons. UN E-Government Development Index (EGDI) serves as a guide to enhance competitiveness and performance implementation of government programs in a competitive countries, including Russia [1].

Russia's position in the international index is a key indicator of the governmental program "Information Society" (2011-2020). At the same time there is a legitimate question of whether international methodology in its original form can be applicable to estimate e-government and e-participation parameters within a particular country and its regions.

Under the order of the Ministry of Telecom and Mass Communications of the Russian Federation, the research team had conducted monitoring of federal and

regional executive authorities web-sites. In order to follow the key performance indicators of Russian ICT Ministry this research was aimed to assess the level of online services at regional level by using UN EGDI methodology.

2. Literature review

E-government measurement techniques are not a new topic in scientific literature. There is urgent discussion among scholars about what techniques could measure the success in this field appropriate. In world practice, United Nations E-Government Development Index is the most popular ranking used for cross-countries comparison.

Benchmarking and ranking tools, such as proposed index, are used by decision makers when devising information and communication policies and allocating resources to implement those policies [2 Heeks]. At the same time the researchers recognize its limitations in several fields. For that reason, more e-Government benchmarking studies focus on supply-side not back-office [3]. Rorissa claimed that "e-Government benchmarking methods become more problematic, and the critics' views more telling, when they move beyond objective, supply-side criteria (e.g., services offered via websites) to include calculated indices, psychometric measures, or other subjective indicators (e.g., human development index and internet use)" [4]. In the case of the EU, its e-Government benchmarks are simple, inexpensive, fairly transparent and replicable, and widely accepted and used [5].

For the purpose of finding best solutions for measurement, Rorissa [4] collected data from 582 e-Government websites sponsored by 53 African countries. The applied frameworks respond, in part, to the need for continuous assessment and reconsideration of generally recognized and regularly used frameworks. The authors recognized that almost all African governments now have some presence on the Web, including fully fledged e-Government web portals. However, the current status of e-Government services in African countries is not well documented in detail. They have computed e-Government indexes, produced e-Government rankings, and compared rankings to previous ones. A clear picture that emerges from the analysis and results is that although progress has been made, there is a long way to go, to bridge not only the North‑South divide when it comes to e-Government services but also among the various sub-regions [6].

E. Pnopoulou et al. [7] developd a framework for evaluating websites of public authorities based on 4 axes: website general characteristics, content, e-services and e-participation. They found the high level of content and general characteristics development (such as navigation, accessability, privacy). E-services and e-participation aspects are the least developed.

Huang et al. in their paper showed a close correlation between usability and credibility, as e-government websites with a high usability were perceived as having higher credibility, and vice versa [8].

Holzer and his team conducted U.S. survey of municipal websites (2 largest cities from each state) using the following parameters for assessment: privace/security, usability, content, services, citizens and social engagement [9]

Zhao et al. studies threats and opportunities of e-government web-security and used a combination of three methods – web content analysis, information security auditing, and computer network security mapping – for data collection and analysis.

The findings indicate that most state e-government sites posted privacy and security policy statements; however, only less than half stated clearly what security measures were in action [10].

Wu and Guo conducted specialized analysis of e-government performances of 31 provincial government websites in China using DEA [11] providing assessment of e-government efficiency differences both in different individual provinces and in different districts. The lead to the conclusion that the less developed western provinces achieve a higher efficiency mean than the eastern and middle ones in China. R. Medaglia [12] analyses publications on e-participation, identifying them as belonging to a single scheme consisting of the following categories: e-participation activities, e-participation actors, e-participation effects, contextual factors, e-participation evaluation. Fedotova et.al assesses the level of electronic participation (e-participation) initiatives that are promoted by government authorities in Portugal at local and national levels [13]. A. Prosser [14] investigates e-participation on the European Union level with the focus on legal basis and technical possibilities for citizens participation. L. Bershadskaya and colleagues conducted comparative analysis of the USA, the UK and Russian e-petition portals and distinguished strong and weak traits of each portal [15].

The scholars identified several technical issues with indexes and pointed to the need for a statistical tool that could be used to evaluate and guide the development of e-government ranking systems. Whitmore A. suggests to apply factor analysis to the ranking of e-government programs. The importance and high visibility of e-government rankings highlights the need for the use of a more sophisticated set of tools in order to ensure that "the rankings fully measure what they purport to and are not an artifact of their own design methodology" [16].

A review of the scientific research allows one to summarize that the development of e-participation mechanisms constitutes a global trend. At the same time there is a lack of studies focused on deep methodological analysis, as well as finding the borders of international techniques application.

3. Research Methodology

Monitoring of Russian executive authorities' websites was based on UNDESA methodology (UN agency which forms EGDI) [1], which has been selected by Russian authorities as key points. The web-monitoring allows to evaluates 4 stages of Online services index development:

- emerging online presence,
- enhanced presence,
- transactional presence,
- connected presence.

According to UN methodology, we assessed national portal, e-services portal as well as the websites of the related ministries of education, labor, social services, health, finance and environment protection.

Criterion "presence / absence" (1 point- yes, 0 - no) was taken as evaluation metric for complex assessment by researchers. The following approach was applied during the research: the presence of some indicators has opened opportunities to get points for its high-quality presentation (acess from home page, descriptions, analytical tools etc.).

This evaluation brought more detailed assessment of web-sites opportunities and functions.

In accordance with survey methodology each web-site could get maximum 38 points for 4 stage indicators. The 4th stage was assessed on web-sites only, becuase online services portals couldn't consists of such element in Russia.

As a result, web monitoring methodology has been corrected due to clarifiaction the list of criteria in two groups of objects: governmental web-sites and online services portals (Table 1).

Table 1. Web-monitoring structure: distribution of max possible points for stages and web-resources

Indicators (yes-1, no-0)	Total points	
	online services portals	web-sites
Stage 1. Emerging Online presence	8	12
Stage 2. Enhanced presence	11	11
Stage 3. Transactional presence	26	6
Stage 4. Connected presence	0	38
TOTAL	45	67

Online services index at regional level was calculated according to the formula used in UN survey. The final score was determined by summing the scores for all four stages of online services development.

The fourth stage of online services assessment relates to e-participation tools and methods, such as:

- calendar listing of upcoming e-participation activities;
- e-participation tools for public opinion assessment;
- feedback;
- archive on responces by government to citizens's questions, queries and inputs;
- e-voting on public policy issues;
- regulations and law online discussions with citizens;
- links to social networks;
- Open Data page or link to portal;
- Open Budget publication.

The authors considered e-particiaption development through these indicators achievement, corresponding to UN e-information, e-consultation and e-decision-making parts.

The authors have apllied specified approach for government portals and online services portals, taking into account rationality and necessity of certain indicators presence on each of them. Specially trained group analyzed the web-sites. If in the process of assessing the implications questions and ambiguities emerged, the group came together for discussion and final decision.

4. Findings

665 websites of federal and regional authorities and public services portals were assessed during the research. Web monitoring was carried out in two stages: in January and August 2014. After the first stage the methodology for assessment was corrected: evaluation criteria were separated between administration sites and online services portals. These changes were adopted in accordance with new trends of UN methodology. The results of the first stage of monitoring showed that the Ministry's website, under Russian law and practice, should not be present indicators presented in the portal services. For example, the website of the authority couldn't have user's personal account, and e-participation tools couldn't be presented at online services portal. etc. In this paper all graphs and tables with results belong to the 2d final stage of monitoring based on changed methodology.

4.1. Stages of Online Services Development

The monitoring results for federal and regional levels were calculated into online services indexes using the same approach of calculation as proposed in EGDI (fig.1). According to research results, federal level websites have demonstrated higher level of online services presentation.
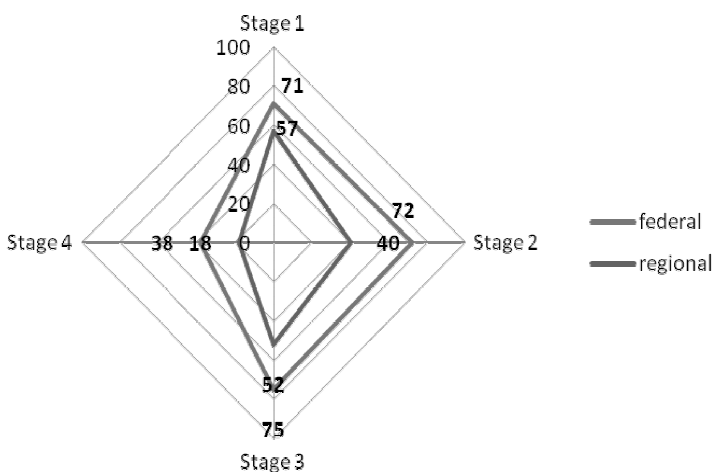
Figure 1. E-Government development Index, %. Source: http://pluto.netlight.ru/russia/e-goverment/

Online services index value of at regional level varies from 0.00 to 0.93. The first group (with the highest index value from 0.80 to 0.93) includes 9 regions; the second (0.60 to 0.79) - 32, the third (0.40 to 0.59) - 35, the fourth (lowest index value - less than 0.39) attributed 9 subjects of the Russian Federation.

4.2. eParticipation tools assessment

The study showed that the fourth stage hasn't been as much developed as other stages of online services in Russia (Table. 2). Information services and applications from citizens are the most popular functions at all sites. Special attention in Russia at federal

and regional levels is paid to open data (45%) and open budget (33%) publication. However, more sophisticated analytical tools that allow citizens to participate in political decision-making electronically are almost absent at all the analyzed web-sites.

Table 2. Presence of Online Services 4 Stage indicators, %

Indicators	The percentage of web-sites with indicators
Public e-box	more than 50%
E-participation tools to obtain public opinion Open Data page Open Data in machine- readable formats (XML, RDF, xls, CSV, HTML) Availability to attach files Opinion polls Calendar listing of upcoming e-participation activities Open Budget	31-50%
Links to social networks FAQ section Research results, analytics Regulations and law online discussions with citizens Archive on responses by government to citizens' questions, queries and inputs Safe identification system Access to special features from home page E-voting on matter of public policy Feedback analysis Blogs Chat rooms	10-30%
Open Budget in machine-readable formats GIS, maps Page for IT-developers Additional analytical tools, online calculator Web casting Demand assessment (popularity rankings etc.) Applications based on Open Data API on open budget API on public reports	less than 10%

Source: Web-monitoring results, 2014

According to web monitoring, there is lack of publication the feedback of citizens treatments on official web-sites. It's recommended to publish frequently asked questions and targeted answers to citizens' complaints, to develop API to access information about visits on sites. There are several elements of e-participation tools on web-sites (polls, chat rooms, blogs etc.). At the moment it's difficult to track complaints, there are no links to e-petition portal (Russian public initiative), to single portal of information disclosure and open data portal.

As stated in regular UN reports, Russia has managed to make a giant leap in EGDI ranking in 2012 thanks to reinforce the creation of the federal online services portal. In the ranking, published in 2014, the Russian Federation has retained 27th place, which indicates the need for careful analysis of the factors that may contribute to the positive dynamics. The authors recommend to develop e-participation tools in Russia in order to achieve the same significant progress that has been reached by online services portal creation.

5. Conclusions

Web-monitoring conduction in Russia, the adjustment methodology and research results allowed researchers to draw the following conclusions:

1. International methods can be applied to assess e-government and e-participation development at the national level. These methods are suitable for the assessment process at the federal level. For research at regions and municipalities it's needed to clarify the methodology in accordance with regional specific. Even taking into account EGDI as landmark of development (stated at federal level), its indicators should not copy blindly at regional level. Following the UN methodology was determined by Russian authorities indicators. While conducting research in other countries its important to study the program documents first with the purpose to find KPIs related to the specific area. The list of assessment indicators couldn't miss its criterion also.
2. The study revealed some limitations on the applicability of web monitoring techniques for different types of information resources: governmental websites and online services portals. This clarification should be considered by researchers from those countries where government services are presented on online services portals onle and services not related to government sites.
3. Additional statistical tolls should be applied to determine regional difference. In this study the authors applied regional division into four groups distingiushed by progress in online services implementation. This division is closed to normal distribution.
4. It's strongly recommeneded to develop coordination between the responsible authorities to publish departmental statistics on their websites, as well as the presentation of statistics and web-monitoring results on a single public web-site. Using this update data the researhers could ad multy-factor analysis for comlex index development.
5. For e-participation development, it's necessary to optimise web-sites for mobile devices, to create useful mobile applications using electronic services, payment services (applications must be demanded by citizens and should be developed for all major mobile platforms - iOS, Android, Windows Phone, Windows 8), as well as to present API for open data and other public information.

Testing the UN methodology at national and regional levels is extremely useful for the construction of monitoring systems for e-services and e-participation assessment at the level of inter-state associations, such as the Eurasian Economic Union. The authors believe that combination of supply-side assessment with demand-side measurement (polls, downloads, registrations, citizens' feedback) could bring a complex picture of e-government and e-participation development.

References

[1] UN E-Government Survey 2014 (2014), http://unpan3.un.org/egovkb/en-us/Reports/UN-E-Government-Survey-2014

[2] R. Heeks, Understanding and measuring eGovernment: International benchmarking Studies, *UNDESA Workshop "E-participation and E-Government: Understanding the present the creating the future"* (2006), 1-46.

[3] D. Janssen, S. Rotthier, K. Snijkers, If you measure it they will score: An assessment of international e-Government benchmarking, *Information Polity* 9 (2004), 121‑130.

[4] A. Rorissa, D. Demissie, T. Pardo, Benchmarking e-Government: A comparison of frameworks for computing e-Government index and ranking, *Government Information Quarterly* 28 (3) (2011), 354-362.

[5] C. Codagnone, T.A. Undheim, Benchmarking e-Government: Tools, theory, and practice, *European Journal of ePractice* 4 (2008), http://www.epractice.eu/files/4.2_0.pdf

[6] A. Rorissa, D. Demissie, An Analysis of African E-Government services websites, *Government Information Quarterly* 27 (2) (2010), 161-169.

[7] E. Panapouou, E. Tambouris, K.Tarabanis, A framework for evaluating websites of public authorities, *Aslib proceedings: New Information perspectives* 60 (5) (2008), 517-546.

[8] Z. Huang, M. Benyoncef, Usability and credibility of e-government websites, *Government Information Quarterly*n31 (4) (2014), 584-595.

[9] M. Holzer, M.Fudge, R. Shick, G. Stowers, A. Manoharan, U.S. Municipalities E-Governance Survey 2010-2011, *Rutgers* (2011).

[10] Zhao, J.J., Zhao, S.Y.: Opportunities and threats: A security assessment of state e-government websites. Government Information Quarterly, 27 (1), 49-56 (2010)

[11] J. Wu, D. Guo, Measuring E-government performance of provincial government website in China with slacks-based efficiency measurement, *Technological Forecasting and Social Change* (2015)

[12] R. Medaglia, eParticipation research: Moving characterization forward (2006-2011), *Government Information Quarterly* 29 (2012), 346-360.

[13] O. Fedotova, L. Teixeira, H. Alvelos, E-participation in Portugal: evaluation of government electronic platforms, *Procedia Technology*, 5(2014), 152-161.

[14] A. Prosser, eParticipation on the European Union Level, *Technology-Enabled Innovation for Democracy, Government and Governance*, Springer Berlin Heidelberg (2013), 1-8.

[15] L. Bershadskaya, A. Chugunov, D. Trutnev, E-Participation Development: A Comparative Study of the Russian, USA and UK E-Petition Initiatives, *Proceedings of International Conference ICEGOV 2013 ACM Press* (2013), 73-76.

[16] A. Whitmore, A statistical analysis of the construction of the United Nations E-Government Development Index, *Government Information Quarterly* 29 (1) (2012), 68-75.

Adoption of Common Service Centre with presence of intermediary for delivery of e-Government services: A conceptual framework

Rajesh SHARMA[a,1] and Rajhans MISHRA [b]

[a] *Research Scholar, Indian Institute of Management Indore (India)*
[b] *Assistant Professor, Indian institute of Management Indore (India)*

Abstract. Government of India is making large investments for providing e-Government services to citizens under its National e-Governance Plan (NeGP) 2006. Common Service Centres (CSCs) are envisaged as one of the pillars of the delivery mechanism of e-Government services in rural areas where the availability of Internet, literacy levels of citizens and PC penetration is rather low. This channel of delivery is quite different from that of developed countries because citizens do not interact directly with technology. An intermediary called as the Village Level Entrepreneur (VLE) delivers the services to the citizens and technology is manifested to the citizens by results of his interactions at the CSC. The study aims to develop a conceptual framework for adoption of CSCs by the citizens in this unique context. A qualitative study is carried out to identify the factors that can augment the technology adoption constructs. The study reveals that quality of service delivered at CSC may be an important determinant for adoption of CSCs. The proposed conceptual framework is intended to form the basis of future empirical studies in similar context in India and in other developing countries.

Keywords. Common Service Centre, adoption, e-Government, intermediary, rural, India, quality of service

1. Introduction

The governments all over the world are adopting Information and Communication Technology (ICT) for transforming government administration [1] in the form of e-Government. Governments are looking at e-government as a mode of delivery that has the potential for reducing the governance costs by minimizing the wastage, eliminating corruption by improving transparency and promising a better future to the citizens by opening up opportunities for reduction in rural poverty and inequality [2]. Accordingly, huge investments are being made in promoting e-government with the objective of achieving effective delivery of government services.

---

[1] Corresponding Author.

The mode of delivery of e-government services in rural areas of developing countries needs to be quite different from that of developed countries to overcome the handicap of low Internet penetration and low computer literacy of these areas. For example, Government of India has introduced the concept of Common Service Centres (CSCs) which are ICT enabled front end service delivery points at the village level for delivery of government services (www.csc.gov.in). These CSCs are manned by a Village Level Entrepreneur (VLE) involved in delivering the service. As compared to developed countries, the important contextual differences are presence of intermediary, involvement of public-private partner and demographic differences in respect of education, income and computer literacy. These differences point to the possibility that the determinants of adoption of Common Service Centres for availing e-Government services may be quite different from developed countries where delivery of government services is directly a result of interaction of citizens with the e-Government websites.

Research initiatives in adoption of e-government services in the context of citizens of rural areas in developing countries have been rather sparse [3]. Recent meta-analysis of existing research on e-government adoption studies [4] reveals out of 63 relevant empirical studies in this field till 2013, only two studies were reported from India [5][6] and both were in urban context only. Role on intermediary in delivering e-Government service has been explored in only one study [7] which was carried out in urban context in the city of Madinnah in UAE. No study has specifically looked at the context involving rural population, presence of intermediary and role of public-private partnership for delivering e-Government service.

The above discussions point to a research gap in respect of empirical studies in the context where the services are delivered through intermediaries rather than directly being availed by citizens. Empirical studies based on such framework are likely to lead to new insights for stimulating adoption of e-government in rural areas. The findings can be put to use by the practicing managers for reducing the digital divide between rural and urban areas in developing countries.

Accordingly, the objective of this research is to develop a conceptual framework of acceptance of common service centres by the citizens of rural areas for availing e-government services. The research contributes to the existing body of knowledge on e-government adoption by proposing a framework that takes into account unique contextual factors of delivery of e-government services in rural India. The framework can be used for studying the phenomena of adoption of e-government in other developing countries too that have a similar delivery mechanism.

Mixed method approach has been adopted to ground the research in practical perspective [8][9]. For this purpose, the relevant determinants of adoption are obtained from the literature by looking at various theories of technology adoption. Based on theoretical foundation of literature and interpretive analysis of interview results [10], a framework is proposed to explain adoption of CSCs in Indian context.

2. Literature Review
The literature review is carried out in two parts. In the first part, prior research in e-Government adoption is reviewed to identify the theories of technology adoption that

have been used in various studies. In the second part, delivery mechanism of e-Government services in rural India is reviewed.

2.1 Prior research in e-Government adoption

Several adoption models such as Diffusion of Innovation, Theory of Planned Behaviour, TPB[11], Technology Adoption Model, TAM[12], Perceived characteristics of innovation, PCI [13], D&M IS Success Model, Unified Theory of Acceptance and Use of Technology, UTAUT [14] etc. have been empirically tested in different contexts. Survey papers [15] [16] provide a gist of studies related to technology adoption studies in various contexts. Some other models and theories that have been used in these studies are SERVQUAL model [17], big five personality model [18] and actor network theory [19].

Review of existing literature specifically on e-Government adoption has been carried out which reveals that TAM have been extended with constructs from other theories such as TPB, DOI, PCI, web trust model [20] etc depending upon the context of service delivery.

Another widely used theory of adoption is the UTAUT model which has integrated eight earlier adoption models and theories. It has been used in several empirical studies by extending it with constructs that are not part of UTAUT such as trust [21], perceived security control [22], perceived risk and optimism bias [23].

Apart from extensive use of TAM and UTAUT as the base model for studying e-Government adoption, some authors have based their research on other theories such as Big Five Model [24], DOI [25], DeLone & McLean IS Success Model [26], Theory of Planned Behaviour [27] and SERVQUAL model [28].

2.2 e-GOVERNMENT DELIVERY MECHANISM IN INDIA

National e-Governance Policy: To realize the potential of use of IT in governance, National e-Governance Policy (NeGP) was unveiled in 2006 with the vision to *"make all Government services accessible to the common man in his locality through common service delivery outlets and ensure efficiency, transparency and reliability of such services at affordable cost"* (www.negp.gov.in). The policy emphasized on process re-engineering with centralized initiative and decentralized implementation with emphasis on Public Private Partnership (PPP) for ensuring fully electronic delivery of services. Expenditure of nearly Rs 40,000 crores was envisaged by the government for the initiatives under NeGP.

Role of Common Service Centres: As per NeGP, Common Service Centres (CSC) was envisioned as the primary delivery channel of e-government in India in order to alleviate the difficulties of low Internet penetration and poor infrastructure in rural areas. They were envisaged to provide shared ICT facility having computers and Internet connectivity for public access of various information services. The CSC operator (called as the Village Level Entrepreneur or VLE) is expected to provide hand-holding functions and act as an interface between the citizens and the e-Government portals.

The CSC scheme is based on the Public Private Partnership (PPP) model that envisages a 3-tier structure consisting of the VLE catering to a cluster of 3-4 villages; the Service Centre Agency (SCA), that will be responsible for a division of 500-1000 CSCs; and a State Designated Agency (SDA) identified by the State Government responsible for managing the implementation in the entire State (Website www.deity.gov.in).

3. Qualitative study to identify additional constructs

Literature review in preceding section reveals that there are several unique features in rural India such as presence of intermediary at Common Service Centre (CSC), lack of reliable ICT infrastructure, lack of computer literacy among the citizens and dependence on PPP partner for providing government services. The model for explaining adoption behaviour needs to take into account these salient features in addition to the constructs explored in technology adoption research.

In order to ground the research in practical perspectives, in-depth interviews were held with stakeholders such as citizens, VLEs, SCA and government officers to elicit their views on factors that may influence intention to use CSC. . The responses were recorded and common themes were generated by the interpretive approach which combines the perspectives of different stakeholders into a common set of beliefs. According to [29], theoretical saturation is generally achieved after sample size of 12. However, in this case we collected responses of 15 stakeholders and stopped when no new information was forthcoming [30]. The responses were coded and segregated into common themes as summarized in Table1:

Table 1: Themes of determinants of CSC adoption

S. No.	Statement related to adoption of CSC	Underlying Theme
1	Accessibility of CSC location from villages	Ease with which services at CSC can be used
2	Distance of CSC from village	
3	Flexible timing of CSC (i.e., no strict office hours)	
4	Simplicity of processes at CSC	
5	Accuracy of information made available by the CSC	Usefulness of CSC
6	Instant completion of work without the need for follow-up visits	
7	No requirement to pay bribes	
8	Number of services available at CSC	
9	Whether the information/certificate made available by CSC is accepted by other organizations	
10	Behavior of VLE	Service quality at CSC
11	Comfortable environment at CSC such as fan, light, water etc.	
12	CSC should not have long queues	

13	Timely completion of service as per promise	
14	VLE should be helpful	
15	VLE should be knowledgeable	
16	Whether proper feedback is given regarding likely date of fulfillment	
17	Positive word of mouth about the CSC	Influence of opinion of society on intention for using CSC
18	Reputation of the CSC in the village	
19	Reputation of the VLE in the village	

4. Proposed Conceptual Model

Most of prior research in technology adoption has involved direct interface of technology with the users. The situation in present context is different because citizens at CSCs do not directly handle the technology. It can be a matter of debate whether the constructs of technology adoption should apply in this context. According to [31] [32], *"technology is a design for instrumental action that reduces the uncertainty in the cause-effect relationships involved in achieving a desired outcome"*. Therefore any action that helps in achieving the objective with less ambiguity can be termed as a technology. In the present context, CSC enables service delivery in rural areas without the requirement of citizens to visit headquarters located far away from their village. The new system of CSC (consisting of technology and the intermediary) reduces the uncertainty in achieving a desired outcome (availing government service) and therefore can be viewed as a technology even though the citizens do not interact with the website themselves. This scenario clearly indicates that the technology adoption constructs should form the basis for the conceptual model.

It is necessary to include other constructs in the framework that emerge from the qualitative study. First is the behaviour of VLE with the users, second is timely provision of the services and third relates to physical facilities at the CSC. These dimensions relate to quality of service delivered at the CSCs. Accordingly the conceptual model proposed in Figure 1 includes these important constructs apart from the technology adoption constructs, namely, ease of use of CSC, usefulness of CSC and social influence.

Figure 1: Proposed Conceptual model
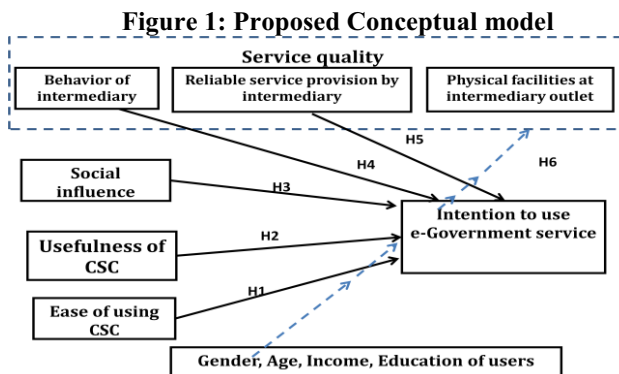
These constructs are described as follows:

Ease of use of CSC: In our context it is defined it as *"the degree of ease associated with obtaining government services through CSCs"* and this leads to the hypothesis H1: *Ease of using CSC is positively related to the behaviour intention for using CSC to avail e-Government services.*

Usefulness of CSC: In our context, it is defined as '*the degree of ease associated with obtaining government services through CSCs'* and leads to hypothesis **H2: *Usefulness of CSC is positively related to the behaviour intention for using CSCs.***

Social Influence: In our context it is defined as *"the degree to which an individual perceives that it is important for others to believe that he or she should obtain government services through the CSC"* and leads to hypothesis **H3: *Social influence is positively related to the behaviour intention for using CSCs*** .

Behaviour of intermediary: It is defined as *"the empathy and responsiveness that is displayed by the intermediary towards the users approaching CSC for availing e-Government services"* and this leads to hypothesis **H4: *Greater empathy and responsiveness shown by the intermediary to the users is positively related to the behaviour intention for using CSCs.***

Reliability of services provided by intermediary: It is defined as *"the extent to which reliable services are provided by the intermediary to the users"* and this leads to hypothesis **H5: *Reliability of services provided by intermediary is positively related to the behaviour intention for using CSCs.***

Physical facilities at intermediary outlets: It is defined as *"the extent to which the equipments and ambience at the CSC are perceived to be attractive by the users"* and this leads to hypothesis **H6: *Physical facilities at intermediary outlet are positively related to the behaviour intention for using CSCs.***

We posit that H4, H5 and H6 are captured by dimensions of service quality, namely, tangibles, reliability, responsiveness, assurance and empathy that are proposed in the SERVQUAL scale. Dimensions of empathy and responsiveness can measure perception of users about behaviour of VLE while timely provision of service without uncertainty is captured by dimensions of reliability and assurance in the SERVQUAL scale. Similarly, extent to which equipment and ambience are perceived to be attractive by the users is captured by the dimension of tabgibles in SERQUAL. In accordance to the findings in prior research, it is felt that SERVPERF scale will be appropriate in the present context where the respondents are not much educated.

Age, sex, income and education: Based of prior literature, it is proposed that the demographic factors such as age, sex, income and education will have moderating relation with ease of use, usefulness of CSC and social influence.

5 Conclusions and implications for research and practice

In the paper, we have presented a conceptual model for identifying the determinants for adoption of common service centres for delivery of e-Government services in rural areas. The conceptual model presented in Figure 1 takes into account the unique e-Government delivery context where technology is manifested to the citizens by success of the transactions at the CSC and the intermediary plays a crucial role in determining the quality of service. The study extends the body of knowledge in the domain of

technology adoption by providing a framework for assessing intent to use a different kind of delivery channel of e-Government services, namely the common service centres. Apart from the CSC, there are other innovative upcoming technology-based channels such as SMS and mobile apps on smart phones which are expected to dominate the service delivery in future. The research process by which the model in this paper has been developed can be replicated for further research in adoption of services delivered through these new delivery channels.

References

[1] Abramson, M. A., & Means, G. (2001). E-government 2001, The Price water house Coopers endowment series on the business of government.

[2] Schware, R. (2000). Information technology and public sector management in developing countries: present status and future prospects. Indian Journal of Public Administration, 46(3), 411-16.

[3] Imran, A., & Gregor, S. (2007). A comparative analysis of strategies for egovernment in developing countries. Journal of Business Systems, Governance and Ethics, 2(3), 89-99.

[4] Rana, N. P., Dwivedi, Y. K., & Williams, M. D. (2013). A meta-analysis of existing research on citizen adoption of e-government. Information Systems Frontiers, 1-17.

[5] Sahu, G. P., & Gupta, M. P. (2007). Users' acceptance of e-government: A study of Indian central excise. International Journal of Electronic Government Research (IJEGR), 3(3), 1-21.

[6] Ojha, A., Sahu, G. P., & Gupta, M. P. (2009). Antecedents of paperless income tax filing by young professionals in India: An exploratory study. Transforming Government: People, Process, and Policy, 3(1), 65–90.

[7] Al-Sobhi, F., Weerakkody, V., & El-Haddadeh, R. (2011). The Relative Importance of Intermediaries in eGovernment Adoption: A study of Saudi Arabia. In M. Janssen, H.J. Scholl, M.A. Wimmer, and Y-H. Tan (Eds.): EGOV 2011, LNCS 6846, 62–74."

[8] Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. Educational researcher, 33(7), 14-26.

[9] Mingers, J. (2001). Combining IS research methods: towards a pluralist methodology. Information systems research, 12(3), 240-259.

[10] Lin, F., Fofanah, S. S., & Liang, D. (2011). Assessing citizen adoption of e-Government initiatives in Gambia: A validation of the technology acceptance model in information systems success. Government Information Quarterly, 28(2), 271-279.

[11] Ajzen, I. (1985), ``From intentions to actions: a theory of planned behaviour'', in Kuhl, J. and Beckmann, J. (Eds), Action-Control:From Cognition to Behaviour, Springer-Verlag, Heidelberg, pp. 11-39.

[12] Davis F. A., Perceived usefulness perceived ease of use and user acceptance of information technology, MIS Quarterly 8 (1989) 318–339.

[13] Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. Information systems research, 2(3), 192-222.

[14] Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS quarterly, 27(3).

[15] Sarkar, J. (1998). Technological diffusion: alternative theories and historical evidence. Journal of economic surveys, 12(2), 131-176.

[16] Chuttur M.Y. (2009). "Overview of the Technology Acceptance Model: Origins, Developments and Future Directions," Indiana University, USA . Sprouts: Working Papers on Information Systems, 9(37). http://sprouts.aisnet.org/9-37

[17] Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A conceptual model of service quality and its implications for future research. the Journal of Marketing, 41-50.

[18] Costa, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. Personality and individual differences, 13(6), 653-665.

[19] Latour, B. (2005). Reassembling the social-an introduction to actor-network-theory. Reassembling the Social-An Introduction to Actor-Network-Theory, by Bruno Latour, pp. 316. Foreword by Bruno Latour. Oxford University Press, Sep 2005. ISBN-10: 0199256047. ISBN-13: 9780199256044, 1.

[20] McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: an integrative typology. Information systems research, 13(3), 334-359.

[21] Barua, M. (2012). E-governance adoption in government organization of India. International Journal of Managing Public Sector Information and Communication Technologies, 3(1), 1-20.

[22] Carter, L., Shaupp, L. C., Hobbs, J., & Campbell, R. (2011). The role of security and trust in the adoption of online tax filing. Transforming Government: People, Process and Policy, 5(4), 303-318.

[23] Carter, L., Shaupp, L. C., Hobbs, J., & Campbell, R. (2011). The role of security and trust in the adoption of online tax filing. Transforming Government: People, Process and Policy, 5(4), 303-318.

[24] Venkatesh, V., Sykes, T. A., & Venkatraman, S. (2014). Understanding e‑Government portal use in rural India: role of demographic and personality characteristics. Information Systems Journal, 24(3), 249-269.

[25] Rokhman, A. (2011). e-Government adoption in developing countries; the case of Indonesia. Journal of Emerging Trends in Computing and Information Sciences, 2(5), 228-236.

[26] Floropoulos, J., Spathis, C., Halvatzis, D., & Tsipouridou, M. (2010). Measuring the success of the Greek Taxation information system. International Journal of Information Management, 30(1), 47–56

[27] Hung, S. Y., Tang, K. Z., Chang, C. M., & Ke, C. D. (2009). User acceptance of intergovernmental services: An example of electronic document management system. Government Information Quarterly, 26(2), 387-397.

[28] Chee-Wee, T., Benbasat, I., & Cenfetelli, R. T. (2008, January). Building citizen trust towards e-government services: do high quality websites matter?. In Hawaii International Conference on System Sciences, Proceedings of the 41st Annual (pp. 217-217). IEEE.

[29] Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. Field methods, 18(1), 59-82.

[30] Mason, M. (2010, August). Sample size and saturation in PhD studies using qualitative interviews. In Forum Qualitative Sozialforschung/Forum: Qualitative Social Research (Vol. 11, No. 3).

[31] Eveland, J. D. (1986). Diffusion, technology transfer, and implementation thinking and talking about change. Science Communication, 8(2), 303-322.

[32] Thompson, E. P. (1967). Time, work-discipline, and industrial capitalism. Past and present, 56-97.

Towards a Business Model for SMS-Based Government Services

LANZA, Beatriz B. B. [a], GIL-GARCIA, J. Ramon [b], GIMENEZ, Fernando A. P. [c]

[a] *Universidade Federal do Paraná UFPR,Companhia TIC Paraná CELEPAR, Curitiba, Brazil. Center for Technology in Government at UAlbany, USA* bialanza@pr.gov

[b] *University at Albany, State University of New York, USA. Centro de Investigación y Docencia Económicas, Mexico* jgil-garcia@ctg.albany.edu

[c] *Universidade Federal Paraná UFPR, Curitiba, Brazil* gimenez@ufpr.br

Abstract. Mobile government could be conceptualized as the use of diverse mobile technologies to provide information and services to citizens, businesses, and other stakeholders. Governments around the world, particularly in developing countries are finding mobile strategies to be useful ways to reach to their citizens, including disadvantaged groups. In fact, individuals are increasingly using mobile phones for accessing information and this trend is becoming stronger with the improvement of long-range technologies and the high penetration of mobile devices in developing countries. However, their use to deliver government services to citizens remains generally low. Since smart devices are not as widely spread as more traditional mobile phones, it is necessary to think about commonly available and low cost technologies such as Short Message Service (SMS) as a useful alternative for governments around the world. Therefore, based on some foundational concepts from configuration theory, this paper proposes a conceptual business model for SMS-based government services. We think that this model not only contributes to theory, but could also provide guidance to public managers for them to decide when and how to implement SMS-based government services.

Keywords. SMS-based Government Services, Electronic Government, Mobile Government, Short Message Service, Business Model, Configuration Theory

1. Introduction

Government-delivered services with universal access through internet are essential requirements for electronic government (eGov) [7]. Mobile phones have become the most rapidly adopted technology in history and the most popular and widespread personal technology in the world [25]. Mobile government (mGov) is important in many developing countries because of the extensive use of mobile devices by citizens and this puts great pressure upon its implementation [16]. Just to cite an example, in Brazil there are 37 percent more cell phones than inhabitants [29], but only 31 percent of the phones have internet access. Forty one percent of the population use traditional equipment and not smartphones. The most important use is to receive and send SMS. This service corresponds to 66 percent of mobile phone use in Brazil [6]. In 2009, over 70 percent of world's population used SMS and Voice technologies showing that other countries have similar situations [14].

Government agencies use technology to offer services and information and to contact people, but mostly through old and sometimes obsolete bureaucratic structures [4]. While there are many worldwide success stories with the use of SMS, this service re-

mains incipient when compared to the increase in number of mobile phones in many countries [19]. Governments and researchers apparently ignore the potential value generated by the use of the simple and low cost SMS technology. Few have looked at mobile phones under the business and information system perspective [27], with the mobility business model (BM) receiving very little attention [2]. There is a need to examine issues related to mGov from different perspectives. It seems that mGov may substantially influence the generation of emergent complex strategies and tools for eGov [16].

Interdependent imperatives like environment, structure, leadership, strategy [21] and sustainability [12] help to shape organizational settings when considered holistically [20]. Patterns and generalized forms define the components related to each imperative that are combined into gestalts or configurations [21], The absence of an appropriate BM be behind the low use of SMS by governments [18], but its use may be long lived and extensive providing all mobile devices embrace it in all government levels. The aim of this paper is to propose a conceptual BM for SMS-based government services. It is based on a configurational approach to organizational design [21] [22] [23] and recent developments on business model [26].

The paper is organized in five sections, including the foregoing introduction. The second section describes the different approaches to studying e-government, m-government, and SMS. The third section explains some important concepts of Configuration Theory and their relationships with the generation of a Business Model. Section four proposes and discusses a preliminary BM Structure for SMS-based government services. Finally, section five provides some final comments and suggests areas for future research about this topic.

2. Electronic Government, Mobile Government, and SMS

Electronic government aims to provide citizens and businesses a better access to government information and services, improve the quality of services, reduce costs, and create opportunities for participation [11]. So, eGov could play a very important role in this new governance model, with the participation of a wide network of actors [8]. eGov could be seen as the supply side and the digital divide could be seen as the demand side of the same socio-technical phenomenon [13]. If government pays attention to the technology used by citizens in their daily lives, a small part of the complexity of eGov may be resolved or at least better understood. However, mGov requires more than technical feasibility for enterprise-wide adoption in government. It requires the ability to discuss issues related to economic, social and, especially, political aspects [8].

2.1 Mobile Government

This paper uses mGov as government-delivered services for citizens on mobile phones. mGov is an essential platform of communication between government and citizens. The mobile services value chain includes the government (1), which is expected to lead in developing and improving services. It also includes the (2) hardware industry, software and solutions; (3) telecommunications companies, with infrastructure and services; (4) brokers (aggregating companies specialized services in integrating networks and billing) and finally at the end, (5) the citizen [10] [18]. The value chain integration can be further complicated by the lack of clarity of its actors and their roles [9].

Users need mobile government services to be accessible and delivered anywhere and anytime. It will cause the government activities to achieve greater mobile base more conveniently [16]. Mobile devices are a common part of everyday life for most citizens, so there is great potential in the use of mobile phones as a government service delivery channel. It is important to understand that government efforts to use digital media can be unfruitful if they do not consider citizens' access and the necessary skills to meaningfully use of some of these information and services [8]. So, one of the most basic technologies embedded in all cell phones could be the most efficient way of government to communicate with citizens and provide them with information and services [20].

2.2 A Short Message Service and Government

Since 2001 SMS is being successfully exploited by private companies worldwide. Although encryption of SMS messages is relatively safe, once sent, it is almost impossible to see or modify the message content; ii) accessibility: the success of mGov will depend largely on the number of citizens using it and it is known that socioeconomic factors such as income, education level, gender, age, disability, language and regional differences affect the attitude of citizens towards mGov [16].

Any country with a moderate to high level of mobile phone coverage ought to be able to benefit from the SMS use [1]. The SMS is one of the highly used and well-tried mobile services with global availability within all GSM/CDMA networks (Global System for Mobile Communications/Code Division Multiple Access), limited to the transmission of secure plain text between subscribers and server for different purpose [15].

However, this technology does not replace, and cannot replace, the internet or any other form of government relationship, but it could complement the services already offered in person or electronically [16]. Thus, SMS could become an alternative channel for governments services, which can be used to expand the supply of new or improved services, particularly those already available electronically. There is reasonable agreement about the possibilities of use of SMS technology to optimize the delivery process of combined services with other practices of eGov.

While there are many success stories about government-delivered services through the use of SMS, the proportion is still far from the increase of mobile phones. This proves to be a critical factor which prevents all three government levels to fully incorporate delivery services through SMS. While there are many success stories about government-delivered services through the use of SMS, the proportion is still far from the increase of mobile phones. This proves to be a critical factor which prevents all three government levels to fully incorporate delivery services through SMS. Through the lens of Configuration Theory we will try to identify the relationships present in the value chain of mGov and investigate the existence of possible relationships between variables and the perceptions of power involved in this process.

3. Configuration Theory and Business Models

This section briefly presents some basic concepts from configuration theory and business model, which are later used to propose a conceptual business model for SMS-Based Government services.

3.1 Configuration Theory

Configuration Theory (CT) proposes that the environment, structure, leadership, strategy, and sustainability are he imperatives that help shape organizational settings [21] [12]. For the development of a configurational perspective in organization studies, Miller argued that these variables [21]: (1) were widely accepted paradigms in organizational analysis; (2) ad solid empirical demonstrations; and (3) had already played a central role in the generation of organizational settings repeatedly.

Mintzberg [24] also adds that this perspective in organizations writes its history based on periods of stability and transformation. According to Pereira and Toni [28], due to the great changes taking place, society suffers disruptions in its technological and social systems, leading to most changeable and uncertain environment that require diverse combinations of elements of an organization. For Gimenez [12] this approach indicates the existence of multiple ways to achieve organizational success, but there is a limited number of viable patterns of behavior that can be followed by organizations to accomplish their goals.

The complexity of the organizational governance environment requires that each of its characteristics is studied and understood within a continuum, where each topic should be examined in a unique way in its context, and the environmental variables and the organizational reality are considered. There are various gears and levers in a decision-making structure, large acts are the consequences of innumerable and often conflicting smaller actions by individuals at various levels of government [3]. Therefore, the multidimensional rationality proposed by CT seems to be highly adequate to support the BM building process for mGov. It can include proper contexts and result in an actionable model [27].

3.2 Business Models

In 2008, the BM concept was first used for mGov [10]. However, those were generic references and no detailed models were discussed. Diniz and Cunha [10] argue that the success of mobility in government depends largely on the existence of applications focused on the needs of the citizen and the massive investment in services provision to citizens via mobile devices. On the other hand, these authors also point out that the lack of preparation for this type of service could significantly hinder efforts of government organization.

mGov could be an emerging channel to create new spaces and participation of society, involving the development of innovative models. However, there is no structural BM to work with the most basic technology embedded in any type of mobile phone [17].   There are still barriers to current BM change in the organizations. In order to change this situation it is necessary to deploy new tools, changing organizational processes and adopt a new business posture [5]. It is also necessary to establish a BM that explicitly attempts to complement state bureaucracy in the provision of public services and supports government action [18], because the relationship is much more complicated in government than in the private sector.

For designing the BM for SMS-based government services the vision of Osterwalder and Pigneur [26] is adopted. Their proposal shall be implemented through the organizational structuring of processes and systems, including nine components. Each of these components is briefly describe in the following section. This was chosen because its components come, in a simple way, to present the five imperatives proposed in the CT, that help shape organizational setting. The literature shows that this model

has been used more by the private sector, but taking more or less emphasis, all components can and should be the agencies of government, without exception.

4. Proposing a preliminary BM Structure

This section presents how the adoption of two modes of reasoning from TC – synthesis and analysis – may be applied in the development of a BM for SMS-based government services. This paper presents an ongoing research project, which will be based in multiple case studies using quantitative and qualitative evidence. In addition, in terms of data collection we will use official government documents, general media, interviews, and questionnaires.

4.1 Using synthesis to understand mGov cases

First, in this proposal the synthesis lens is used to understand holistically the mGov cases that have occurred in the past and they are in the present. This way proposed by Miller [21] is adequate to deepen understanding of the historical relationships and sequence among the five imperatives, how and why this imperatives might change during the course of the projects's lifecycle, especially in government agencies where there are constant transitions or top leadership changes. The lifecycle concept suggests an ordering that describes some common transitions among imperatives [20]. Then, for each case its lifecycle is designed and compared with the patterns and configurations suggested by Miller and Gimenez, to verify common transitions among imperatives over lifecycle, identify phase of cycle (birth, growth, maturity and revival), factors initiating, and resulting imperatives. The following is a brief description of the first four imperatives advocated by Miller [21] and the last by Gimenez [12].
   Environment - the proponents of the environmental imperative argue that organizations must adapt to their environments. The uncertainty and unpredictability government policies the nature and degree of competition from stakeholders, and the fast changes of technologies create important constraints that determine which modes of structure and strategy are viable. **Structure** - the formal structure of organizations, as the hierarchy, formal rules and procedures, and managerial authority could influence strategy and decision making and these structures can be implications in performance of organizations. **Leadership** - this imperative shows the influence that chief executive officer's behavior can have on their organizations, how significant an impact the leader can have on the strategies, structure, and culture of an organization. **Strategy** - this imperative defends that the organization has the power to seek out the best niches of the market and the strategic choice operates as managers determine product scope, and consistent with market needs. **Sustainability** - this imperative advocated a balance between economic, social and environmental interests. Focuses on partnerships, cooperation, Coopetition and participative management, with recursive interaction between organization and environment reflected in the structure and strategy [12].
   Second, the analysis is used to measure the components present in the cases using the components suggested [26]. For example, investigate whether each of the nine components, previously cited, are present and how they are related among them and to the specific contexts of the cases.
   Key Activities - for this components we suggest three categories: (1) production - system development, management and delivery of SMS as a whole; (2) production of resolution - consulting, knowledge management, updating and technical training; (3)

platform / network - networks of services provided by operators, platforms and more of these platforms of brokers, government IT companies, software. **Partners** - there are several types of partnerships, including the strategic alliances between competitors and non-competitors; ventures. In this component should also ensure that the government can have control of the content of this information. **Key Resources** - this component describes the most important physical, human and financial resources for the BM operation, for the government to create and offer your Value Proposition, reach the target audience, keep sustainable relationships with citizens. **Cost Structure** - this cost structure would be based on low-cost structures. It should contain: fixed cost (structure, system); variable cost (SMS amount, new system implementations); economy of scale (purchase SMS). Also it is must reflect how the costs would be shared if it were centralized corporate purchasing SMS in government. **Customer Relationships** - this describes the types of relationships that the government provides to citizens, explains the type of relationship whether personal or automated. **Customer Segments**- this component defines the different groups of people or segments that government should serve and which to ignore. This component is possible to exhibit clearly the target audience is those who own a mobile phone. **Value Propositions** - this component describes the services that create value specifically for the needs of people segment you want to achieve. May be similar to existing ones, but with additional features and attributes. The values considered may be quantitative (price, service speed) or qualitative (citizens' experience with your device, efficiency, accessibility). **Channels** - this seems to be simpler than the model for the private sector because the public and the delivery channel are very specific, well-defined. But it is necessary to pay attention that citizens need to adhere to the service, for that there should be specific legislation. **Revenue Streams** - the public sector is constantly challenged to implement the principles of private enterprise and generally BM is referred to something that it is financial gain. It presents some questions: i) Only the government should pay the costs in the generation of these new services? ii) the SMS providers should not pay part of the costs, since they will increase their profits when the government actually use SMS corporately? The Figure 1 shows how is has been used the TC e BM approaches to propose BM for SMS-Based Government Services.
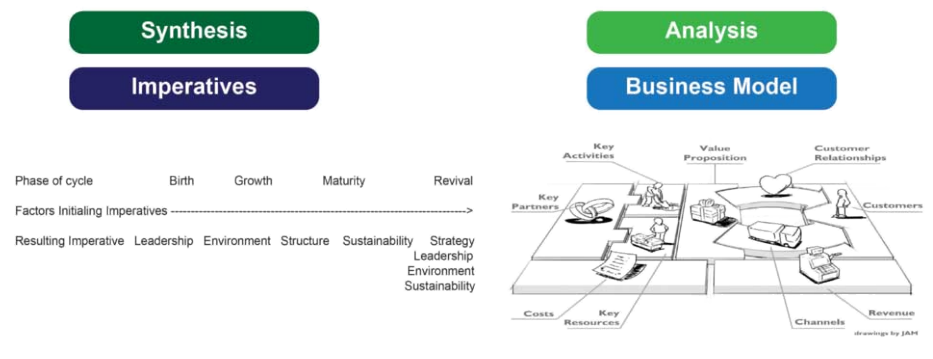
Figure 1 - Configuration Theory and Business Model
Source: adapted from Miller (1987); Osterwalder&Pigneur (2011)

5. Final Comments

This study helps to identify in what conditions traditional mobile phones and the use of SMS can help government reach to their citizens, including disadvantaged groups. It also identifies other reasons why the technology embedded in all cell phones is not used enterprise-wide by government, such as (1) that lack of a clear and actionable BM for this technology; (2) the lack of clarity of the actors and the roles in the mGov supply chain; and (3) the lack of a better use policy and price exclusive to government. More specifically, this paper proposes a conceptual BM for SMS-based government services. It proposes the use of the five elements of configuration theory (environment, structure, leadership, strategy and sustainability) and the BM approach to structure the processes associated with the use of SMS in government-delivered services. In the analyzed literature BM and mGov are treated only as descriptions of how business is done. In other studies, there is only a mention of BM in mGov, but no research was found that presented a proposal of a BM structure for mGov using SMS.

Again, Cunha and Miranda [8] emphatically argue that the use of eGov systems to provide services should consider citizen access and their ability to use those services and information. This is one of the strongest reasons to defend the use of SMS-based government services because all cell phones (traditional and smart) have this technology and it can receive information even if the device holder has no credit (this happens in several countries such as Brazil). This research aims to contribute to the BM and mGov literatures, but also to present recommendations for countries where the conditions are similar to the Latin American reality. We argue that even at an early stage, the proposed approach shows a model potentially useful for governments, especially in developing countries with a clear increase of mobile phones and high penetration among disadvantaged groups. This model, following appropriate methodological procedures, can later be empirically evaluated and contribute to the consolidation of a Business Model for SMS-Based Government Services.

References

1. Aker J.C, Mbiti IM. Mobile phones and economic development in Africa. CGD Working Paper 211. Washington DC: Center for Global Development; Jun 2010. Available from: http://www. cgdev.org/ files/1424175_file_Aker_Mobile_wp211_FINAL.pdf
2. Al-Debei, M. M.; El-Haddadeh, R.; Avison, D. Defining the Business Model in the New World of Digital Business, In Proceedings of the 14th Americas Conference on Information Systems, AMCIS'08, Toronto, Canada, pp. 1-11, 2008.
3. Allison, G. Conceptual models and the cuban missile crisis. The american political science review, vol. LXIII (3:689-718), September, 1969.
4. Bertot, J. C.; Jaeger, P. T.; Hansen, D. The impact of polices on government social media usage: Issues, challenges, and recommendations. Government Information Quarterly, 29(1), 30-40, 2012.
5. Chesbrough, H. W. Business model innovation: Opportunities and barriers. Long Range Planning, 43: 354-363, 2010.
6. CGI. Comitê Gestor de Internet do Brasil. TIC domicílios e empresas 2012 - Pesquisa sobre o uso das Tecnologias de informação e comunicação no Brasil. Available: http://www.cetic.br/publicacoes/2012/ tic-domicilios-2012.pdf [2014, Oct 26]
7. Cunha, M. A.; Annenberg, D.; Agune, R. M. Prestação de serviços públicos eletrônicos ao cidadão. In: KNIGHT, P. T.; FERNANDES, C. C. C.; CUNHA, M. A. (Org.). E-desenvolvimento no Brasil e no mundo: subsídios e Programa e-Brasil. São Caetano do Sul: Yendis, 2007. p. 559-584, 2007.
8. Cunha, M. A.; Miranda, P. O uso de TIC pelos Governos: uma proposta de agenda de pesquisa a partir da produção acadêmica e da prática nacional. Organizações&Sociedade, América do Norte, 20, out. 2013. Available: http://www.portalseer.ufba.br/index.php/revistaoes/article/view/8843/6301 [2014, Mar 6].
9. Diniz, V.; Gregório, A. Do eGov governo eletrônico para o MGov Cidadania Móvel. In: KNIGHT, P. T.; FERNANDES, C. C. C.; CUNHA, M. A. (Org.). E-desenvolvimento no Brasil e no mundo: subsídios e Programa e-Brasil. São Caetano do Sul: Yendis, 2007. p. 688-702, 2007.

10. Diniz, V.; Cunha, M. A. M-Gov ou Cidadania Móvel. In: Foro Iberoamericano sobre Estrategias para Implementar la Carta Iberoamericana de Gobierno Electrónico - Gobierno móvil, CLAD, 2009. Available: http://old.clad.org/documentos/otros-documentos/foro-iberoamericano-sobre-estrategias-para-implementar-la-carta-iberoamericana-de-gobierno-electronico-gobierno-movil/view [2015, Jun 11].
11. Gil-García, J. R.; Pardo, T. A. E-government success factors: Mapping practical tools to theoretical foundations. Government Information Quarterly 22.2: 187-216, 2005.
12. Gimenez, F. A. P. A Sustentabilidade e a Inovação como um novo desafio para a Gestão das Organizações. Tese para o concurso de professor titular do Departamento de Administração Geral e Aplicada, Setor de Ciências Sociais Aplicadas, Universidade Federal do Paraná, 2012.
13. Helbig, N.; Gil-García, J.R.; Ferro R. Understanding the complexity of electronic government: Implications from the digital divide literatureGovernment Information Quarterly, 26 (1) (2009), pp. 89–97.
14. Jahanshahi, A.A.; Khaksar, S. S.; Yaghoobi, N.; Nawaser, K. Comprehensive model of Mobile Government in Iran, Indian Journal of Science and Technology, Vol. 4, No. 9, pp.1188-1197, 2011.
15. Jain, M. K.; Anjana J. Financial Transaction Security Using Mobile SMS. International Journal of u- and e- Service, Science and Technology Vol.8, No.2, pp.365-374 http://dx.doi.org/10.14257/ijunesst.2015.8.2.35, 2015.
16. Kuschu, I.; Kuscu, H. From E-government to M-government: Facing the Inevitable? In the proceeding of European Conference on E-Government (ECEG 2003), Trinity College, Dublin, 2004.
17. Lanza, B. B. B. Dinâmica de relacionamento entre atores de projetos governamentais: o caso do m-Gov do Paraná. Novas Edições Acadêmicas, ISBN13: 978-3639750782, 152p., February 2015.
18. Lanza, B. B. B.; Cunha, M. A. Relations among Actors in Governmental Projects: The Case of Paraná mGov. International Journal of E-Services and Mobile Applications (IJESMA), 5(3), 25-42. DOI: doi10.4018/jesma.2013070102, 2013.
19. Lanza, B. B.B., Gil-Garcia, J. R., Gimenez, F. A. P. (2015, May). The value of low cost technologies: proposing a business model for SMS-based government services. In Proceedings of the 16th Annual International Conference on Digital Government Research (pp. 312-313). ACM.
20. Meyer, A. D.; Tsui, A. S.; Hinnings, C. R. Configurational approaches to organizational analysis. Academy of Management Journal, v. 36, n. 6, p. 1175-1195, 1993.
21. Miller, D. The genesis of configuration. Academy of Management Review, v. 12, n. 4, p. 686-701, 1987.
22. Mintzberger, H. The structuring of organizations (The theory of management policy), USA, 1979.
23. Mintzberger, H. Organization design: fashion or fit? Harvard Business Review, p. 103- 116, jan/febr 1981.
24. Mintzberger, H.; Ahlstrand, B.; Lampel. J. Strategy safari. New York: The Free Press, 1998.
25. OECD. International Telecommunication Union. M-Gov: Mobile Technologies for Responsive Governments and Connected Societies, OECD Publishing. http://dx.doi.org/10.1787/9789264118706-en, 2011.
26. Osterwalder, A.; Pigneur, Y. Business Model Generation - inovação em modelos de negócios: um manual para visionários, inovadores e revolucionários. Alta Books, 2011 p.280, 2011.
27. Pateli, A.; G. Giaglis. A Research Framework for Analysing eBusiness models. European Journal of Information Systems 13: 302-314, 2004.
28. Pereira, B. A. D.; Toni. D. D. A busca da racionalidade multidimensional da organização através da teoria da configuração. Revista Eletrônica de Administração. Porto Alegre: ed. 29 vol. 8, n. 5, set-out, 2002.
29. TELECO. Estatísticas Celulares no Brasil. Available: http://www.teleco.com.br/ncel.asp [2014, Oct 26].

Harmonising the public service models of the Points of Single Contact using the Core Public Service Vocabulary Application Profile

Nikolaos LOUTAS[1], Michiel DE KEUZER[1], Konstantinos TARABANIS[2],
Miguel ALVAREZ-RODRIGUEZ[3], Peter BURIAN[3]

[1]*PwC EU Services, Brussels, Belgium*
{firstname.lastname}@be.pwc.com
[2]*Information Systems Lab, University of Macedonia, Thessaloniki, Greece*
kat@uom.gr
[3]*Directorate-General Informatics, European Commission, Brussels, Belgium*
{firstname.lastname}@ec.europa.eu

Abstract. In the process of implementing the Services Directive, Member States have implemented electronic Points of Single Contact (PSCs) offering public services categorised according to business events. The objective of this paper is to present a commonly agreed data model for describing business events and public services with a particular focus on the electronic PSCs. The paper describes the definition and verification of this data model, called Application Profile of the ISA Core Public Service Vocabulary (CPSV-AP), which includes 16 classes and 85 properties. Moreover, the paper describes the model's mapping to the data models used in the PSCs of 10 MS.

Keywords: public service, business event, service directive, point of single contact, semantic interoperability, CPSV-AP

1. Introduction

In the process of implementing the Services Directive, Member States have implemented electronic Points of Single Contact, in the form of eGovernment portals that allow businesses to:

- Find information about business events and related public services, i.e. the rules to be followed, the prerequisites to be fulfilled, the formalities to be completed and the legislation that is governing a particular business event and its related public services; and
- Execute the public services online (wherever possible).

These electronic PSCs are currently facing several challenges, such as:

- Lack of coordination between the electronic PSCs within the same country and fragmentation of responsibilities;
- Heterogeneous, monolingual, descriptions of public services and business events;
- Administration-centric vs. business centric-approach; and
- National vs. cross-border public service provision.

There is thus a strong need within the MSs for harmonising the way business events and related public services are described. This can be achieved by means of a common data model for representing business events and public services. In our work, such common data model is defined as an Application Profile of the ISA Core Public Service Vocabulary1 (henceforth referred to as the CPSV-AP). An Application Profile2 is a specification that re-uses terms from one or more base standards, adding more specificity by identifying mandatory, recommended and optional elements to be used for a particular application, as well as recommendations for controlled vocabularies to be used.

2. Methodology

Work reported in this paper has been conducted according to the ISA process and methodology[3] for developing Core Vocabularies. The process involved the set-up of a Working Group (WG) and the publication of drafts of the specification for external review. The CPSV-AP was developed under the responsibility of the European Commission's ISA Programme[4] which was also chairing the WG. The WG[5] was responsible for defining the specifications and was established from (part of) the members of the EUGO Network[6] and TIE Cluster representatives. The representatives from 10 MSs (Austria, Estonia, Finland, Greece, Netherlands, Latvia, Lithuania, Poland, Spain and Sweden) participated in this process. Apart from email communication, the WG members participated in 4 webinars.

In practice, the specification of the CPSV-AP followed a four-step process:

Step 1: Following a bottom-up approach, we started by reviewing and analysing the state-of-the-art in the 28 MSs concerning the models being used for describing business events and public services on the electronic PSCs. This analysis led to the documentation of the classes and properties of the model being used in each MS.

Step 2: The participating 10 MSs were then asked to review and validate the analysis of the documented data model for their country.

Step 3: Subsequently, the PSCs' data models for describing business events and public services were compared in order to identify differences and commonalities. As a result, we identified possible new classes and properties for the CPSV-AP. These were suggested as a proposal to the WG. The CPSV-AP also includes a set of recommended controlled vocabularies for different properties with a primary focus on the identification of a common controlled vocabulary for public service types, which can then be linked to key business events. These controlled vocabularies were also subject to the approval of the WG.

Step 4: Furthermore, we proceeded to the creation of a (machine-readable) mapping of each PSC model to the classes, properties and values of the controlled vocabularies of the CPSV-AP, in order to harmonize the national data models and the common data model. The 10 participating MSs were then asked to contribute to and validate these mappings. Finally, a public review period was organised.

---

[1] https://joinup.ec.europa.eu/asset/core_public_service/description
[2] http://dublincore.org/documents/2001/04/12/usageguide/glossary.shtml#A
[3] https://joinup.ec.europa.eu/node/43160
[4] http://ec.europa.eu/isa
[5] https://joinup.ec.europa.eu/node/104619
[6] http://ec.europa.eu/internal_market/eu-go/index_en.htm

3. Related work

All countries nowadays, including the 10 European Member States that were surveyed in the context of this work, have created models for describing public services. A commonly agreed view for public service is that it can be defined as a "*set of deeds and acts performed by or on behalf of a public administration for the benefit of a citizen, a business or another public administration*"[7]. Researchers in the past years, such as the Government Enterprise Architecture [[2]], have also been busy with studying and modelling public services. With the advent of eGovernment, the question on how to organise and present public services and information on public services at government portals emerged. The need for an efficient and citizen-centric way to organise public services on electronic PSCs led to the introduction of the Life Events (LE) term [[3],[4]]. This term has been defined as referring "*to the government services needs at specific stages in life*" [[5]] and as describing "*situations of human beings where public services may be required*" [[6]]. Similarly, the Business Events (BE) term – also called business episodes or business life-events – has been introduced for describing public services referring specifically to businesses [[4]].

Most public administrations have adopted the aforementioned terms and offer their services categorised according to LEs and BEs. This categorisation of public services is also used in the majority of PSCs. Consequently, the commonly agreed data model proposed in this paper intends to document public services in the context of BEs that comprise the business life cycle. Typical examples of such business events are[8]:

- Starting a business, e.g. starting a new activity or company;
- Starting cross-border business e.g. registering a company on abroad ;
- Doing business, e.g. staffing, reporting and notifying authorities;
- Closing a business, e.g. selling the company, closing a branch.

4. The Core Public Service Vocabulary Application Profile

The CPSV-AP (see Figure 1) has 16 classes and 85 properties. To indicate the minimum requirements to comply with the CPSV-AP, the classes and properties are being classified as being mandatory or optional. A minimal implementation of the CPSV-AP at least provides information on the mandatory properties of the mandatory classes. Optional classes can still have properties that are indicated as mandatory, if the particular class is used. 4 mandatory classes are defined, i.e. Business Event, Public Service, Input and Formal Organisation classes, and 27 mandatory properties.

Following, we discuss briefly one by one the classes of CPSV-AP:

Business Event class. This class represents a Business Event. A Business Event is a specific situation or event in the lifecycle of a business, which relates to one or more needs or obligations of that business at this specific point in time. A Business Event requires a set of public services to be delivered in order for the associated business needs or obligations to be fulfilled.

---

[7] e-Government Core Vocabularies, v1.1 https://joinup.ec.europa.eu/asset/core_vocabularies
[8] http://ec.europa.eu/idabc/servlets/Doc2bb8.pdf?id=1675
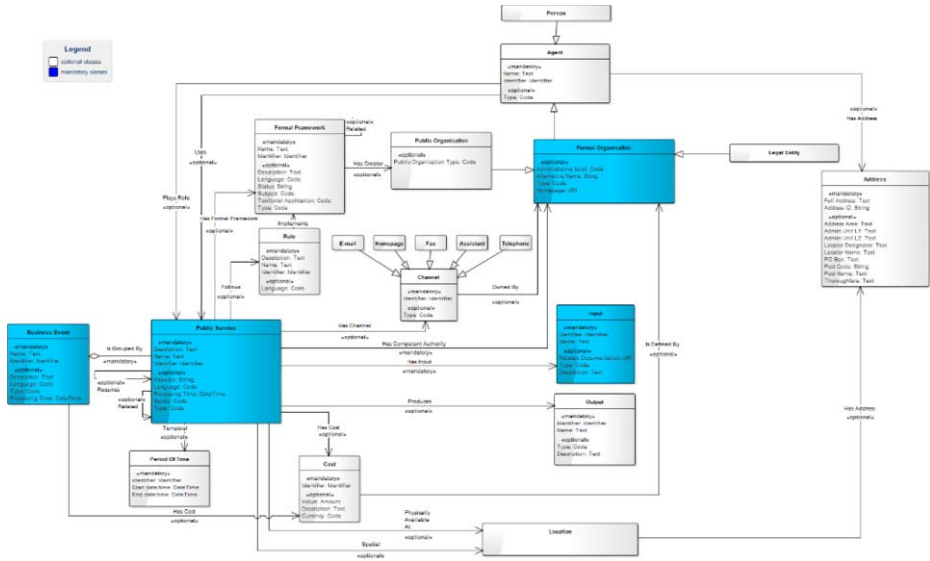
Figure 1. Graphical representation of the relationships between the classes and properties of the CPSV-AP

Public Service class. This class represents the Public Service itself. A Public Service is the capacity to carry out a procedure and exists whether it is used or not. CPSV-AP adopts the definition of the Public Service as defined in Section 3.

Input class. Input can be any resource, e.g. document, artefact. In the context of Public Services, Inputs are usually administrative documents, applications, etc.

Output class. Outputs can be any resource, e.g. document, artefact, produced by a Public Service. In the context of a Public Service, the Output documents an official documentation of the Competent Authority (Formal Organisation) that permits/authorises/entitles an Agent to (do) something.

Cost class. The Cost class represents any costs related to the execution of a Public Service or to all Public Services related to a Business Event which the Agent consuming it, needs to pay.

Channel class. The Channel class represents the medium through which an Agent provides, uses or otherwise interacts with a Public Service.

Period of Time class. The Period of Time class represents an interval of time that is named or defined by its start and end dates. This interval indicates from when to when a Public Service is executable.

Rule class. The Rule class represents a document that sets out the specific rules, guidelines or procedures that the Public Service follows.

Formal Framework class. This class represents the legislation, policies that lie behind the rules that govern the service. The definition and properties of the Formal Framework class in the CPSV-AP are aligned with the ontology included in "Council conclusions inviting the introduction of the European Legislation Identifier (ELI)"[9].

Agent class. The Agent class is any resource that acts or has the power to act. In some countries' legislation the concept Person is a class for anyone that can be legally

---

[9] http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52012XG1026%2801%29

represented and can thus have both Natural Person and a Legal Person (organisation) as subclasses. In the context of this specification a Natural Person is described through the "Person" class, defined in the Core Person Vocabulary[10], and the Legal Person through the "Legal Entity" class, defined in the Core Business Vocabulary[11].

Formal Organisation class. The Formal Organisation class has been defined in the Organization Ontology[12]. It represents an Organisation which is recognized, with associated rights and responsibilities.

Public Organisation class. The Public Organisation class represents a Formal Organisation that is owned by and managed by a state's government.

Person class. The Person class represents a natural person. A natural Person can be the user of a particular Public Service.

Legal Entity class. The Legal Entity class has been defined in the Core Business Vocabulary[13] and represents a business that is legally registered. A Legal Entity is able to trade, is legally liable for its actions, accounts, tax affairs, etc. This makes legal entities distinct from the concept of organisations, groups or sole traders. This vocabulary is concerned solely with registered legal entities and does not attempt to cover all possible trading bodies. A Legal Entity can play different roles related Public Services. The Legal Entity can be a user of a particular Public Service but can also be the Competent Authority of the Public Service.

Location class. The Location class represents an identifiable geographic place. The Address class has been defined in the context of the Core Location Vocabulary[14].

Address class. The Address class represents an address for the representation of a Location. The representation of addresses varies widely from one country's postal system to another. Even within countries, there are almost always examples of addresses that do not conform to the stated national standard.

5. Mapping of the PSC data models to CPSV-AP

The mapping of the PSC data models to the CPSV-AP was based on the review and analysis on the way information is provided on the PSC on the one hand, and the CPSV-AP on the other hand. The objective of the mapping exercise was to promote the harmonization between the business event and public service models used in the 10 MSs and the commonly-agreed European data model, i.e. the CPSV-AP. Each mapping can be of five types (Figure 2), where A is the foreign class or property and B is the class or property from CPSV-AP.

A discussion on the mapping results for each of the participating countries follows:

Austria. Austria has 9 PSCs, while there is also a main Austrian PSC operating as an "umbrella" of the individual PSCs and as a gateway for easily finding each regional PSC. Our analysis included therefore one Austrian data model. The Austrian data model includes 6 classes (i.e. Business Event, Public Service, Input, Formal Framework, Formal Organisation, Location) of the CPSV-AP and 41 properties. All 6 classes have been mapped as "exact match" and 39 of the properties are also mapped as "exact match"; the rest 2 properties are a "related match" and a "close match".

---

[10] https://joinup.ec.europa.eu/asset/core_person/description
[11] https://joinup.ec.europa.eu/asset/core_business/description
[12] http://www.w3.org/TR/vocab-org/
[13] https://joinup.ec.europa.eu/asset/core_business/description
[14] https://joinup.ec.europa.eu/asset/core_location/description
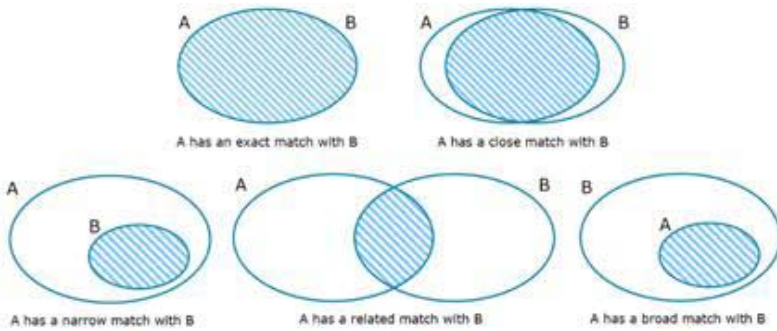
Figure 2. Mapping types

Estonia. The Estonian data model includes 5 classes of the CPSV-AP and 37 properties. 4 classes of the Estonian model are an "exact match" to CPSV-AP classes (Public Service, Input, Formal Framework, Formal Organisation) and 1 class (Business Event) is a "narrow match". More specifically, information and services on the Estonian PSC is categorized according to Topics which is a term with a wider meaning than Business Events. Hence, the "narrow" mapping used between Business Event and Topic. As regards properties' mapping, 32 properties have an "exact match" with CPSV-AP properties, 1 property has a "close match", 1 a "related match" and 3 a "narrow match" due to their direct relation with the Topic class.

Finland. The Finnish data model includes 9 classes (i.e. Business Event, Public Service, Input, Cost, Rule, Channel, Formal Framework, Formal Organisation, Agent) of the CPSV-AP and 80 properties. The Finnish PSC has the Service class which is an "exact match" to Public Service, but it also considers permits as a special subtype of services. The rest 7 classes have an "exact match" with CPSV-AP and 1 class (Agent) has a "close match". As regards the properties, most of them are "exactly" mapped; there are 2 properties with a "close match", 4 with a "related match" and 9 with either a "broad" or a "narrow match" due to the Service/Permit issue described previously.

Greece. The Greek data model includes 4 classes (i.e. Business Event, Public Service, Input, Location) of the CPSV-AP and 14 properties. The Business Event class of the CPSV-AP is only closely matched to the Provision Method of the Greek PSC which defines whether a service refers to business establishment, to cross-border business or both. The rest 3 classes are all an "exact match" to the relevant CPSV-AP classes. As regards properties, 12 are an "exact match" and 2 a "related match".

Netherlands. The concept of Public Service does not appear in the Dutch PSC. For this reason, the Dutch representative has additionally provided the description of the data model used in public administrations in the Netherlands[15]. In total, the Dutch data model includes 4 classes (i.e. Business Event, Public Service, Formal Framework, Address) and 25 properties. 2 of the classes (Public Service, Address) are an "exact match" with CPSV-AP and the rest 2 a "close match". As regards properties, 22 of them are an "exact match", 2 a "close match" and 1 a "related match".

Latvia. The Latvian data model includes 7 classes (i.e. Public Service, Input, Cost, Channel, Rule, Formal Framework, Formal Organisation) and 53 properties. The Latvian PSC is organized to support two types of services: e-Services and Services. Both of these have been mapped with the Public Service class of the CPSV-AP, the former as a

---

[15] https://www.logius.nl/ondersteuning/samenwerkende-catalogi/#c8604

"narrow match" and the latter as an "exact match". The rest classes are an "exact match" apart from the Contact class that is "close match". As regards properties, 30 of them are an "exact match" and the rest 23 properties are either a "close" or a "related match" with CPSV-AP.

Lithuania. The data model includes 6 classes (i.e. Business Event, Public Service, Input, Rule, Formal Framework, Formal Organisation) and 30 properties. The information on the Lithuanian PSC is organized according to Business Events and Permits. For this reason the Permit class is "broadly" matched to Public Service class. All other classes are an "exact match" to CPSV-AP. Regarding properties, 33 are an "exact match", 6 a "close match" and 1 a "related match" with CPSV-AP.

Poland. The Polish data model includes 6 classes (i.e. Business Event, Public Service, Input, Rule, Formal Framework, Formal Organisation) and 62 properties. All classes are an "exact match" apart from the Step (of the procedure) class that is "close match" matched to Rule class of the CPSV-AP. As regards properties, 52 of them are an "exact match", 9 a "close match" and 1 a "related match" with CPSV-AP.

Spain. The Spanish data model includes 11 classes (i.e. Business Event, Public Service, Input, Output, Cost, Period of Time, Formal Framework, Formal Organisation, Public Organisation, Location, Address) and 60 properties. All classes are an "exact match" with the CPSV-AP apart from the Public Service class that is "closely" matched to CPSV-AP. As regards properties, 50 of them are an "exact match", 4 a "close match", 4 a "related match" and 2 a "narrow match" with CPSV-AP.

Sweden. The Swedish data model includes 4 classes (i.e. Business Event, Public Service, Input, Formal Organisation) and 23 properties. The Swedish PSC is built around the concept of Permits which are considered the result of the Public Services. For this reason the Permit class is "broadly" matched to the Public Service class. The rest classes as well as all properties are an "exact match" with CPSV-AP.

Mapping PSC data models with CPSV-AP revealed some interesting results. First, the national data models describing Business Events and Public Services do not cover the whole set of classes or properties defined in the CPSV-AP. Nevertheless, all national data models had additional concepts (mostly properties but sometimes also classes) not defined in the CPSV-AP. This is attributed to the fact that PSCs want to provide detailed information to users, and it cannot be considered as a deficiency of the CPSV-AP

Second, this mapping exercise lead to a few updates of the CPSV-AP that was originally defined. One such important update is the connection of the Public Service and the Formal framework classes. These classes were originally related only through the Rule class. However, the mapping revealed that usually Public Service and Formal framework classes were directly related. Another update is the specification of the Channel class to Email, Homepage, Fax, Telephone and Assistant. The latter was inserted purely as a result of the PSCs' analysis. Finally, the Cost class is also the result of the mapping exercise. Initially, cost was only related to Public Service as a property. However, the PSC analysis showed that cost is often related with Business Event and Formal Organisation classes as well.

6. Conclusions and future work

The CPSV-AP model can be used as the native data model for publishing business event and public service descriptions on a PSC or any other platform containing public service information. Also, it could serve as a common layer for mapping different existing data

models and seamlessly exchanging information on public services across platforms within and across borders. The CPSV-AP should be seen as a generic data model that can be further extended and tailored to the national context or to be suitable for other purposes like Public Service Portfolio management.

The adoption of CPSV-AP largely depends on its implementation in tools for the creation, publication, searching and finding, and exchange (mapping, exporting and harvesting) of these descriptions. In order to facilitate this adoption, future work will focus on the functional specifications of the following reusable tools: (a) a tools for creating machine-readable descriptions of public services; (b) tools for mapping the classes, properties and controlled vocabularies of any data model to CPSV-AP; and (c) tools enabling to harvest public service descriptions, by tagging the content with HTML tags (RDFa) and crawling the pages, or by using a web service that can be called for exporting and harvesting the information.

These tools will stimulate further usage of CPSV-AP towards a more user-centric and harmonised way of publishing information and descriptions of public services and towards easy exchange of this information between different platforms, both within the context of a country and cross-border.

Acknowledgements

References

[1]   SPOCS LSP (2011), Points of Single Contact Research Study, Available at: http://ec.europa.eu/internal_market/services/docs/services-dir/studies/spocs_en.pdf
[2]   Peristeras V. and Tarabanis K. (2008) "The Governance Architecture Framework and Models", chapter in Pallab Saha (eds.) "Advances in Government Enterprise Architecture" Hershey, PA: IGI Global Information Science Reference.
[3]   Tambouris E. and Tarabanis K. (2008), Understanding and Scoping Life Events, International Journal of Electronic Governance, Vol.1, No.2, pp.139-154
[4]   Tambouris E., Tarabanis K., Izdebski W., Momotko M. (2006), "Life Events Revisited: Conceptualization and Representation Using Generic Workflows", Proceedings of the 5th EGOV International Conference, DEXA2006, Krakow
[5]   Pieterson, W., Ebbers, W., van Dijk, J. (2005), The opportunities and barriers of user profiling in the public sector, Lecture Notes in Computer Science, p.3591
[6]   European Commission (2003), The Role of eGovernment for Europe's Future, COM (2003) 567 Final, Brussels, 26 September

---

[16] European Commission. Interoperability for European Public Administrations (ISA). Accessing Member State information resources at European level. http://ec.europa.eu/isa/actions/01-trusted-information-exchange/1-3action_en.htm

PhD Colloquium Papers

This page intentionally left blank

Government Crowdsourcing: the role of trust and community in creating public value

Ann O'BRIEN[a,1]

[a] *Discipline of Business Information Systems, J.E. Cairnes School of Business & Economics, National University of Ireland Galway*

Abstract. Inspired by the e-Participation framework of Saebo, the activities of online political discourse and decision making are studied in the context of crowdsourcing (platforms which enable the wisdom from the crowd) as an underlying technology; and are examined though the public administration theory of creating public value. Due to the shared locus of power, crowdsourcing is situated at the intersection between government led and citizen led e-Participation. Previous research has shown that a high Sense of Community has positive and significant effects on both internal and external efficacy; translated to government crowdsourcing, it would mean that design features that encouraged a Sense of Community in government crowdsourcing would be best placed to create public value by promoting civic engagement, and enabling the creation of a public that could understand and act in their own interests. As citizens are the arbiters of public value, the second phase of the research looks at the outcomes of government crowdsourcing process from the perspective of the citizen. Proposing that there is a hierarchy of participation; that there are different routes to participation; allowing many different actors with different values to deliberate and learn from the process in different ways according to their needs and time constraints. A key contribution is in the novelty of the subjects being investigated – as these two constructs, crowdsourcing and public value have to date not been researched together. The outcomes of the research will guide interactions and connections in government crowdsourcing communities through the technical, legal and social capabilities offered to community members, similar to the way that architecture guides people in physical spaces.

Keywords. Crowdsourcing, Public Value, Trust, Sense of Community

1. Research Aim

Recent research by both Fung and Li have highlighted divergent views from similar research areas of interest; those of political analysts and technologists in e-Participation; both highlight the merits of integrating research [1, 2] to better understand complex problems and lead to more complete solutions. This research aims to incorporate the views of both the public administration perspective using public value as a lens, and the technology perspective with the process of crowdsourcing to explore the effects of design decisions on government crowdsourcing. A key principle of government crowdsourcing

---

[1] a Discipline of Business Information Systems, J.E. Cairnes School of Business & Economics, National University of Ireland Galway

is that of citizen participation, and by seeking the wisdom of the crowd the co-creation of value. Earlier research has concentrated on business crowdsourcing [3], whereas this research examines design decisions of government crowdsourcing to determine the key features of government crowdsourcing that influence the participation of citizens and the achievement of stated goals. Knowledge network effectiveness [4] is used as a basis for a framework of government crowdsourcing, to help to identify where public value can be created from both intrinsic (ends in themselves) and instrumental (means to an end) perspectives.

Trust is a multi-faceted construct which is both a prerequisite [5] to and an outcome of e-participation; trust is also identified as a central component in the achievement of public value [6]. This research brings together and extending two distinct areas of research identified by Bannister and Connolly [7] that of the broader public administration literature who argue that the restoration of citizen trust in traditional government is at the core of public sector modernization. Yet, e-government literature, has emphasized, trust in e-government specifically rather than government in general. Carter and Belanger emphasize that citizens must have confidence in both the government and the enabling technologies; trust is also essential to maintaining ongoing relationships and countering uncertain situations [8]; The construct trust will be examined as a prerequisite of participation, an enabling factor to the creation of public value and an outcome of the crowdsourcing process. By integrating public administration and technologist views outcomes of this research should aid governments in policy making and in developing their crowdsourcing platforms to improve the quality and effectiveness of engaging with citizens.

2. Theory

Crowdsourcing is a term that refers to a diversity of practices and is subject to many definitions. The term originally coined by Jeff Howe, states that crowdsourcing "is a business practice that means literally to outsource an activity to the crowd"[9]. More than that crowdsourcing can be seen as a distributed problem solving and production model and as a public participation tool for governance and planning [10]. Crowdsourcing exists at the intersection between government led and citizen led e-Participation. Recognizing the duality of e-Participation as the integration of government led and spontaneous citizen-led e-Participation [11] raises further questions of power distribution and value recognition. In crowdsourcing the locus of control resides between the organization and the online community [10]. The public value paradigm recognizes citizens as the arbiters of value [12] this opens up a route to agenda setting and decision making for citizens that facilitates the co-creation of value.

The requirements of government crowdsourcing differ considerably from business crowdsourcing in particular equity and education; where balancing the need for broad inclusiveness and informed understanding is a key challenge [13]. In the interest of fairness and inclusiveness it is important to use a broader definition of the digital divide; one that goes beyond access to technology and takes into account motivational access, material access, skills access, and usage access and context of technology adoption [14]. e-Participation is taken to mean the technology-mediated interaction between the civil society sphere and the formal politics sphere and between the civil

society sphere and the administration sphere, with the focal point being the citizen [15]. It is acknowledged that an e-Participation framework needs to recognize both the complexity of the domain, and the need to use methods relevant to all those with an interest in the outcome [16]. While recognizing that knowledge is discursive, contingent, and changeable and that it emerges through interaction [17]. E-Participation projects are undertaken with the aim of increasing transparency in governance, the creation of legitimacy and even revitalizing democracy [18]. Influential frameworks to understand and measure success include; Net Benefits of e-government success from a citizen perspective [19] and e-cognocracy a concept of citizen participation that combines direct or participative democracy and liberal or representative democracy to cognitive ends; along with the EF$_3$ framework it has been used to evaluate the success and impact of e-participation experiences [20].

Government crowdsourcing facilitates citizen participation, also a shared locus of power and multiple stakeholders. It is of strategic importance to identify differing interests among stakeholders and to analyse stakeholders' interests and power bases [21]. So that by mapping e-government benefits against stakeholder groups, useful insights into the benefits that are important across stakeholder groups can be discovered [22]. Equally important is the practice of obtaining stakeholder feedback at all stages of the design process [23].

Public value comes from delivering benefits that extend beyond narrow monetary outcomes to include that which is valued by society generally [24]. Public value research is seen as a perspective to link the individual and society [25]. As new public management have fallen short of their goal; public value can be seen as a new paradigm to change the way ICT enabled reforms are addressed [26]. The government has a special role as guarantor of public values, public management in the broader sense with increased emphasis on collaborative governance, democracy and citizenship [27]. A key tenant of the public value paradigm is enabling citizens to be responsive as partner in the state and take up the duty of active participation [6]. E-Participation provides new forms of civic involvement and this has in turn led to revised concept of volunteerism and citizenship in the US and we-Government in the EU [28] it calls on governments to provide the public with the technology tools that enable citizens to create public value.

2.1 The public value of crowdsourcing and the role of trust

One of the key premises in creating public value [24] is that of individuals coming together to 'form a public' [12, 29]. In this way individuals form groups to enter into dialog and learn from each other in so doing they decide which values are most important, and co-create value [30]. Crowdsourcing enables this dialog and exchange of views [31]. It represents the intersection of technology tools, problem-solving techniques, and the participatory spirit of citizen engagement [32]. The public value paradigm calls on the public themselves to identify what is considered 'valuable' so that citizens are not just beneficiaries of public interventions but also client obligatees [33]. To create a sense of online community; collaboration must take place, a key skill required is trust, the cornerstone of building trust is cooperation [34] and the ability to build relationships with others. To participate in government crowdsourcing a citizen must first trust the site and

by participating will confer legitimacy on the crowdsourcing and its outcomes. While acknowledging that virtual policy communities challenge both the primacy of politicians, and civil servants position of privileged information and knowledge; and the established practices within public administration [28].
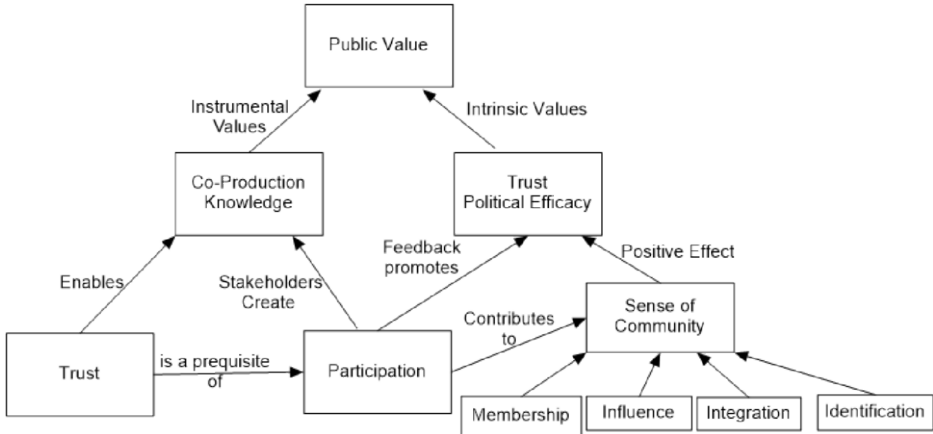

Figure 1. Proposed model of Public Value in Government Crowdsourcing

Figure 1 above shows two different types of public values; instrumental values [35] (means to an end) that are derived from the co-production of knowledge, made possible by multiple stakeholders who trust the process of government crowdsourcing enough to participate. The second strand consists of intrinsic values [36] (ends in themselves) from the normative argument; these are developed alongside instrumental values. Here the shared locus of power and the development of the virtual community of government crowdsourcing could enable a greater Sense of Community. Four main elements contribute to Sense of Community; that of membership, a sense of belonging, influence, when a person believes they can make a difference to their group; integration and fulfilment of needs, values in the group are shared and needs fulfilled and finally shared emotional connection or shared history among members of the community [37]. Political efficacy occurs in two forms: internal efficacy refers to beliefs about an individuals' own ability to influence the political process and external efficacy is concerned with beliefs about the responsiveness of government officials to the concerns of citizens Anderson's research shows that a high Sense of Community has positive and significant effects on both internal and external efficacy and trust [38]; translated to government crowdsourcing, it would mean that design features that promoted a Sense of Community in government crowdsourcing would be best placed to create public value by promoting civic engagement and enabling the creation of publics (or communities) that could understand and act in their own interests [39].

Virtual community plays a central role in the crowdsourcing process; they are the bedrock upon which the wisdom of the crowd is gathered. Virtual communities can be highly prolific producers of social capital [40] and the original four factor structure of Sense of Community proposed by McMillan and Chavis can be validated in virtual

communities [41]. Therefore, the Sense of Community Index is a measurement that opens the constructs of trust and political efficacy to greater scrutiny, which can then guide design decisions of the government crowdsourcing process.

2.2 A hierarchy of e-Participation

For Fogg and Eckles [42] the sequential nature of persuasion on successful web services lead to a pattern of behavioural goals or behaviour chain. These phases are made up of one or more goals called target behaviours. Users are not all expected to contribute in the same way; but they are exposed to a variety of influence strategies that lead them to create value for themselves and others [42]. From discovery through superficial involvement and on to true commitment, these phases also hold true to e-Participation. When viewed through Arnstein's typology of participation or 'rungs of a ladder' from non-participation up to tokenism where citizens' voices can be heard but decisions are made which disregard their voice. On to placation where citizens have an advisory position, and partnership where trade-offs are made between powerful and 'ordinary' citizens. Only on the top of the ladder, the rungs of delegated power and citizen control can citizens have a major role in decision-making processes [43]. This categorisation resonates with the positioning of the locus of control in crowdsourcing and suggests that there might be a hierarchy of e-Participation that incorporates the role of social media, context and change allowing for thick or thin engagement as appropriate. Extending from liking and tagging leading to discovery, on to voting or sharing in superficial involvement and finally creating informed comments or starting new discussions/threads in true commitment. Although, Fung [44] reminds us that sometimes it is more appropriate for members of the public to have a consultative role; also how political theory distinguishes between aggregative and deliberative decision making [45]. This highlights the different routes participation could take; allowing many different actors with different values to deliberate and learn from the process and begin to act more like citizens [46], in their own way, depending on the context.

3. Methodology

The existing theory on creating public value [24] holds as one of its central premises the importance of forming a public as a way that individuals create public value and become arbiters of public value [12]. Crowdsourcing fits well into this concept as it consists of creating a community that seeks knowledge from the 'crowd'. In constructing a participation architecture to act as a model for a comparative study of multiple projects it will be possible to identify the best practices along with problems and weaknesses. By examining a crowdsourcing community as a knowledge network, it is possible to break down the design dimensions that focus on developing, distributing and applying knowledge. The design dimensions of knowledge networks encompass strategic, structural and tactical issues; this influences dynamics and behaviours within the knowledge network, which in their turn affect the outcomes achieved [4]. In a business crowdsourcing context the participation architecture of a technical community is determined not only by its technical architecture, but also by community design decisions made by the community's leaders [3]. Governments must decide to what extent the outcome of crowdsourcing will be implemented into the current policy making

framework or risk alienating their citizens who have participated. Government crowdsourcing can be examined by comparing the application of these dimensions, because they offer a holistic approach for achieving collaboration, network cohesion and broad connectivity and help yield important outcomes of coordination and learning [4]. While allowing Bryson's design science inspired, iterative design guidelines [47] for public process participation to remind us that e-Participation must fit with the broader policy process.

This research will follow a deductive approach using case studies; some will be chosen on the basis that similar results will be expected, this literal replication would lead to using projects with participatory goals, or those from the EU only. A contrasting group will be chosen where the contextual factors are expected to be different, for theoretical replication [48] those that are deliberative in nature, or from the US. The US have the most experience of using crowdsourcing in government and operate to international best practice. Other case studies will be used outside of the US to ensure that; the political, cultural and administrative context is also examined. For example as well as US case studies such as Keep the Web Open and Regulation Room, EU case studies such as completed crowdsourcing projects (Cockpit – Palerno, Tilburg, Athens), Smart City initiatives e.g. (Smartip - Manchester, Cologne, Bologna and Ghent) and the ongoing Route to PA or UK Open Government Manifesto. Increasingly secondary data sources open up new areas of possibility; as demonstrated by a recent study of social media and public value [49].

There will be an in depth examination of the community's website and analysis of literature pertaining to each project followed by semi-structured interviews, that will focus on understanding how sponsors approached the prospect of building community. Interviews will cover the following domains: (1) when and how the crowdsourcing community was founded; (2) how they prepared for and designed the crowdsourcing community; (3) explore their experience in building the community; and (4) what metrics were used to measure the outcomes of the crowdsourcing efforts. Using the proposed model of government crowdsourcing; case studies will be used in phase one to examine design decisions of knowledge management and community building within the crowdsourcing project. Seeking answer the question: how can government crowdsourcing be designed to create public value and trust? Phase two uses a quantitative questionnaire to examine the views of citizens regarding the outcomes of government crowdsourcing. Looking at the context of individual citizens' views by using the Sense of Community Index; along with their level of interaction with the government crowdsourcing site will help to highlight areas where they identify public value has been created.

4. Prospective results

Using the proposed model of government crowdsourcing to examine how government crowdsourcing can be designed to create public value and trust; will lead to an improved understanding of how design influences participation; and help to develop a more explicit causal model [50]. Analyzing case studies using the proposed model will help identify the role of trust and community; the tensions, dynamics and behaviors within government crowdsourcing that affect outcomes and the creation of public value. In phase two, by using the experiences of citizens as the arbiters of public value [12] to

examine the outcomes of government crowdsourcing, a clearer picture will emerge of how different levels of citizen engagement with government crowdsourcing create public value and trust. In this way constraints and enabling factors will be identified using public value as a lens.

References

1. Fung, A., H. Russon Gilman, and J. Shkabatur, *Six Models for the Internet + Politics.* International Studies Review, 2013. 15 (1): p. 30-47.
2. Li, Y., *Think tank 2.0 for deliberative policy analysis.* Policy Sciences, 2015. 48 (1): p. 25-50.
3. West, J. and S. O'Mahony, *The Role of Participation Architecture in Growing Sponsored Open Source Communities.* Industry and Innovation, 2008. 15 (2): p. 145-168.
4. Pugh, K. and L. Prusak, *Designing effective knowledge networks.* MIT Slone Management Review, 2013. 55 (1): p. 2013.
5. Putnam, R.D., *The Prosperous Community.* The American Prospect, 1993. 4 (13): p. 35-42.
6. Stoker, G., *Public Value Management.* American Review of Public Administration, 2006. 36 (1): p. 41-57.
7. Bannister, F. and R. Connolly, *Trust and transformational government: A proposed framework for research.* Government Information Quarterly, 2011. 28 (2): p. 137-147.
8. Belanche, D., et al., *Trust transfer in the continued usage of public e-services.* Information & Management, 2014. 51 (6): p. 627-640.
9. Howe, J., *Crowdsourcing: How the power of the crowd is driving the future of business.* 2008, London: Random House.
10. Brabham, D.C., *Crowdsourcing.* 2013, Cambridge MA: Mit Press.
11. Porwol, L., A. Ojo, and J. Breslin, *On The Duality of E-Participation–Towards a foundation for Citizen-Led Participation*, in *Technology-Enabled Innovation for Democracy, Government and Governance*. 2013, Springer. p. 211-225.
12. Moore, M. and A. Fung, *Calling Publics into Existence: The Political Arts of Public Management*, in *Ports in a Storm: Public Management in a Turbulent World,*, J.C. Donahue and M.H. Moore, Editors. 2012, Brookings Institution Press: Washington, DC. p. 180–97.
13. Epstein, D., M. Newhart, and R. Vernon, *Not by technology alone: The "analog" aspects of online public engagement in policymaking.* Government Information Quarterly, 2014. **31**(2): p. 337-344.
14. van Dijk, J.A.G.M., *Digital divide research, achievements and shortcomings.* Poetics, 2006. **34**(4–5): p. 221-235.
15. Sæbø, Ø., J. Rose, and L.S. Flak, *The shape of eParticipation: Characterizing an emerging research area.* Government Information Quarterly, 2008. 25 (3): p. 400-428.
16. Macintosh, A. and A. Whyte, *Towards an evaluation framework for eParticipation.* Transforming Government: People, Process and Policy, 2008. 2 1): p. 16-30.
17. Chadwick, A., *Bringing E-Democracy Back In Why it Matters for Future Research on E-Governance.* Social Science Computer Review, 2003. 21 (4): p. 443-455.
18. Moreno‑Jiménez, J.M. and W. Polasek, *e‑democracy and knowledge. A multicriteria framework for the new democratic era.* Journal of Multi‑Criteria Decision Analysis, 2003. 12 (2‑3): p. 163-176.
19. Scott, M., W.H. DeLone, and W. Golden, *Understanding Net Benefits: A Citizen-Based Perspective on eGovernment Success*, in *Thirtieth International Conference on Information Systems (ICIS 2009)*. 2009, AIS: Phoenix, Arizona. p. 1-11.
20. Espes, C.P., M.A. Wimmer, and J.M. Moreno-Jimenez. *A Framework for Evaluating the Impact of E-Participation Experiences*. in *Electronic Government and Electronic Participation: Joint Proceedings of Ongoing Research, Posters, Workshop and Projects of IFIP EGOV 2014 and EPart 2014*. 2014. IOS Press.
21. Crosby, B.C. and J.M. Bryson, *Leadership for the common good: Tackling public problems in a shared-power world*. Vol. 264. 2005: John Wiley & Sons.

22. Rowley, J., *e-Government stakeholders—Who are they and what do they want?* International Journal of Information Management, 2011. 31 (1): p. 53-62.
23. Grimsley, M. and A. Meehan, *e-Government information systems: Evaluation-led design for public value and client trust.* European Journal of Information Systems, 2007. 16 (2): p. 134-148.
24. Moore, M., *Creating Public Value - Strategic Management in Government*. 1995, Cambridge, MA: Harvard University Press.
25. Meynhardt, T., *Public Value Inside: What is Public Value Creation?* International Journal of Public Administration, 2009. 32 (3-4): p. 192-219.
26. Cordella, A. and C.M. Bonina, *A public value perspective for ICT enabled public sector reforms: A theoretical reflection.* Government Information Quarterly, 2012. 29 (4): p. 512-520.
27. Denhardt, J.V. and R.B. Denhardt, *The New Public Service: Serving, not Steering*. Third ed. 2011: ME Sharpe
28. Linders, D., *From e-government to we-government: Defining a typology for citizen coproduction in the age of social media.* Government Information Quarterly, 2012. 29 (4): p. 446-454.
29. Dewey, J., *The Public and its Problems*. 1927, New York: Henry Holt.
30. Bovaird, T., *Beyond engagement and participation: User and community coproduction of public services.* Public Administration Review, 2007. 67 (5): p. 846-860.
31. Farina, C.R. and M. Newhart, *Rulemaking 2.0: Understanding and Getting Better Public Participation*. 2013, IBM Center for the Business of Government.
32. Brabham, D.C., *Crowdsourcing as a model for problem solving an introduction and cases.* Convergence: the international journal of research into new media technologies, 2008. 14 (1): p. 75-90.
33. Alford, J. and J. O'Flynn, *Making Sense of Public Value: Concepts, Critiques and Emergent Meanings.* International Journal of Public Administration, 2009. 32 (3-4): p. 171-191.
34. Putnam, R.D., *Bowling alone: America's declining social capital.* Journal of Democracy, 1995. 6 (1): p. 65-78.
35. Nabatchi, T., *Putting the "Public" Back in Public Values Research: Designing Participation to Identify and Respond to Values.* Public Administration Review, 2012. 72 (5): p. 699-708.
36. Nabatchi, T., et al., *Democracy in motion: Evaluating the practice and impact of deliberative civic engagement*. 2012: Oxford University Press.
37. McMillan, D.W. and D.M. Chavis, *Sense of community: A definition and theory.* Journal of Community Psychology, 1986. 14 (1): p. 6-23.
38. Anderson, M.R., *Community Psychology, Political Efficacy, and Trust.* Political Psychology, 2010. 31(1): p. 59-84.
39. Moore, M. and J. Bennington, *Public value: Theory and practice*. 2011, Palgrave Macmillan.
40. Zwass, V., *Co-Creation: Toward a Taxonomy and an Integrated Research Perspective.* International Journal of Electronic Commerce, 2010. 15 (1): p. 11-48.
41. Abfalter, D., M.E. Zaglia, and J. Mueller, *Sense of virtual community: A follow up on its measurement.* Computers in Human Behavior, 2012. 28 (2): p. 400-404.
42. Fogg, B. and D. Eckles, *The behavior chain for online participation: how successful web services structure persuasion*, in *Persuasive Technology*. 2007, Springer. p. 199-209.
43. Arnstein, S.R., *A ladder of citizen participation.* Journal of the American Institute of planners, 1969. 35 (4): p. 216-224.
44. Fung, A., *Varieties of Participation in Complex Governance.* Public Administration Review, 2006. 66 (6): p. 66-75.
45. Cohen, J., *Deliberation and democratic legitimacy*, in *The good polity: normative analysis of the state*, . Editor. 1989, Basil Blackwell: Oxford. p. 67-92.
46. Fung, A., *The Principle of Affected Interests and Inclusion in Democratic Governance.*, in *Representation: Elections and Beyond*, J.H. Nagel and R.M. Smith, Editors. 2013, University of Pennsylvania Press.: Philadelphia.
47. Bryson, J.M., et al., *Designing public participation processes.* Public Administration Review, 2013. 73 (1): p. 23-34.
48. Yin, R.K., *Case Study Research: Design and methods*. 2013: Sage publications.
49. Osmani, M.W., et al. *The Public Value of Social Media in the UK Public Sector*. in *Electronic Government and Electronic Participation: Joint Proceedings of Ongoing Research, Posters, Workshop and Projects of IFIP EGOV 2014 and EPart 2014*. 2014. IOS Press.
50. Fulla, S. and E. Welch. *Framing virtual interactivity between government and citizens: A study of feedback systems in the Chicago Police Department*. in *System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on*. 2002. IEEE.

Integrated Public Service Delivery in Agile e-Government System through Service Oriented Enterprise Model in the Context of India

ANU PAUL[a,1]
[a]*Research Scholar, School of Computer Science*

Abstract. The efficient public administration becomes a universal phenomenon in developed and developing countries. It aims better, faster, convenient, and cost-effective services to citizens, business, and employees. The revolution of ICT with the citizens changing demands enforced e-Government (EG) researchers and practitioners for developing new skills, competencies, capabilities, and knowledge to respond to changes in its service delivery environment. This requirement rooted to achieve an agile EG system (EGS). Since the business environment of enterprises is undergoing rapid changes, researchers are trying to find an appropriate approach toward the creation of an enterprise agility to cope with this vicissitudes. Transforming the traditional enterprise model to Service-Oriented Enterprise Architecture (SOEA) is one of the best state-of-the-art solutions for an organization to get agile. It is the evolution of Enterprise Architecture style and it integrate the business and IT logic to create an agile enterprise system. This research adopted this philosophy and applied in the EGS. The researcher proposed a Service-Oriented e-Government Enterprise Architecture to achieve agile EG system in view of Indian context. Government of India has taken various EG initiatives with high importance given to citizen-centric service delivery, but they are islands of attempts in the absence of interoperability. The proposed approach is capable to solve the interoperability and integration issues of EGS and is illustrated using a public service provided by Government of India.

Keywords. E-Government Development in India, Interoperability and Integration, Agile e-Government System, Service Oriented e-Government Enterprise Architecture, SCA, DSM.

1. Introduction

The e-Government (EG) is defined as the application of electronic means in transactional exchanges of government services and information between government, employee, citizen, and business [25]. The focal aspire of EG is to offers better, faster, convenient, and cost-effective services to citizens, business, and employees in a single window [3]. This necessitates the EG become a seamless government or matured EG. It is evolving through four different maturity stages emergence, interactive, transactional, and then connected or integrated. I.e., it is a transformation from disseminate the digitized

---

[1] *Corresponding Author: Mahatma Gandhi University, Kottayam, Kerala, India, anupaul71@gmail.com.*

information of each government agencies to the integration these information silos. The final stage requires Integration (INT) of governmental departments across the country and it is apparent with the EG Interoperability (EGIOP). The entities of EG System (EGS) work together and share information for the business motive of citizen-centered service delivery. Accordingly, the citizenry act as service consumers and multiple government agencies are the service providers [16]. The national portal is acting as a platform for provision and accession of public services [25]. Hence, EG is an enterprise and its logical structure is the EG Enterprise Architecture (EA) [21][24]. Universally, many countries have viewed EA as a solution for EGIOP [2][12][21].

The revolution of ICT with the citizens changing demands enforced EG to develop new capabilities to respond to the changes in its service delivery environment. Universally it is a big challenge for governmental organizations to keep abreast with citizens changing demands. Information is scattered across organizations and leads to conflicts of value, goal, and culture exist between partners. Since EG is an automated enterprise with IT oriented multifaceted service-centric system, EA confines the integration requirements of software components [26]. Service-Oriented Architecture (SOA) manifestation led some governments to pertain this in their EGs, but it limits business artifacts of an enterprise [9]. The concurrent use of EA and SOA in EGS is most beneficial for integration of business requirements [9][26]. As an automated enterprise logic consists of Business Processes (BP) and application logic, EGS should comprises of centralized public service processes with the existing application logic that belong to various departments at central, state and local level. This is pragmatic only when the EG is viewed as a Service Oriented Enterprise (SOE). This concept leads to an agile EGS resulted the efficiency, re-usability, non-redundancy, citizen friendliness, rapid service delivery, and citizen's active participation in governance process [16]. This is an evolutionary movement than a revolutionary [5].

The agility of EGS is defined as the quality that determines its responsiveness, competency, quickness, and flexibility [1][27]. These capabilities enable EGS to respond proactive and innovative changes of the citizens needs on a timely basis. It is the provision to permit adding or removing independent systems from the integrated structure without major modifications that increases the time-to-market of new services. The EGS agility is based on the flexibility, adaptability, reactivity, robustness, and integration [27]. Information sharing and integration is one of the crucial approach for increasing organizational efficiency and performance [29]. The challenges are encountered when achieving cross-boundary information sharing and integration and it increases the complexity of Enterprise System (ES). A centralized SOE architectural style is one of the state-of-the-art tool for enterprise information system integration [10][11][29]. Since EG is viewed as an enterprise, the agility in EGS can be achieved in the same way as the ES progresses towards service-oriented direction [27]. This research studies how cross-boundary information sharing and integration are carried out in EGS using the service-oriented mechanism. It explores how EGIOP&INT issues are solved in view of an agile EGS through a Service-Oriented e-Government Enterprise Architecture Framework (SOeGEAF) [11] in Indian context and then modeled public service integration in a more pragmatic way. The researcher progressed with a practical solution for the proposed framework and models by the Service Component Architecture (SCA) standards which is a programming model for service creation and implementation [4][30], and the Dependency Structure Matrix (DSM) methodology [28][8] to integrate the service components (SCs) from disparate platforms. Thus the public services within SOeGEA would be instantiated by the SCs in a frame of SCA.

2. Research Background

2.1. Context and motivation of Research

Globally, the EG Development (EGD) is similar in countries, but in India it is more complex due to the multi-tire administrative structure, diversity of culture, and different process management methods in various governmental departments. There are several EG initiatives across the country at the center, state and grass-root level. The diverse governmental agencies are developing their own systems to provide better services to their stakeholders, without knowing other department's systems and also without considering the integration. They are islands of attempts and many of them providing excellent services, but some are still using legacy systems as manual process. These application silos caused multiple similar software systems with duplication of data and efforts. The ventures of EG departments are designed, developed, and implemented independently by the line departments. This caused the duplication of design artifacts, application assets, data archives, and human efforts which increases the secure storage requirements of confidential data and its density, redundancy, cost due to man hours, and deliberate responses due to slow decision making process. In the existing context, the citizen need to interact with each governmental departments to get their public services. The government processes are inter-linked by the multiple government agencies, and poor performance in service delivery due to deprived coordination among these agencies. Because of its inherent organizational fragmentation, citizen have several obstacles to access single window e-services. This context leads to our research problem.

2.2. Research Problem

The main objective of this research is to develop a solution for EGIOP in order to get integrated public services in Indian context using the state-of-the-art paradigms. To reach the research objective we present the following main research question, "How to provide single-window public services to citizenry in an agile e-Government system?" Single-Window access is meant for the interactions between government and citizens done 24X7 from anywhere through a national portal of centralized EGS. Agile EGS meant for the efficient, reusable, non-redundant, integrated, and rapid service delivery with citizen friendliness. The following sub-questions help us to answer the main research question.

- RQ1 - What are the existing e-Government Development scenarios in India?
- RQ2 - What are the possibilities of EA and SOA adoption in Indian EGS?
- RQ3 - How to solve the Interoperability and Integration issues of EGS in India?
- RQ4 - How to integrate the public services in the agile EGS perspective?
- RQ5 - How to manage the service dependency of service components?

Single-window public services provision in an agile EGS is the most innovative stage of EG while traditional government transformed to EG through different maturity stages. To attain a solution of the research question, the researcher has done a profound exploration by the five aforementioned sub-questions for the existing context, up-to-date developments, and current stage of EG in RQ1; then the need and possibility of the proposed elucidation by RQ2; how to achieve matured EGS by existing developments in RQ3; options of advanced tools to get agile EGS and also to manage the service dependency by RQ4 and RQ5.

2.3. Research Approach

The objective of the research were achieved through the Design Science Research (DSR) methodology. The research process begin with the review of the literatures which provides context for the importance and meaning of the research topic and questions. The DSR method is concerned with the construction and evaluation of an IT artefact [7]. Hevner et al. list guidelines that DSR should follow, it is presented in Table 1. In addition to the literature review, the researcher accomplished three phases such as analysis, framework formulation, and prototyping.

Table 1. Application of design science guidelines in this research.

Guideline name and description	Meeting the guideline in this research
Guideline 1: Design as an Artifact	The produced artifact such as framework and models are fruitful for EGIOP&INT issues in view of an agile EGS of India.
Guideline 2: Problem Relevance	EGIOP&INT is one of the most essential matter in research and also in practice to the evolution of EG maturity. Since the findings and conclusions are based on Indian context, it is applicable for policy makers globally to guide them towards taking correct strategic decisions towards a mature EGS.
Guideline 3: Design Evaluation	The constructed artifacts were evaluated by prototyping a public service in EG domain.
Guideline 4: Research Contributions	The research has sound contributions and the subsequent section conferring them.
Guideline 5: Research Rigor	EG as a multidisciplinary area of research rises the volume of literature that hasten the convolution of literature survey. Prototyping is comparative when modelling in view of transformation of an automated enterprise into the SOE.
Guideline 6: Design as a Search Process	Due to the extensive research process, the resultant artifacts are completed in iterations. Each iteration results in the inclusion of new practical concept and is one step closer to finding solution for EGIOP&INT issues, as well as reuse of public service processes by managing its dependency.
Guideline 7: Communication of Research	The framework and models in view of EG in India has been presented in the National Conference on e-Governance (NCeG) in India which is a platform for policy makers, practitioners, industry leaders and academicians. The other perceptions also has been communicated to the appropriate audience.

3. Related Work

Many countries have realized the need to develop a national EGS to deliver integrated public services as a part of EG transformation. This is concerned with the solutions for EGIOP&INT issues which have been studied for years (Scholl, H., 2007; Scholl & Klischewski, 2007; Janssen, M., 2007; Guijarro, L., 2007; Lallana, E. C., 2007; Estevez, E., 2007; Kubicek, H., 2008; Gottschalk, P., 2009; IDABC, 2010; Saha, P., 2010; Chourabi, H., 2011; Janssen, M., 2011; Scholl, H., 2012). It requires systems integrated across different functions in both information and service and also real one-stop service for citizens and businesses (Wimmer, M., 2002; Estevez, E., (2007); Dias, G. P., 2007; Apostolov, M., 2008). Numerous literature attested the globally widespread use of EA and SOA for EGIOP&INT issues. Literature proposes the concurrent use of comparable

technologies EA and SOA in an administrative system is most useful (Scott J. Dowell, 2007, B. Chakravarti, 2008; M.Ibrahim, 2009). Many literature specifies the successful use of agile capabilities in large scale ESs. There is already much work in the use of SOA in enterprise integration and also to create agility in systems, especially in the EG domain (Sweden, E., 2006; Votis, K., et.al., 2006; Grant, G., 2007; Soumia, A., 2011; Weerawarana, S. H., 2012; Soumia, A., 2013; Almahamid, S. M., 2013; Budiardjo, E. K., 2013; Ramtohul, A., 2014). The literature disclosed that SOA is prerequisite for enterprise agility, but it confines the business artifacts of an enterprise.

The concurrent use EA and SOA leads to view an enterprise in a new perspective SOE and is required to solve the EGIOP&INT issues along with the effective governance process in order to increase agility in ESs. Thus it need to be restructured in a service-oriented way by the transition of enterprise into SOE models (M. Mircea, 2011a). Many literature discloses the service and service-oriented modeling of systems (Arsanjani, A., 2008; Engels, G. and Assmann, A., 2008; Assmann, M., 2009; Lee, J., 2010; Tohidi, H., 2011; Hulstijn, J., 2011; Rafati, L., 2012; Mircea, M., 2012; Afsarmanesh, H., 2012; Fazlollahtabar, H., 2013; Juric, M. B., 2013; Yousef, R., 2014; Al-Khanjari, Z., 2014). But, these attempts are conceptual frameworks and architectural models that describes principles for services structuring with publishing methods to facilitate enterprise integration. It doesn't state practical solution for service creation and integration. This research modeled SOeGEA in a more pragmatic way to achieve agile EGS by specifying a programming model SCA to create, implement, compose, and reuse public services that are defined in multiple languages, deployed in multiple container technologies, with multiple service access methods (Chappell, D., 2007; Fiadeiro, J., 2012; KejriWal, M. K., 2014) along with DSM for representing and analyzing (T. R. Browning, 2002; Jacob, J., 2011; Brown, N., 2011; Jacob, J., 2012) public service integration. Though, this is based on Indian context, it is valid towards a mature EGS for countries globally.

4. Outcomes

In the first phase of the research, three analysis were conducted to explore the EGD in India. The first study investigated key EG initiatives across the country by a content analysis method of secondary sources. This study concluded that India have developed all the amenities which facilitate GEA to solve the EGIOP&INT issues in perspective of a centralized EGS [13][14]. The EGD methods and programme management approaches of the country [22][23][19] were scrutinized in the second analysis by direct interview method and content analysis of primary source materials. This investigation revealed the earnestness of EGIOP to the integration of department wise information silos [18][17]. The third analysis evaluates the global position of India based on the surveys carried out by the United Nations during 2003 to 2014 by documentary research method and it uncover the lacking for further development to reach at a matured EGS [20][15].

This research proposed a framework for EGIOP by using the existing amenities of India [21][18][17] in the second phase. Figure 1 shows this Government Enterprise Architecture Framework (GEAF) that comprises SOA based portal as a platform for all the stakeholders to interact with EG agencies in the role of service providers. Service Delivery Gateways are acting as the Government Service Bus instead of an Enterprise Service Bus that is the heart of GEAF. In view of EG as a SOE, GEAF turned into a SOeGEAF this bring agility in EGS [11]. Thus the public services can be instantiated by the SCs in a frame of SCA [30][4], this research modelled a service architecture in this

perspective [17][16]. Business functions of SOE are provided as a series of business services and are assembled together to create a particular business need [6]. Thus, integration of SCs to form a business services requires the establishment of relationships between them and the flow of deliverables among them should be coordinated. This is very challenging process and we proposed DSM method [28][8] to overcome this in the case of public services. This approach offer a view of architectural conformity of SC dependency management while integrating public services using SCA.
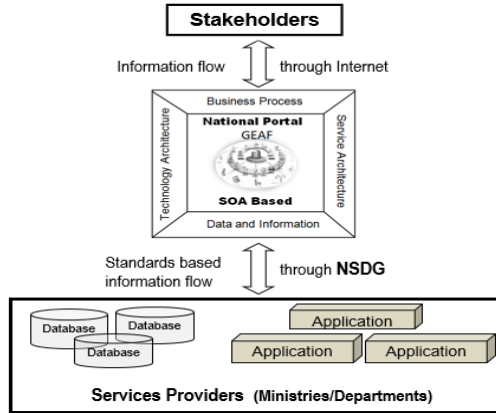
Figure 1. Government Enterprise Architecture Framework

The third phase prototyping, illustrate our approach using a citizen-friendly public provision 'Passport Service' provided by GoI in the EG domain. The researcher anticipated public services are the BP in SOeGE system [18][17][16]. Passport is delivered by central administrative government agency 'Passport Department of India', established under the Ministry of External Affairs. The applicant's identity is verified in local Police Station which is come under the Ministry of Home Affairs and has no direct control by central Department. These departments should be coordinated to issue passport. According to the componentization of BP into SCA modelling elements [6], passport service is decomposed into 3 business services, 5 corresponding composites and its SCs. Figure 2 shows SCA model and it articulates integration of related business functions by the configuration and administration of service elements. Then researcher proposed and designed a dependency model of these SCs using DSM to integrate the business service of passport service. In the proposed SOeGE system, the passport can deliver in time by completing all the formalities with the coordination of the concerned departments involved in the process under the monitoring of the central EGS.

5. Conclusions

The EGIOP&INT is a searing research problem for many researchers universally. The changing demands of citizens enforced agile EGS and it brings EGIOP&INT. Transforming the traditional enterprise model to SOEA is one of the best state-of-the-art solutions for organizations to get agile system and this research applied this philosophy towards the EGS in the perspective of India. A centralized structure for interoperability across governmental agencies in a country is a rapport for public service integration, will

increases efficiency of one-stop service delivery to citizen and also eases evolution to the EG maturity. In view of EG as a multidisciplinary area of research, this encompassing the advanced proficiencies such as SOE, SCA and DSM that belongs to various fields such as ESs, Software Architecture and System Engineering respectively.
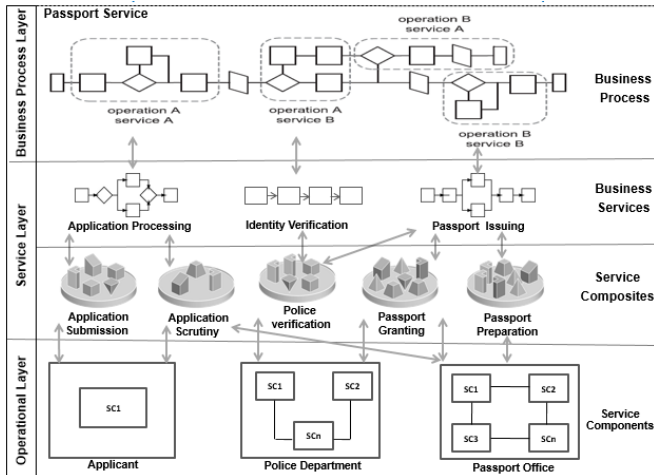
Figure 2. Service Architecture Model of Passport service in SOeGE System

Researching a solution for EGIOP&INT by constructing a framework and models is very complex and inconsistent, but quite relevant. This study has several limitations as it considers only the functional features of EGS in the perception of IT and the illustration is using only one public service. It doesn't considering the non-functional facts such as workflow management, taxonomy and ontology of services, management style, organization culture, political issues, practical difficulties as EGS of India such as loathe in the supremacy of center over the states in terms of decision making, etc. Despite the above limitations, this study is quite rare in related research universally and it is a first attempt to make a link between SOE, SCA and DSM to achieve an agile EGS so as to solve the EGIOP&INT issues. Therefore, the aforementioned limitations are the opportunities for future research.

References

[1] Aggoune Soumia, Imache Rabah, Khadraoui Abdelaziz and Mezghiche Mohamed, A Components Oriented Method for Evaluation of e-Government Information Systems Agility, *Research Journal of Recent Sciences,* 2(7), pp: 57-65, 2013.

[2] Ajay Ahuja and Neena Ahuja, Why EA is must for One-Stop e-Governmentt, In *Emerging Technologies in E-Government*, G.P Sahu (Edn.), CSI-SIGeGov Publications, 2008, pp: 235-241.

[3] C. Shailendra, Jain Palvia, and Sushil S. Sharma, E-government and egovernance: definitions/domain framework and status around the world, In *Foundations of E-government*, A.Agarwal et.al Ed. CSISIGeGov Publications, 2008, pp: 1-12.

[4] David Chappell, Introducing SCA, Chappell & Associates, 2007.

[5] Haitham Abdel Monem El-Ghareeb, A. M. Riad, Aligning Service Oriented Architecture and Business Process Management Systems to Achieve Business Agility, *Egyptian informatics Journal*, 9(1), pp: 1-14, 2008.

[6]   Hamid Tohidi, Modelling of business services in service oriented enterprises, *Procedia Computer Science* 3 (2011), pp: 1147–1156, 2011.

[7]   Hevner AR; March ST; Park J & Ram S., Design science in information systems research, *MIS quarterly,* 28(1), pp: 75–105, 2004.

[8]   Jacob, J. and Varghese, K., A Model for Product-Process Integration in the Building Industry Using Industry Foundation Classes and Design Structure Matrix, In Proceedings of Construction Research Congress, 2012, pp. 582-590. doi: 10.1061/9780784412329.059.

[9]   Mamdouh Ibrahim, IBM Technical library, *Service-Oriented Architecture and Enterprise Architecture*, https://www.ibm.com/ developerworks/webservices/library/ws-soa-enterprise.

[10] Marinela Mircea and Anca Ioana Andreescu, A Roadmap towards Service-Oriented Enterprise, *Journal of Organizational Management Studies*, Vol. 2012 (2012), DOI: 10.5171/2012.628284.

[11] Martin Assmann, Gregor Engels, Transition to Service-Oriented Enterprise Architecture, In R. Morrison, D. Balasubramaniam, K. E. Falkner (eds.): Proceedings of the Second European Conference on Software Architecture (ECSA 2008), Paphos (Cyprus). Springer, LNCS, vol. 5292 (2008), pp: 346-349.

[12] Pallab Saha, *Enterprise Architecture as Platform for Connected Government*, Report of Government Enterprise Architecture Research Project, NUS Institute of Systems Science, 2010.

[13] Paul, A., e-Governance the Existing Indian Scenario, In Proceedings of the National Seminar on Advanced Computing & Information Management, Kerala, India, 2010a, pp: 3-8.

[14] Paul, A., e-Governance, a Paradigm shift in India, *International Journal of Interdisciplinary Studies and Research: Baselius Researcher*, 11(1), pp: 63-73, 2010b.

[15] Paul, A., Paul, V., Analysis of e-Government Development in India, *IJCCSE* 2(2), pp: 225-231, 2015c.

[16] Paul, A., Paul, V., IT Perspective of Service-Oriented eGovernment Enterprise, *International Journal of Computer, Information, Systems and Control Engineering*, 9(1), , 153 – 158, pp: 2015a.

[17] Paul, A., Paul, V., A Conceptual Framework for Interoperable e-Government System in India, In Compendium of 17th National Conference on e-Governance (NCeG), Government of India, 2014b, pp: 45-50.

[18] Paul, A., Paul, V., A Framework for e-Government Interoperability in Indian Perspective, *International Journal of Computer Information Systems and Industrial Management Applications* (IJCISIM), 6 (2014a), pp: 582-591.

[19] Paul, A., Paul, V., Challenges and Strategies of e-Governance in India, *International Journal of Interdisciplinary Studies and Research: Baselius Researcher*, 12(2), pp: 472-480, 2011c.

[20] Paul, A., Paul, V., e-Government Development Scenario in India based on United Nations Surveys, *International Journal of Current Research*, 7(4), pp.15451-15461, 2015b.

[21] Paul, A., Paul, V., The e-Government Interoperability through Enterprise Architecture in Indian Perspective, In A. Abraham et.al (Eds.): IEEE Proceedings of WICT, India, 2012, pp: 645-650.

[22] Paul, A., Paul, V., The Implementation issues of e-Governance in India, In the Proceedings of the National Seminar on Modern Trends in Electronic Communication & Signal Processing, Kerala, India, 2011a, Excel India Publishers, pp: 43-48.

[23] Paul, A., Paul, V., The Project Management Model for e-Governance in the Context of Kerala State, In A. Abraham et.al (Eds.): Springer Proceedings of ACC Part II CCIS 191, 2011b, pp. 201-209.

[24] Raghunath Mahapatra and Sinnakkrishnan Perumal, Enterprise Architecture as an Enabler for E-Governance: An Indian Perspective, In Pallab Saha (Eds.): Handbook of Enterprise Systems Architecture in Practice, IGI Global Publications, 2007, pp: 272-289.

[25] S P Kulshrestha, Public Service Delivery System and E-governance, CSI Communications, 37(4), pp: 18-20, 2013.

[26] Scott J. Dowell, Enterprise Architecture within the Service-Oriented Enterprise, In Pallab Saha (Eds.): Handbook of Enterprise Systems Architecture in Practice, IGI Global Publications, 2007, pp: 382-399.

[27] Shahani Markus Weerawarana, Kanchana Thudugala and Wasantha Deshapriya, An Agile Approach towards eGovernment Solution Procurement and Implementation: An Experience from Sri Lanka, *Journal of e-Government Studies and Best Practices*, Vol. 2012 (2012), DOI: 10.5171/2012. 549264.

[28] T. R. Browning, Process Integration Using the Design Structure Matrix, *System Engineering*, 5(3), pp: 180-193, 2002.

[29] Weiss, P., Modeling of Service-Oriented Architecture: Integration of Business Process and Service Modeling, *Information Sciences and Technologies Bulletin*, ACM Slovakia, 2(2), pp: 79-92, 2010.

[30]   Z. Ding, Z. Chen, and J. Liu, A rigorous model of service component architecture, J. ENTCS, vol. 207, April, 2008.

Virtual applications: analysis of forums and mechanisms for participation in virtual environments and the public policy process

Larissa Galdino de MAGALHÃES[a,1]

[a] *Institute of Philosophy and Human Sciences*
University of Campinas

Abstract. The debate on the internet and democracy seeks to highlight the importance of the contributions of the theories of participation and deliberation with respect to innovative digital experiences pro-democracy. However studies treat analysis watertight, escaping the perception of the established correlation between the context and the mechanisms of political practice everyday. Assuming that the use of the internet can impact or influence political participation, to include virtual environments and its manner of use as a significant impact on public policy process, and not necessarily in response to possible democratic deficiencies pointed out by the literature current. The project deals with the bodies and mechanisms of participation in virtual environments, virtual appliances in the national context, therefore, relies on the perspective of institutionalized participation on the Internet experiments, and the prospect of gradual integration into the assembly's agenda, policy formulation , decision making, implementation and evaluation of public policy process. In other words analyzes virtual applications created by governments on the internet, into the process of public policy, emphasizing specifically its manner of use, either function assigned to the phases or cycles of the policy process, or on its function "politics" of formation of public opinion, voting behavior and communication flow. The analytical scenario speculates that the virtual applications, while instances and mechanisms of participation in virtual environments can help to foster functions in the "cycle of public policies", although secondarily, admits the possibility of those virtual appliances, in its manner of use, do not have role in the process, but political function.

Keywords. Internet, public policy, participation.

1. Introduction

This project deals with the relationship between information and communication technologies (ICTs) and political institutions in the democratic context of recent decades. Analyzes, specifically, the function or ways of using channels, events, devices, tools or applications created by governments in virtual environments (internet) and its insertion in the process of public policies, essentially in character instances and mechanisms of participation. It is therefore proposed that the understanding of the function of political participation experiences in the public policy process, from its size and type of policy, related to some stage, cycle or function, creating new participatory processes exclusively in virtual environments.

---

[1] E-mail: larissagms@yahoo.com.br.

According to Medeiros and Guimarães (2004), the depletion of conventional management platforms was driven by the need for governments to increase their revenues by improving its internal processes; and the pressures of society for the government to perfect their spending acting with transparency and quality in the provision of services to citizens and changing the approach to public management. Soon the use of ICTs for political platform and through new institutional arrangements of public administration, provides the participation and / or resolution of various social actors in state institutions, whether through design, implementation, regulation, implementation, monitoring and decision of public policy.

The various institutionalized forms of participation of civil society in interaction with the state as participatory budgeting, public policy councils, conferences, participatory master plans, committees and thematic committees, and others, can be interpreted "as different ways of incorporating citizens and civil society associations in the deliberations on policy (AVRITZER 2008, p. 45) ". [2] These institutionalized processes are located in participatory arrangements "(...), so common elements of public administration, but with differing combination of mechanisms of representation, participation and deliberation in the public policy process (NASCIMENTO, 2012, p .70)."

Virtual applications created by government are being inserted or added to the experience, forums and mechanisms for participation, and competes be regarded as engendered in the public policy process, relating them to their functions or forms of use, which supports the interactive approach analysis. Its operation depends on endogenous factors (institutional design; dynamics of participation and scope of the policy) and exogenous factors (civil society organization, public management, political, political project, technical and bureaucratic constraints, political will, actors, etc.).

So the strategy that guides this, consider a new object, seen from the public policy cycle, and a secondary circumstance, reflects on the political process. Therefore the research assumes virtual applications as instruments incorporated into the policy process, whose function can be linked to any stage of the process, promotes democratic potential, or can be used to renew the ways of popular participation and making government decisions (FREY, 2002), contributing to the democratic refreshment, insofar as it involves the most diverse sectors of civil society in addition to the representative institution (COLEMAN, 1999). Later, postulates the perspective on how to use virtual applications, as they do not have functions in the public policy process, but "political" functions, establishing relations with the political landscape, public opinion, communication flows, social representation, voting behavior or political decision. Thus, although governments use tools that enable a distinct relationship between citizen and the state, not always these features ensure the primacy of civil society, therefore, "are used as traditional media outlets, promoting one-way information without worrying about hosting, the potential permitted by the Internet, the contributions that citizens can and want to offer (BRAGATTO, 2009, p.152). "

1.1. Internet and expansion of democracy

Considering the relationship between the internet and means or forms of state closer to the citizen, it is speculated that "new technologies would therefore have the potential to build a new relationship between government and citizens, to be assumed that enable public administration more efficient, democratic and transparent (MEDEIROS, GUIMARÃES, 2004, p.49)."

For some it is precisely the promotion techniques different conditions of the Internet relative to other traditional platforms of communication, which invigorates expanding expectations of democratic practices (NEGROPONTE, 1995; MORRIS, 2000; FROOMKIN, 2004; STANLEY, WEARE, 2004; KLEIN, 1999; MCNIVEN, MARCHE, 2003; DELLI CARPINI, 2000). However, others disagree with this positive relationship, when considering the uneven access to technologies, the dubious quality of information, the threat to freedom and privacy of users (BUCHSTEIN, 1997; SUNSTEIN, 2001; DAVIS, 1999; DEAN, 2003; WILHELM 2002).

The discussion of internet and politics, two affiliations include: the institutional aspect and the social aspect. According to Gomes (2007) the literature on digital democracy shows a differentiation as to the emphasis on society and the State. The social aspect involves about civic engagement, public sphere, online deliberation, restoring relations with the capital.The institutional component operates in three ways: the study of the digital configuration of democratic institutions, parliaments digital cities; activities transmitted from the state to citizens, such as online public services and e-government; or initiatives toward the citizen to the state, including opportunities for participation or enabling "(...) the inputs from the citizens in the form of votes, responses to polls, decisions and budget suggestions, registration and discussion of opinions in forums electronics, etc.) (GOMES, 2007, p. 11)".

The different interpretations of democracy are related to different political theory platforms. Roughly speaking, with the inclusion of the Internet in politics, criticism of the composition of democracy are transferred to the debates on democratic practices in virtual environments, so there are different divisions in this debate. The understanding of the impacts and effects of the Internet and political issues also have variations and evaluations are related to the chosen theoretical platforms, as in the case of participation, effectiveness or benefits of democracy.

Even the most progressive currents are cautious about the positive effects of the Internet on the public sphere, and participation potential in the democratic environment, and the Internet can help make complex political issues between society and state, allowing a greater flow of information, strengthening communal ties, expanding political participation and promoting new forms of relationship between society and state institutions (CHADWICK, 2003; ESTER, VINKEN, 2003).

But another line warns of the dangers as the Internet marketing process, the lack of credibility of the information available (BUCHSTEIN, 1997; DEAN, 1997; BARNETT, 1997). Buchstein (1997, p.259) ponders the negative effects of the Internet use of political ends, as it can encourage the private lives of public affairs coming into collision with the publicly expressed citizenship. Still, for the author, lacking effective controls for the propagation of polarized speeches and messages, disregarding inequalities and address issues such as gender and race.

A final group of researchers recognized the potential of democratic growth of the Internet, but aware that the effects and impacts of political issues in the network will depend on the organization and its use (COLEMAN, 1999; BARBER, 2000). Papacharissi (2002, p.10) highlights the need to speculation the Internet and similar technologies can revolutionize the public sphere or become part of the current democratic environment, especially in this context in which traditional political activities require revitalization.

This differential scenario also boosted the demand for changes in the nature and structure of democratic institutions. The confluence of conditions and technological innovations ultimately reverberate the demands for participation of the

population and of the State itself to promote forms of political participation on the Internet. It is this analytical purpose which deals this research.

1.2. Public policies: the cycle and its limits

According to Souza (2006), public policy is a field of knowledge which aims to analyze the government into action, and that action therefore focuses on the role of governments, and directs his gaze to the conflicts around the context, interests. It is therefore denomination and object of this research, at the same time, which involves a diverse range of stakeholders, decision-making levels and involves subsequent processes of decision or proposition, formulation, implementation and evaluation.

The explanatory models [10] on public policy, for purposes of analysis stands out, a priori, the model of "public policy cycle" as analysis feature of public policies, i.e, heuristic model. Prioritizing interactive approach of the three dimensions that make up the policy analysis, institutions, political dispute (power relations) and the public policy itself, the heuristic model of the public policy cycle defined phases function as device for critical analysis of the political processes under way. This, since it is important to know in what context they operate policies (the specifics) and the arrangement constitutes in each case to enable analyzes that reflect the political debate in question and which are not simple reproductions of rational political models and general (BAPTIST, REZENDE 2011, p. 161).

The heuristic model of the public policy cycle is to find viable answers to the research problem, that is, although does not correspond exactly to what happens in practice works as an analytical subsidy to the political process. Its relevance stems from the fact of its phases have certain conformity with the following elements of the political-administrative process, and can be investigated regarding its actors, relationships, resources, networks and practices typical of each phase. Although systemic perspective of the public policy cycle that "breaks" the process at different stages have been criticized on the causality between one stage and another; as the sequence and direction of the stages and empirical studies (CAPELLA, 2006), for purposes of this research is considered that "to divide the public act on partial political and administrative process phases of problem solving, the 'policy cycle' is over revealing a heuristic model rather interesting to analyze the life of a public policy (FREY, 2000, p.226)."

The idea of public policies cycle Howlett and Ramesh (1995) as a process of reformulation of times and not as accurate analytical mechanism guides the analysis Improved Model, condensing stages of the policy process into five stages: (1) Mounting the agenda; (2) policy formulation; (3) decision making; (4) implementation and (5) evaluation.

Considering the discussions on the size of the Internet and political participation as relevant to the public debate and influence in public decision-making (GOMES, 2011), this research project seeks to understand the role of virtual applications, as bodies and mechanisms participation in virtual environments, integrated into the public policy process, casting lights, specifically on the analysis of public policy and its relationship with the Internet, without however disregarding the possible political use of covert applications in the public policy cycle. Understanding the summarized above models and approaches enables contribute to understand the reasons why public policy is designed, the trajectory and the role of the actors, the conflicting relationships, groups and institutions involved in public policy, and then to

understand the role of virtual applications in the process of public policy, or how to use the policy. It is, in short, the intellectual challenge of this research.

2. Objectives

The main objective is the realization of such objects and their function in the process of public policy or its policy function, through the survey and mapping of different practices or digital experiments.This means understanding the main points of such practices, namely: when ?, where ?, how ?, why ?, and running both function in the public policy cycle, as possible forms of political use.

That is, it is here analysis using two different standards, in which, first, the perception of the functions of virtual applications as will be approaching the specified function of the public policy cycle - ie if the virtual appliance connects to the time of policy formulation or its monitoring, or review etc.

Second, considering the possible fragility and non-compliance of virtual appliance to the public policy cycle, we will consider its existence as that linked to external goals to cycle and thus closer to what we call "political use" . In this research field, will prevail the vision of virtual applications as mechanisms for expanding social base and political mobilization.

Thus, the first specific objective of the project seeks to understand how these initiatives relate to the public policy cycle functions. Therefore, the main objective is to analyze the role of virtual applications of government, while instances and mechanisms of participation in virtual environments, considering the 'production cycle' of public policies. Based on the data collected and preliminary results, the second specific goal is understanding the ways of use of the bodies or mechanisms of participation in virtual environments that have no role in the public policy cycle, but use forms related to politics, referring to the 'political role' of the application.

3. Material and methods

The theoretical-methodological problem at hand is to establish connections between function in the cycle of public policies through virtual applications when negative, in addition to point out the ways to use the applications in politics. No cumulative contribution that connects these goals and substance to the hypothesis, which supports the operation of mixed method research.

On the relationship between cause and effect, what changes is the public policy process as it is composed of the dependent variables (functions, type of policy, actors, networks, interests, rules and regulations, and political context) that they operate in environments Virtual in different ways, so the independent virtual applications the cycle. But considering ways to use virtual applications, they become dependent on interests, molds, power relations, built by the policy. Therefore, the articulation of assumptions, and location functions or forms of uses can illustrate the variation of the causes, modes, associations, and produce certain effects while political participation tools in virtual environments or as an instrument of world politics (applications virtual) (MAHONEY, 2003; REZENDE, 2011; SCHMITTER, 2008).

From the point of view of order of search problem of this approach uses: the quantitative methods, translating values and qualifying information categories; and qualitative, interpreting phenomena and attributing meanings or functions in the analytical process through direct data collection.

Considering the proposed work plan, research objectives and hypotheses, to establish correlations and verification of admissible assumptions, defining the contours

and limits of this research based on methods and materials, through the following stages:

1st Stage of knowledge and understanding of the research:

1.1.   Exploratory research: preliminary study's research objective providing greater accuracy and understanding of subsequent research. This approximate phase favors research with little cumulative studies, whether through observation, patterns, ideas and hypotheses and also the refinement of theories and concepts to be applied and further developed (PIOVESAN, TEMPORINI, 1995);

1.2.   Research universe: (a) spatial area: the phenomenon of observation locus is the internet, and the chosen spatial parameter is the national context, considering the federal, state and municipal levels. (B) time frame: the period in which the phenomenon is circumscribed is between 2014 and 2015, however can go back in time, to show the time series of a particular phenomenon related to the virtual appliance.

1.3.   Literature review: theoretical basis for the treatment of the topic and the research problem, from material published in journal articles, books, newspapers, and material available on the internet, through bibliographical and documentary research, enabling the discussion of ' state of the art '.

2nd Stage of application and analysis of the survey:

2.1.   Monitoring objects through the field study: carried out through direct observation of the activities of governments at the national level on the internet, either through official channels, but also on social networks (more informal channels of interaction) capturing objects (virtual applications mediated by the idea of forums and mechanisms for participation in virtual environments) to occur interpretation of that reality.

2.2.   Data organization for the Virtual Application Mapping: to map virtual applications, considering the breadth of virtual scenarios, considered key-strategic prospecting elsewhere, such as scientific journals, studies and reports on the main themes, websites, portals, blogs, social networks, discussion lists, forums, and search engines, mostly used, like Google, and other academic environment, such as Periodicals Portal of Capes, Google Scholar and Sciencie Direct.

2.3.   Implementation of mapping: meets specific criteria, based on the strategic key of virtual applications and analytical dimensions already established by the state of the art (and referenced in the review).For research, the use of the phases or stages of the cycle of public policy to name the function of the applications it is intended; already for ways to use other literature has been listed related to political marketing, electoral behavior, media and public opinion.

2.4.   Database: the construction of the database will be used the software Statistical Package for Social Sciences (SPSS) version 20, for this type of encoding. To obtain the virtual data is guided by two methods: guided navigation and the evaluation framework (set of variables / phases that meet the cycle), the material will be backed up. Clarifies that the Framework evaluation (TONHATI, 2007; SILVA, 2009; MAGALHÃES, 2013) comprises a number of criteria, built on literature (above) for virtual analysis of the function. It therefore indicators that assess and identify the construction, functions, design, information practices, content, services, transactions, and online participation. Already guided navigation ensures that the route analysis and evaluation through the items, criteria and parameters (contained in the framework) is carried out properly and effectively. The questions that guide the framework of evaluation and compilation of the data, compiled for the following categories of virtual data base will be used: case name; Location and population; promoter; political

system; realization (temporality); mode - online or mixed; the participants: number and selection technique (if any compilation); Application role in the public policy cycle (SAMPAIO, 2014). In cases of applications related to the forms of use when it comes to politics, initially will be classified as 'Other', and as the seat is being consolidated through observation and systematization of data will be possible to set parameters or standards on how to use, so the materials and methods listed in knowledge and understanding stage of research will be constantly taken up during the analysis.

3rd research synthesis stage:

3.1. Analysis Consolidation: standards and / or comparative analysis compiled the data through virtual exploration, and located the duties of the public policy cycle, can systematically analyze becomes each of these stages, making use of explanatory models and key Theoretical tracked in the previous steps (debate with bibliography), establishing causal relationships able to respond if the virtual applications are innovations that can help to foster the process of public policies encouraging democratic political participation through virtual environments.

4th Stage of research evaluation:

4.1. Dissemination of results: still on the data and forms of storage, files will be made of logs, usually in electronic form, based on analytical categories and using SPSS or Excel. For data analysis purposes I suggest three streams: the reduction of the data from the categories; Exposure data from the typologies; and then check on the basis of inferences evidence and theoretical premises (MILES, HUBERMAN, 1984).

4.2. Final elaboration of the text: the forms of analysis and use of certain theoretical and methodological strategies, and analytical resources in question, it is understood that variations on the run, and in the public policy process is influenced by the function or form of use of virtual applications . Thus the final drafting of the text rests in the forms of analysis of the results, taking as input the concurrent debate of knowledge production and research.

4. Shape analysis of the results

Considering the above program, the proposal is to think of a connection between virtual applications in their functionality in the public policy cycle. Moreover, not all created and forged tool in a virtual environment by governments achieve the function that whether or called, so the application would be out of the public policy process, because its function in the cycle is not real and effective. But this application may be related to the world of politics, the constraints and limits of doing politics, the interests of governments, the flow of communication, speeches, political and public communication strategies, marketing, and so are not disregarded in the proposed target and in the background, it is intended to analyze the ways to use these virtual applications connected to its political use.

Finally, to take three-dimensional analysis to the study of public policy in order to substantiate the goal, which presupposes: institutional dimension (polity), procedural dimension (politics) and the material dimension (policy/policies).

Coleman (1999) Flip the internet has no functionality without people to do something with it, and its manner of use is relevant. Lee Salter (2004) in his article on the structure and forms of internet use on deliberative democracy, brings important information to the understanding that the internet offers many tools and possibilities, but it depends on agents for their development and definition purposes. In view of this, to analyze the virtual applications that are not located in the public policy cycle, but in politics, operates the prospect of "forms of use", based on the concept that the

technologies are developed to meet certain uses and purposes, so home use corresponds to a need, and the need, an interest. What this project shows is that often the technologies (virtual applications) can be prepared and used for purposes other than those anticipated. Moreover, Salter (2004) emphasizes that the choice of a form of use or another depends on the agents.

References

BIMBER, B.. *The Internet and Political Transformation: Populism, Community, and Accelerate Pluralism*. In: Polity, v. 31, n. 1, pp. 133-160, 1998.

CHADWICK, Andrew. *Internet Politics – States*, Citizens and new communication technologies. Oxford University Press, 2006, EUA.

_____. *Bringing e-democracy back in*: Why it matters for future research on e-governance. In: Social Science Computer Review, Londres, v. 21, n. 4, p. 443-455, 2003.

COLEMAN, S., & BLUMLER, J. (2009). *The Internet and Democratic Citizenship: Theory, Practice and Policy*. Cambridge. Cambridge University Press.

COLEMAN, S. *Can the new Media Invigorate Democracy?* In: Political Quarterly, Oxford, v. 70, n.2, p. 16-22, 1999.

DEAN, Jodi. (2003). *Why the Net is not a Public Sphere*. Constellations, 10 (1), p. 95-112.

_____. *Virtually Citizens*. In: Constellations, Oxford, v. 4, n. 2, p. 264-281, 1997.

DELLI CARPINI, M. "Gen.com: Youth, Civic Engagement, and the New Information Environment". Political Communication, 17: 341-349, 2000.

ESTER, Peter & VINKEN, Henk. *Debating Civil Society On the Fear for Civic Decline and Hope for the Internet Alternative*. International Sociology, 18(4), pp.659–680, 2003.

FROOMKIN, A. M. (2004). *Technologies for Democrac*y. In: Peter Shane (ed.) Democracy Online: The Prospects for Political Renewal Through the Internet. New York: Routledge.

HOWLETT, M., RAMESH, M., 1995. *Studying Public Policy*: Policy Cycles and Policy Subsystems. Oxford University Press, Oxford.

KINGDON, John. Agendas, Alternatives, and Public Policies. Boston: Little, Brown. 1984.

MACINTOSH, ANN. *E-democracy and e-participation research in Europe*. In: Digita lGovernment. Springer US, 2008. p. 85-102.

_____.; WHYTE, A. *Towards an evaluation framework for eParticipation. Transforming Government: People, Process and Policy*, Vol. 2, N. 1, p. 16-30, 2008.

_____ . (2003). *Using Information and Communication Technologies to Enhance Citizen Engagement in the Policy Process*. In: OECD. Promise and Problems of E-Democracy: Challenges of Online Citizen Engagement, OECD, Paris, pp. 19-58.

PAPACHARISSI, Zizi. (2002). *The virtual sphere:* the internet as a public sphere. New media & society, 4 (1), p.9-27.

ROSE, Jeremy; RIOS, Jesus; LIPPA, Barbara. *Technology support for participatory budgeting*. Int. J. Electronic Governance, Vol. 3, No. 1, 2010.

SÆBØ, Ø.; ROSE, J.; FLAK, L. S. *The shape of eParticipation*: Characterizing anemerging research area. Government Information Quarterly , v. 25, n.3, p.400 – 428,2008.

SÆBØ, Øystein. *eDemocracy tools*. Ad hoc Committee on e-democracy (CAHDE). In.: Electronic democracy ("e-democracy") -Recommendation adopted by the Committee of Ministers of the Council of Europe. Strasbourg, 8-9 October 2007.

SABATIER, PAUL, JENKINS-SMITH, HANK. *Policy Change and Learning: The Advocacy Coalition Approach. Boulder*: Westview Press. 1993.

SALTER, LEE. *Structure and Forms of Use*. A contribution to understanding the 'effects' of the Internet on deliberative democracy. Information, Communication & Society, Londres, v. 7, n. 2, p. 185–206, jun. 2004.

SCHMITTER, PHILIPPE. 2008. *The Design of Social and Political Research*. In Approaches and Methodologies in the Social Sciences: A Pluralist Perspective, edited by Donatella Della Porta and Michael Keating. Cambridge: Cambridge University Press.

SMITH, Graham et al. *Measuring the ‗deliberative quality' of an online experimental mini-public*: methodology and early results. Paper prepared for ECPR General Conference, Potsdam, September 2009.

STANLEY, J. W.; WEARE, C. (2004). *The Effects of Internet Use on Political Participation*: Evidence from an Agency Online Discussion Forum. Administration & Society, 36, 503-527.

SUNSTEIN, Cass. *The daily we: Is the internet really a blessing for democracy*? Boston Review. 2001.

VAN DIJK, Jan AGM. *Digital Democracy*: Vision and Reality. Public Administration in the Information Age, v. 19, p. 49, 2012.

Long-Range and Low Cost Technologies: Proposing Guidelines for the use of SMS in the Public Sector

LANZA, Beatriz B. B.

Universidade Federal do Paraná UFPR,Companhia TIC Paraná CELEPAR, Curitiba, Brazil. Center for Technology in Government at UAlbany, USA bialanza@pr.gov.br

Abstract. Mobile government could be conceptualized as the use of diverse mobile technologies to provide information and services to citizens, businesses, and other stakeholders. Governments around the world, particularly in developing countries are finding mobile strategies to be useful ways to reach to their citizens, including disadvantaged groups. In fact, individuals are increasingly using mobile phones for accessing information and this trend is becoming stronger with the improvement of long-range technologies and the high penetration of mobile devices in developing countries. However, their use to deliver government services to citizens remains generally low. This research argues that the government use of Short Message Service (SMS) to offer information and services to citizens is low because of a lack of a clear and actionable business model for this technology. So, it proposes guidelines for SMS-based government services.

Keywords. Electronic Government, Mobile Government, Short Message Service, Configuration Theory

1. Introduction

Government-delivered services with universal access through internet are essential requirements for electronic government (eGov) [8]. Just to cite an example, in Brazil has only 49 percent of its population using the web regularly [6]. Mobile phones have become the most rapidly adopted technology in history and the most popular and widespread personal technology in the world [22]. Mobile government (mGov) is important in many developing countries because of the extensive use of mobile devices by citizens and this puts great pressure upon its implementation [14]. In Brazil there are 37 percent more cell phones than inhabitants [12] [24], but only 31 percent of the phones have internet access. Forty one percent of the population use traditional equipment and not smartphones. The most important use is to receive and send Short Message Service (SMS). This service corresponds to 66 percent of mobile phone use in Brazil [6]. In 2009, over 70 percent of world's population used SMS and Voice technologies showing that other countries have similar situations [17].

Government agencies use technology to offer services and information and to contact people, but mostly through old and sometimes obsolete bureaucratic structures [3]. While there are many worldwide success stories with the use of SMS, this service remains incipient when compared to the increase in number of mobile phones in many countries [15]. Governments and researchers apparently ignore the potential value gen-

erated by the use of the simple and low cost SMS technology. Few have looked at mobile phones under the business and information system perspective [23], with the mobility business model receiving very little attention [1]. There is a need to examine issues related to mGov from different perspectives. It seems that mGov may substantially influence the generation of emergent complex strategies and tools for eGov [14].

Interdependent imperatives like environment, structure, leadership, strategy [17] and sustainability [11] help to shape organizational settings when considered holistically [16]. Patterns and generalized forms define the components related to each imperative that are combined into gestalts or configurations [17], The absence of an appropriate business model may be behind the low use of SMS by governments [15], but its use may be long lived and extensive providing all mobile devices embrace it in all government levels. The aim of this paper is to propose guidelines for SMS-based government services. It is based on a configurational approach to organizational design [17] [18] [19].

The paper is organized in five sections, including the foregoing introduction. The second section describes the different approaches to studying mobile government, short message service and government and, configuration theory. Section three describes the methodology used. Section four presents empirical setting. Finally, section five provides next steps for this research.

2. Theoretical and Conceptual Approaches

This section briefly presents some basic concepts from c Electronic Government, Mobile Government, Short Message Service and Government, and Configuration Theory.

2.1 Mobile Government

This research uses mGov as government-delivered services on mobile phones. mGov is an essential platform of communication between government and citizens. The mobile services value chain includes the (1) government, which is expected to lead in developing and improving services. It also includes the (2) hardware industry, software and solutions; (3) telecommunications companies, with the infrastructure and services; (4) brokers (aggregating companies specialized services in integrating networks and billing) and finally at the end, (5) the citizen [11] [15]. This chain can be further assisted by research institutes and universities that can provide predictive models and impact analysis [24]. The value chain integration can be further complicated by the lack of clarity of its actors and their roles [10].

mGov is being built on eGov efforts, and there are basically two major issues related to this transition [14]: (1) mGov is inevitable, the main forces that influence its adoption are the technological advances in network infrastructure and wireless internet; (2) mGov will not bring additional costs to eGov - there are services in eGovernment that are the same on mobile platforms. However, the real value of mGov comes to light with services and applications that are only possible using the wireless infrastructure and modern mobile devices.

Some governments recognize the great potential of mobile devices. For example, in Singapore the mGov program is part of the central government's plan, released in June 2011, in order to increase the range and diversity of services provided by government. In recognition of the growing importance and widespread use of SMS for deliv-

ery services in that country semi-annual performance tests to check the SMS response time are conducted [26].

For Cunha and Miranda [9], it is important to understand that government efforts to use digital media can be unfruitful if they do not consider citizens' internet access and the necessary skills to meaningfully use some of these information and services. So, one of the most basic technologies embedded in all cell phones could be the most efficient way of government to communicate with citizens and provide them with information and services [25].

2.2 Short Message Service and Government

While there are many success stories about government-delivered services through the use of SMS, the proportion is still far from the increase of mobile phones. This proves to be a critical factor which prevents all three government levels to fully incorporate delivery services through SMS. While there are many success stories about government-delivered services through the use of SMS, the proportion is still far from the increase of mobile phones. This proves to be a critical factor which prevents all three government levels to fully incorporate delivery services through SMS. Through the lens of Configuration Theory we will try to identify the relationships present in the value chain of mGov and investigate the existence of possible relationships between variables and the perceptions of power involved in this process.

2.3 Configuration Theory

Configuration Theory (CT) proposes that the environment, structure, leadership, strategy, and sustainability are the imperatives that help shape organizational settings [13] [17]. For the development of a configurational perspective in organization studies, Miller argued that these variables [17]: (1) were widely accepted paradigms in organizational analysis; (2) ad solid empirical demonstrations; and (3) had already played a central role in the generation of organizational settings repeatedly.

Mintzberg [21] also adds that this perspective in organizations writes its history based on periods of stability and transformation. According to Pereira and Toni [25], due to the great changes taking place, society suffers disruptions in its technological and social systems, leading to most changeable and uncertain environment that require diverse combinations of elements of an organization. For Gimenez [13] this approach indicates the existence of multiple ways to achieve organizational success, but there is a limited number of viable patterns of behavior that can be followed by organizations to accomplish their goals.

The complexity of the organizational governance environment requires that each of its characteristics is studied and understood within a continuum, where each topic should be examined in a unique way in its context, and the environmental variables and the organizational reality are considered. There are various gears and levers in a decision-making structure, large acts are the consequences of innumerable and often conflicting smaller actions by individuals at various levels of government [2]. Therefore, the multidimensional rationality proposed by CT seems to be highly adequate to support the business model building process for mGov. It can include proper contexts and result in an actionable model [20].

According to Miller [17], synthesis is used to identify patterns and configurations and forms of generalizations and analysis is used to define and measure the components or attributes of a phenomenon. This way proposed by Miller [17] is adequate to deepen understanding of the historical relationships and sequence among the five im-

peratives (environment, structure, leadership, strategy and sustainability), how and why this imperatives might change during the course of the projects's lifecycle. The lifecycle concept suggests an ordering that describes some common transitions among imperatives [17].

3. Methodology

This chapter aims to further develop this empirical research by designing the research methodology to meet the aim, objectives and research questions of this study.

3.1 Research problem, objective and research questions

The problem domain in this research is the government use of SMS to offer information and services to citizens is low because of a lack of a clear and actionable business model for this technology. To aim the objective this study, that is, proposing guidelines for SMS-Based Government Services, the research questions that will be answered are the following: (1) What are the trajectories of mGov State and Local governments, highlighting the transitions between the five imperatives: Environment, Structure, Leadership, Strategy and sustainability? (2) What are the common dimensions and differing dimensions of State and Local Government? (3) What guidelines for SMS-Based Government Service?

3.2 Research type, data collection

This study is based in Multiple Case Studies on any mix of quantitative and qualitative evidence. The research participant are: project managers and mGov's technicians of the State and local government and project managers and mobile services value chain's technicians (software and solutions companies, telecommunications companies and brokers). No part of this research the hardware industry and citizens. The data collection: official government documents, general media, interviews and questionnaires. Analysis of data: documents be analyzed using documentary analysis, interviews by the technique of content analysis and questionnaires for content analysis.

3.3 Using Configuration Theory

The synthesis lens is used to understand holistically the mGov cases that have occurred in the past. This way is adequate to deepen understanding of the historical relationships and sequence among the five imperatives (environment, structure, leadership, strategy and sustainability), how and why this imperatives might change during the course of the projects's lifecycle.

To answer the first question for each case - *What are the trajectories of mGov Project, highlighting the transitions between the five imperatives: Environment, Structure, Leadership, Strategy and Sustainability?* - data will collect via the deeply interviews using 19 dimensions, each one related to its respective imperative, as show Figure 1.
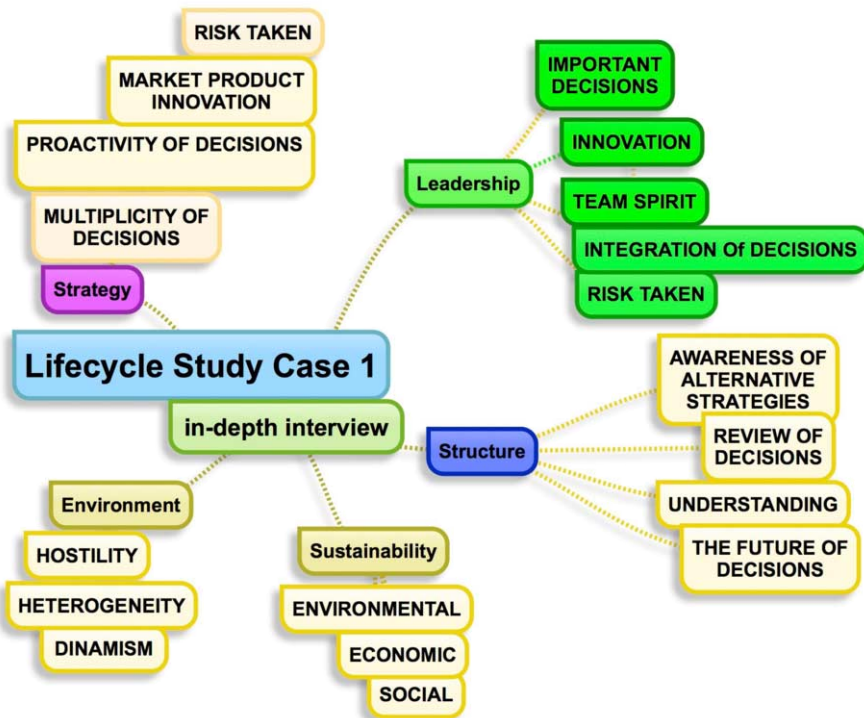
Figure 1 - First Research Question

With data collected will be assembled a lifecycle, from its birth until the date of data collection for each project, showing the sequence among the five imperatives. Then, for each case its lifecycle is designed and compared with the patterns and configurations suggested by Miller, Mintzberg and Gimenez, to verify common transitions among imperatives over lifecycle, identify phase of cycle (birth, growth, maturity, revival and permanence), factors initiating, and resulting imperatives. This perspective permit writes the historical from each case studied based on theirs periods of stability and transformation.

To answer the second research question *What are the common dimensions and differing dimensions of State and Local Government?* Using content analysis to compare the life cycles with the literature researched in each project studied to highlight what are common and what are differing. And, if so, what emerged from the data that has not been presented in the literature.

Finally, I hope I will have enough data to answer the fourth research question - *What guidelines for SMS-Based Government Service* can be proposed from the case studies?

4. Empirical Setting

Brazil is the biggest country simultaneously South America's largest country by geographical size (8 515 767 km²) and both population (202 768 562 inhabitants). The

only Portuguese-speaking country in both North and South America. This Portuguese heritage dates back to the 1700's, when Brazil was first colonized by this European nation. Therefore, there is also a large proportion of Brazilian inhabitants that have an African heritage. Not only European and Asian descents immigrated to Brazil in the 19th century, but also Japanese, Polish, Spanish, Italians and Germans.

In Brazil there are 278.482.125 cell phones devices, these 76,49% pre-paid. There are eight cells companies: Vivo (28,47%), TIM (26,97%), Claro (25,33%), Oi (18,14%), ALGAR (0,43%) Nextel (0,54%), Sercomtel (0,02%), Others (0,10%). In Brazil regardless if you have or not credit in your billing plan, you receive sms. Only the sender pay a charge for it.

Paraná, whose capital is Curitiba (2 million inhabitants), is one of Brazil's southern states, it is also bordered by São Paulo, Argentina and Paraguay and the Paraná River. It is the ninth largest state in terms of the area it covers, which totals at 199.314.9 square kilometers or 76 955.9 square miles. In addition, it is the home of the UNESCO World Heritage Site, the National Park of Iguaçu. The entire state has a population of well over 10 million people.

Celepar is a semi-public-held company, whose largest shareholder is the Government of Paraná, which aim is to develop modernization solutions of public management aimed at continuous improvement of the services that the government provides to citizens Paraná.

5. The Next Steps

The data will be collected in state and local governments, as they are close to the day-to-day interests and needs of citizens. In order to have more robust results, I also suggest that the data should be collected in cities or states with different geographical, political, and economic situations.

References

1. Al-Debei, M. M., & El-Haddadeh, R., & Avison, D. (2008). Defining the Business Model in the New World of Digital Business, In Proceedings of the 14th Americas Conference on Information Systems, AMCIS'08, Toronto, Canada, pp. 1-11,.
2. Allison, G. (1969). Conceptual models and the cuban missile crisis. The american political science review, vol. LXIII (3:689-718), September.
3. Bertot, J. C.,Jaeger, P. T., & Hansen, D. (2012).The impact of polices on government social media usage: Issues, challenges, and recommendations. Government Information Quarterly, 29(1), 30-40.
4. Chesbrough, H. W. (2010). Business model innovation: Opportunities and barriers. Long Range Planning, 43: 354-363.
5. CGI. (2014). Comitê Gestor de Internet do Brasil. TIC domicílios e empresas 2012 - Pesquisa sobre o uso das Tecnologias de informação e comunicação no Brasil. Available: http://www.cetic.br/publicacoes/2012/tic-domicilios-2012.pdf [2014, Oct 26].
6. Crubelate, J. M., Grave, P. S., & Gimenez, F. A. P. (2005). Estratégia como Configuração: uma Versão Construtivista da Estratégia em Organizações. In: ENANPAD – Encontro Nacional dos Programas de Pós-Graduação em Administração, 29., 2005, Brasilia/SF. Anais... Rio de Janeiro: ANPAD.

7. Cunha, M. A., Annenberg, D., & Agune, R. M. (2007). Prestação de serviços públicos eletrônicos ao cidadão. In: KNIGHT, P. T.; FERNANDES, C. C. C.; CUNHA, M. A. (Org.). E-desenvolvimento no Brasil e no mundo: subsídios e Programa e-Brasil. São Caetano do Sul: Yendis, 2007. p. 559-584.

8. Cunha, M. A. Miranda, P. (2013). O uso de TIC pelos Governos: uma proposta de agenda de pesquisa a partir da produção acadêmica e da prática nacional. Organizações & Sociedade, América do Norte, 20, out. 2013. Available: http://www.portalseer.ufba.br/index.php/revistaoes/article/view/8843/6301 [2014, Mar 26].

9. Diniz, V., & Gregório, A. (2007). Do eGov governo eletrônico para o MGov Cidadania Móvel. In: KNIGHT, P. T.; FERNANDES, C. C. C.; CUNHA, M. A. (Org.). E-desenvolvimento no Brasil e no mundo: subsídios e Programa e-Brasil. São Caetano do Sul: Yendis, 2007. p. 688-702.

10. ECONOMIST. (2014). Mobile internet in emerging markets. How the mobile internet will transform the BRICI countries. The Economist Newspaper. Available: http://www.economist.com/node/16944020?story_id=16944020 [2014, Mar 17]. 11

11. Gimenez, F. A. P. (2012). A Sustentabilidade e a Inovação como um novo desafio para a Gestão das Organizações. Tese para o concurso de professor titular do Departamento de Administração Geral e Aplicada, Setor de Ciências Sociais Aplicadas, Universidade Federal do Paraná. 13

12. IBGE. (2014). Instituto Brasileiro de Geografia e Estatística. Projeção da População do Brasil e das Unidades da Federação 2013. Available: http://teen.ibge.gov.br/noticias-teen/5072-projecao-da-populacao-do-brasil-e-das-unidades-da-federacao-2013 [2014, Mar 26].

13. Jahanshahi, A.A., Khaksar, S. S.;, Yaghoobi, N., Nawaser, K. (2011). Comprehensive model of Mobile Government in Iran, Indian Journal of Science and Technology, Vol. 4, No. 9, pp.1188-1197.

14. Kuschu, I., & Kuscu, H. (2004). From E-government to M-government: Facing the Inevitable? In the proceeding of European Conference on E-Government (ECEG 2003), Trinity College, Dublin.

15. Lanza, B. B. B., & Cunha, M. A. (2013). Relations among Actors in Governmental Projects: The Case of Paraná mGov. International Journal of E-Services and Mobile Applications (IJESMA), 5(3), 25-42. DOI: doi10.4018/jesma.2013070102.

16. Meyer, A. D., Tsui, A. S.,& Hinnings, C. R. (1993). Configurational approaches to organizational analysis. Academy of Management Journal, v. 36, n. 6, p. 1175-1195.

17. Miller, D. (1987). The genesis of configuration. Academy of Management Review, v. 12, n. 4, p. 686-701.

18. Miller, D. (1996). Configuration revisited. Strategic Management Journal, v. 17, p. 505-512.

19. Mintzberger, H. (1979). The structuring of organizations (The theory of management policy), USA.

20. Mintzberger, H. (1986). The configurational school: strategy formation as an episodic process. In: Perspectives of Strategic Management.

21. Mintzberger, H. (1990). Strategy formation schools of thought. In: FREDRICKSON, J. (Ed.). Perspectives on strategic management. Boston: Ballinger, p. 105-235.

22. OECD. (2011). International Telecommunication Union. M-Government: Mobile Technologies for Responsive Governments and Connected Societies, OECD Publishing. http://dx.doi.org/10.1787/9789264118706-en.

23. Pateli, A., & G. Giaglis. (2004). A Research Framework for Analysing eBusiness models. European Journal of Information Systems 13: 302-314.
24. Peinel, G,& Rose, T. (2008). Business Models for M-Government Services. In A. Anttiroiko (Ed.), Electronic Government: Concepts, Methodologies, Tools, and Applications (pp. 2526-2541). Hershey, PA: Information Science Reference. DOI:-doi10.4018/978-1-59904-947-2.ch186.
25. Pereira, B. A. D., & Toni. (2002). D. D. A busca da racionalidade multidimensional da organização através da teoria da configuração. REAd - Revista Eletrônica de de Administração. Porto Alegre: ed. 29 vol. 8, n. 5, set-out.
26. SINGAPORE. (2014). mGovernment Collaboration and Initiatives. Available: http://www.ida.gov.sg/Collaboration-and-Initiatives/Initiatives/Store/M-Government [2014, Apr 09].
27. TELECO. (2014). Estatísticas de Celulares no Brasil. Available: http://www.teleco.com.br/ncel.asp [2014, Oct 26].

Towards Exploiting Linked Statistical Open Government Data

Areti Karamanou[a]
[a] *University of Macedonia*

Abstract. A significant part of Open Government Data (OGD) is of statistical nature and is currently available in governmental portals. Moreover, statistical OGD is published as linked data to facilitate the exploitation of dispersed data around the Web. However, the potential of linked statistical OGD data has not been fully realized yet. From one point of view, this is due to the limited and still immature efforts to create new processes, software components, models and methods that efficiently exploit linked statistical OGD so as to fully unleash its value. This paper describes the research aim of a currently immature PhD dissertation that fosters for the definition of a well-defined research question in the field of linked statistical OGD.

Keywords. Open Government Data, Linked Data, Statistical Data

1. Research aim

Public sector produces, collects, maintains and disseminates a wealth of information [11]. The opening up of this information is a priority in Europe but all over the globe as well in order to improve public services, promote public sector transparency and accountability and stimulate amongst others economic growth and innovation. This movement of opening governmental data is also known as the Open Government movement or the PSI Directive[1] and the corresponding data as Open Government Data (OGD). OGD is available on various data portals that have been developed and maintained by national, regional or local public bodies all around the world. Examples include data.gov.uk from the UK, the data.gov from the US at national level and the opendata.paris.fr from the city of Paris at the municipality level.

A significant part of OGD is of statistical nature meaning that it consists of data maintained in an organized and strictly structured form. Statistical OGD contain information regarding indicators such as e.g. the unemployment rate, the gross domestic product growth rate (GDP), the level of internet access in households or the number of people killed at road accidents and may come out of observations made in terms of many different disciplines including economics, social sciences, health and others. Statistical OGD are often modelled as data cubes by employing the RDF Data Cube vocabulary [1]. This model describes data cubes through the definition of their dimensions, measures and attributes.

At the same time, Linked Data technology has been introduced that aim at facilitating the linking of dispersed data on the Web [1]. A significant part of open

---

[1] http://ec.europa.eu/digital-agenda/en/european-legislation-reuse-public-sector-information

statistical data from international organizations (e.g. the World Bank) has already been published under the principles of Linked Data. In the context of statistical OGD, Linked Data can be used to facilitate the exploitation of OGD that are dispersed in different sources on the Web [13]. This exploitation refers to, for example, the performing of advanced data analytics on top of linked statistical OGD or the creation of high-quality visualizations out of linked statistical OGD.

However, the potential of linked statistical (OGD) has been unrealized to a large extend [1]. Although currently literature includes a number of technical solutions [12][16][6][7], processes [16] and methods [12] regarding the creation and exploitation of statistical OGD all these solutions are not able to efficiently address the complexity of OGD. In this context, this PhD thesis aims to contribute in the exploitation of statistical OGD from three different dimensions by (i) inventing, designing and describing new processes regarding linked statistical OGD (process level), 2) designing and developing software tools and components (technological level) and 3) inventing conceptual models and methods regarding the exploitation of linked statistical OGD (semantic level).

2. Methodology

The methodology adopted so far for the implementation of this PhD thesis includes two stages.

Stage 1. Preparatory stage – Understanding the scope of the dissertation. This first stage of the PhD thesis is considered to be particularly important as it targets in obtaining the necessary theoretical background of the fields of my interest. In addition, this stage aims at identifying the gap in the research evidence related to the topic that will formulate a well-defined research question. The introduction to the field of the dissertation started with a thorough literature study of the principles and technologies of open data and linked data. Relative scientific research was identified and studied in order to achieve deep understanding of the literature in the field. In order to achieve an extended and comprehensive literature review Webster and Watson's instructions were followed according to which "*authors could deal with a mature topic where an accumulated body of research exists that needs analysis and synthesis. In this case, they would conduct a thorough literature review and then propose a conceptual model that synthesizes and extends existing research*" [19]. Towards this end a concept-centric approach of literature review was conducted that was not guided by or confined to specific authors, set of journals or geographic regions, but presented a concept-centric character focusing on the key concepts that embrace the field to be reviewed. Several indexed bibliographical databases such as Web of Knowledge and Scopus were searched, ensuring the low loss of significant articles. The result of this process was a a number of scientific works that were extracted and studied with respect to methodology, subjects and methods of analysis. More importantly, this process facilitated the formulation of a comprehensive map consisting of the major topics in the field of open data and linked data. The topic that is considered as a gap in the field and that was selected for this dissertation was the exploitation of linked statistical OGD.

Stage 2. Definition of the research question. The next step is to explicitly define the exact research question that this dissertation comes to answer. The results of the previous step will lead the creation of a well-defined research question that will be the subject of this PhD thesis.

3. Preparatory stage

Statistical OGD can be accessed from the various online portals that have been launched and maintained by many governments and other public authorities around the globe. Table 1 presents a list of portals at an international, continental, national, municipality or state (of the US) level.

Table 1. Examples of OGD portals at different levels of public administration.

Portal URL	Level of Public Administration	Region
data.gov.uk	national	UK
data.gov	national	USA
data.gouv.fr	national	France
Open Data Portal[2]	continental	Europe
opendata.paris.fr	municipality	Paris
dati.gov.it	national	Italy
data.go.jp	national	Japan
data.un.org	international	International
data.maryland.gov	state	State of Maryland
publicdata.eu	continental	Europe

According to the Linked Data principles, Linked Data refers to machine readable data published on the Web that have explicitly defined meaning, it is linked to data sets located in external sources and, vise-versa, can be linked to from data sets located in external resources [1]. At the moment the Linked Data Web provides a compendium of over 300 datasets containing over 30 billion RDF triples, which are interlinked by over 503 million RDF links. Literature shows two main vocabularies that can be used to model statistical data as data cubes i.e. the RDF data cube vocabulary and the Open Cubes vocabulary [7], with RDF being the most popular between them. The RDF data cube vocabulary is also currently a W3C standard for modeling structured data with multiple dimensions including statistical data as RDF. RDF data cube vocabulary also allows the "RDFising" of statistical data according to the Linked Data principles and in a way that it can be linked to related data sets and concepts. According to the RDF data cube vocabulary, the basic structure elements (collectively called components) of a data cube are its (i) dimensions, (ii) measures and (ii) attributes. The dimensions of a data cube are used to identify a specific observation of the data cube. Representative examples of data cube dimensions are the time instance or time period that an observation applies to or the geographical area covered by the observation. Dimensions typically represent concepts taken from a code list so as to be used across different data cubes. Code lists are usually identified using the SKOS vocabulary. The measure components of a data cube represent the phenomenon (or phenomena) under observation such as, for example, an indicator that the statistical dataset describes. Finally, the attribute components of a data cube facilitate the interpretation of the observations of the cube. For example, an attribute component can be employed to express the unit of a measure component e.g. percentage.

A literature review was also conducted so as to identify processes, tools, models and methods that currently exist at the three levels that constitute the scope of this PhD dissertation i.e. the process level, the technological level and the semantic level.

---

[2] http://open-data.europa.eu

Regarding the *process level*, a recent work in literature identifies three major phases in the linked data lifecycle i.e. the Publish, Expand and the Reuse phase [16]. The Publish phase, further consisting of three steps, involves the creation of linked data out of raw data. The Expand phase that consists of two steps regards the expansion of data cubes by linking them with other data cubes. Finally, the Reuse phase that consists of three steps regards the exploitation of data cubes. Additional processes in literature include Hyland and Wood's OGD seven-step linked data lifecycle [8], Auer et al.'s eight-step linked data lifecycle [1] and Villazon-Terrazas's five-step linked data lifecycle [18].

Literature also reveals a number of software components and tools at the *technological level* that have been developed for the publishing and reuse of data cubes. Table 2 presents an extensive list of such tools, widgets and components. Examples include Capadisli et al.'s tool [4] for transforming statistical data from SDMX-ML format to Linked Data as well as Salas et al.'s tool [17] for creating linked data from CSV files and OLAP databases. Moreover, the OpenCube toolkit [14][15] that has been developed during the lifetime of the European funded project OpenCube offers tools for each of the phases of the lifecycle. In particular, the author is one of the mentors, designers and developers of the OpenCube MapView component described in detailed in the next section.

Table 2. A list of software tools for OGD publishing and re-use.

	Software tool/component	Phase of the lifecycle	Purpose
1	OpenRefine RDF Extension[3]	Publish	For creating and exporting RDF.
2	TARQL extension for data cubes[4]	Publish	Converts CSV files to RDF.
5	D2RQ extension for data cubes[5]	Publish	Converts relational tables to RDF.
13	R2RML extension for data cubes[6]	Publish	Integration of data from tabular legacy sources (such as relational databases and CSV files) into the RDF repository as data cubes.
14	OpenCube JSON-stat2qb extension for data cubes[7]	Publish	Automatic conversion from JSON-stat format to RDF using Data Cube vocabulary.
15	Ontowiki's CSVImport extension[8]	Publish	Converts excel sheets as RDF using the RDF Data Cube vocabulary.
16	Ontowiki's OLAP2DataCube extension[9]	Publish	Transforms OLAP databases into RDF.
17	CSV2RDF4LOD[10]	Publish	Production of RDF encoding out of Comma-Separated-Values (CSV)
18	TabLinker[11]	Publish	Converts manually annotated Microsoft Excel workbooks to the RDF Data Cube vocabulary.

---

[3] http://openrefine.org/

[4] http://opencube-toolkit.eu/tarql-extension-for-data-cubes/

[5] http://opencube-toolkit.eu/d2rq-extension-for-data-cubes/

[6] http://opencube-toolkit.eu/r2rml-extension-for-data-cubes/

[7] https://github.com/opencube-toolkit/json-stat2qb

[8] http://aksw.org/Projects/CSVImport.html

[9] https://github.com/AKSW/olapimport.ontowiki

[10] http://logd.tw.rpi.edu/technology/csv2rdf4lod

[11] https://github.com/Data2Semantics/TabLinker

19	Linked Media Framework[12]	Publish	Publishes Excel, CSV and XML files as Linked Datas
20	DB2RDF[13]	Publish	Converts data from relational data model to semantic data model (in RDF and RDFS).
21	D2RQ[14]	Publish	Allows the access of relational databases as virtual, read-only RDF graphs.
22	Karma[15]	Publish	Enables the integration of data from a variety of data sources including databases, spreadsheets, delimited text files, XML, JSON, KML and Web APIs.
23	Datalift[16]	Publish	Publish databases, CSV, XML, RDF, RDFa, GML, Shapefile as Linked Data.
11	OpenCube expander[17]	Expand	To search for compatible cubes and create a new expanded cube by merging two compatible cubes.
12	OpenCube compatibility explorer[18]	Expand	Identify compatible cubes for potential merge and establish typed links between cubes.
3	OpenCube MapView[19]	Reuse	Enables the visualization of RDF data cubes on a map based on their geospatial dimension.
4	Grafter[20]	Reuse	Converts tabular data to RDF.
6	Interactive chart visualization widgets[21]	Reuse	Visualizes RDF Data Cube slices.
7	OpenCube R statistical analysis module[22]	Reuse	To apply statistical analysis methods to data represented as RDF data cubes.
8	OpenCube aggregator[23]	Reuse	Its main role is to compute aggregations of existing cubes using an aggregate function.
9	OpenCube OLAP browser[24]	Reuse	A browser for RDF data cubes.
10	Data catalogue management solution[25]	Reuse	A UI for managing catalogues with multiple datasets.
24	SILK[26]	Reuse	Integrates heterogeneous data sources
25	LIMES[27]	Reuse	Creates links

Finally at the *semantic* level, a number of models and methods exist in literature regarding linked statistical OGD. The RDF data cube vocabulary described above is one of the most important contributions to this level but other significant efforts exist as well. For example, the Linked Statistical Models (limo) [12] is a vocabulary created

---

[12] https://bitbucket.org/srfgkmt/lmf/
[13] http://sourceforge.net/projects/db2rdf
[14] http://d2rq.org/
[15] http://www.isi.edu/integration/karma/
[16] http://datalift.org
[17] http://opencube-toolkit.eu/opencube-expander/
[18] http://opencube-toolkit.eu/opencube-compatibility-explorer/
[19] http://opencube-toolkit.eu/opencube-map-view/
[20] http://grafter.org/
[21] http://opencube-toolkit.eu/interactive-chart-visualization-widgets/
[22] http://opencube-toolkit.eu/r-statistical-analysis-module/
[23] http://opencube-toolkit.eu/opencube-aggregator/
[24] http://opencube-toolkit.eu/opencube-browser/
[25] http://opencube-toolkit.eu/data-catalogue-management-solution/
[26] http://wifo5-03.informatik.uni-mannheim.de/bizer/silk/
[27] http://aksw.org/Projects/LIMES.html

to allow the description of statistical and data mining models in the RDF model. As it will be described in the next section, the author had an active role towards the invention of the limo vocabulary. There are also additional standards such and RDF Schema vocabularies closely related to statistical data as for example the Predictive Modelling Markup Language (PMML) [20] XML standard and the DDI-RDF vocabulary [2] that focuses on raw record-level datasets and describes their structure. Finally, Kalampokis et al. [12] introduce the concept of data analytics on top of distributed linked statistical OGD and describe the technical prerequisites needed to enable effective exploitation of linked statistical OGD in data analytics.

4. Progress so-far

Regarding the progress of this PhD dissertation and towards the understanding of the general domain the author has thoroughly studied the relevant research work found in literature as described in the section above.

Regarding the technological level, as a first step the identified software tools and components described in the preparatory section were identified and tested in order to understand their functionalities. Moreover, as mentioned in the previous section, the author has also significantly contributed in the inspiration, design and development of the MapView tool component of the OpenCube toolkit [14][15]. In particular, the OpenCube MapView component allows the visualization of RDF data cubes on maps according to their geospatial dimension. OpenCube MapView component implements three different types of maps:

- marker map that uses markers to show the values of the measure of observations in different locations on the map.
- bubble map that creates bubbles having seize proportional to the value of the measure of observations in different locations on the map.
- choropleth map that represents locations on the map using polygons and colors its polygon using grater color density for locations with higher measure values.

Figure 1 and Figure 2 illustrates examples of instances of a marker and choropleth map of the OpenCube Mapview component.

In addition, regarding the semantic level, existing models and methods have been thoroughly studied. In addition, the author is one of the creators of limo, the open vocabulary for the description of predictive models according to the linked data principles [12]. limo allows the description of statistical and mining models in the RDF model hence enabling their incorporation on the linked data Web as well as their linking to other resources (datasets, people, organizations and articles).
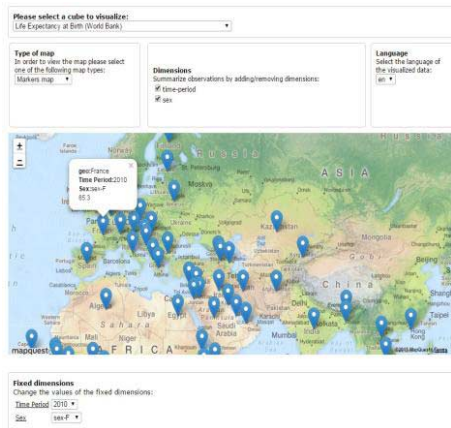
Figure 1. An instance of the OpenCube
MapView marker map.

Figure 2. An instance of the OpenCube
MapView choropleth map.

5. Defining the research question

This stage of the dissertation involves the definition of the exact research question that this PhD thesis comes to answer. As already stated, the current vision of the PhD thesis is to contribute to the following areas of interest in the domain of the exploitation of linked statistical OGD:

- *Semantic level*. This thesis aims to inspire, expand and create models for the description and exploitation of linked statistical OGD. These may include, for example, the expansion of limo so as to describe in detail any type of predictive models or the expansion of the RDF data cube model so as to face possible particularities of linked statistical OGD.

- *Application level*. This regards the technical part of the PhD thesis. In particular at the technical level, this thesis envisions to inspire, design and develop tools that facilitate the easy and efficient exploitation of linked statistical OGD so as to unleash its full potential.

- *Process level*. The vision is to suggest new processes relative to statistical linked OGD.

References

[1] S., Auer, L., Bühmann, C., Dirschl, O., Erling, M., Hausenblas, R., Isele, R., ... & H., Williams (2012). Managing the life-cycle of linked data with the LOD2 stack. In The Semantic Web–ISWC 2012 (pp. 1-16). Springer Berlin Heidelberg.

[2] T., Bosch, R., Cyganiak, A., Gregory, & J., Wackerow (2013). DDI-RDF Discovery Vocabulary: A Metadata Vocabulary for Documenting Research and Survey Data. In LDOW.

[3] C., Bizer, T., Heath, T., Berners-Lee: Linked data - the story so far. International Journal on Semantic Web and Information Systems 5(3), 1–22 (2009).

[4] S., Capadisli, S., Auer & A.C.N. Ngomo (2013). Linked SDMX data. *Semantic Web Journal*, 1-8.

[5] R. Cyganiak and D. Reynolds, "The RDF data cube vocabulary: W3C recommendation," W3C, Tech. Rep., January 2014.

[6] I., Ermilov, I., M., Martin, J., Lehmann, J., & S., Auer, (2013). Linked open data statistics: Collection and exploitation. In *Knowledge Engineering and the Semantic Web* (pp. 242-249). Springer Berlin Heidelberg.

[7] L., Etcheverry & A.A. Vaisman (2012). Enhancing OLAP analysis with web cubes. In *The Semantic Web: Research and Applications* (pp. 469-483). Springer Berlin Heidelberg.

[8] B., Hyland, B. and D., Wood (2011). The Joy of Data - A Cookbook for Publishing Linked Government Data on the Web. Linking Government Data. D. Wood, ed. Springer. 3--26.

[9] V., Janev, V., Mijović, S. & Vraneš, (2013). Lod2 tool for validating rdf data cube models. In *Web Proceedings of the 5th ICT Innovations Conference* (pp. 12-15).

[10] M., Janssen, Y., Charalabidis & A., Zuiderwijk, Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, **29**(4), (2012), 258-268.

[11] E. Kalampokis, E. Tambouris and K. Tarabanis, Open Government Data: A Stage Model, M. Janssen et al. (Eds): *EGOV2011*, LNCS 6846, 235-246, Springer, 2011.

[12] E. Kalampokis, A. Karamanou, E. Tambouris and K. Tarabanis (2013) Towards a Vocabulary for Incorporating Predictive Models into the Linked Data Web, in the *1st International Workshop on Semantic Statistics (SemStats 2013) in conjunction with the 12th International Semantic Web Conference (ISWC2013)*, Sydney, Australia.

[13] E. Kalampokis, E. Tambouris and K. Tarabanis (2013) Linked Open Government Data Analytics, M.A. Wimmer, M. Janssen, and H.J. Scholl (Eds.): *EGOV 2013, LNCS 8074*, pp. 99-110. IFIP International Federation for Information Processing.

[14] E. Kalampokis, A. Nikolov, P. Haase, R. Cyganiak, A. Stasiewicz, A. Karamanou, M. Zotou, D. Zeginis, E. Tambouris, K. Tarabanis (2014) Exploiting Linked Data Cubes with OpenCube Toolkit, *Proc. of the ISWC 2014* Posters and Demos Track a track within 13th International Semantic Web Conference (ISWC2014), 19-23 October 2014, Riva del Garda, Italy, CEUR-WS Vol.1272.

[15] E. Kalampokis, A. Karamanou, A. Nikolov, P. Haase, R. Cyganiak, B. Roberts, P. Hermans, E. Tambouris, K. Tarabanis (2014) Creating and Utilizing Linked Open Statistical Data for the Development of Advanced Analytics Services, *Proc. of the 2nd International Workshop on Semantic Statistics (SemStats2014*) in conjunction with the 13th International Semantic Web Conference (ISWC2014), 19-23 October 2014, Riva del Garda, Italy, CEUR-WS proceedings.

[16] E. Tambouris, E. Kalampokis, K. Tarabanis (2015) Processing Linked Open Data Cubes, E. Tambouris, M. Janssen, H. J. Scholl, M. Wimmer, K. Tarabanis, M. Gascó, B. Klievink, I. Lindgren, and P. Parycek (Eds.): *EGOV2015, LNCS 9248*, IFIP [accepted for publication].

[17] P. Salas, M. Martin, F. Mota, S. Auer, K. Breitman, and M. Casanova, Olap2datacube: An ontowiki plug-in for statistical data publishing, in *Developing Tools as Plug-ins (TOPI), 2012 2nd Workshop on*, June 2012b, pp. 79-83.

[18] B., Villazón-Terrazas, L.M., Vilches-Blázquez, O., Corcho, & A., Gómez-Pérez (2011). Methodological guidelines for publishing government linked data. In Linking government data (pp. 27-49). Springer New York.

[19] J. Webster and RT Watson, Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quart* **26** (2002), 13-23.

[20] D., Wettschereck, & S., Muller (2001). Exchanging data mining models with the predictive modelling markup language. In International Workshop on Integration and Collaboration Aspects of Data Mining, Decision Support and Meta-Learning.

Enhancing students' skills and capabilities to exploit Open Government Data

Maria ZOTOU[a,1]
[a] *University of Macedonia, Thessaloniki, Greece*

Abstract. The current ever-changing world causes economies, trends, technologies and professional domains to constantly shift and transform. This requires that workforce acquire cognitive and transversal skills (e.g. critical thinking, problem solving, ICT usage, capacity to learn) from an early age, i.e. during formal education, so that they will be equipped to properly handle and adapt to these changes and increase their participation in societal or political processes. However, the existing student-centred learning strategies such as Problem Based Learning (PBL) and the utilization of new technologies (e.g. Learning Analytics, Linked Open Data, Semantics etc) often do not show the expected benefits. In the context of this thesis, these issues will be studied and addressed. More specifically, PBL will be extended so as to allow such skills development, inventive tools and methods usage and connection with the web of data. Towards this goal, two main domains are explored in depth: Project Management and learning about Open Government Data (OGD). This paper will focus on the OGD domain.

Keywords. Problem-based Learning, Learning Analytics, Open Data

1. Research aim

Current worldwide conditions include rapid generation of data, information and technologies and the continuous emergence of new professional fields. This requires competent workforce that will be equipped with cognitive and transversal skills, i.e. skills that they can use in order to adapt to these changes, absorb and comprehend all the new information, utilize the novel technologies, transfer across professional domains and contribute to economic growth and boost societal benefits ([1], [10], [5]). Towards this goal, participatory learning pedagogies should be incorporated during education, where as students they can learn how to communicate their ideas, collaborate with their peers and witness in practice how their opinion makes a difference.

A well established learning strategy that allows the development of such as skills is Problem Based Learning (PBL), as it enables active participation of students who "learn by doing". Furthermore, the emergence of Web 2.0. brought on the new version of PBL, namely PBL 2.0., where learners use the newly developed tools (e.g. forum, chat, blogs, wikis etc) in the learning process. However, the utilization of PBL in this present mode has reached its maximum potential. There is a need for this very promising model to be renovated and exploit novel opportunities and technologies that will unleash new benefits and capabilities, i.e. Learning Analytics (LA), semantics, Linked Open Data (LOD) etc.

---

[1] Corresponding Author.

Thus, the main research aim of this thesis is to provide tools and learning methods that will ensure a more up-to-date and connected learning solution and enhance students' transversal skills by:

1. Improving the existing learning strategy of PBL and utilizing novel methods and tools such as LA, Semantics etc. This will bring forth a new data-driven PBL paradigm that will allow students to be guided through the learning process with specific steps and learn to use ICT tools actively.

2. Utilizing novel technologies such as Web 2.0, cognitive tools, learning semantics, Linked Open Data (LOD) etc. These new technologies will allow activities such as semantic annotation of learning resources with descriptions, linkage of learning resources with EU standards, e.g. the ESCO ontology [19], enrichment of the vocabularies that currently exist to describe education concepts etc. It is essential for education to connect with the web of data, and participate in providing meaning to data, enabling resources discoverability and enriching knowledge utilized by observing the worldwide developments.

An important issue for the development and validation of the work carried out is the testing of the developed techniques and tools in domains that encourage active participation. Such a domain is Open Data (OD) and more specifically Open Government Data (OGD). This domain has promised to provide a variety of benefits to society such as to stimulate innovation and contribute to economic growth. Furthermore, OGD can empower citizens for meaningful participation in more informed decision making, influencing public decisions and participating in a balanced dialogue with the governments and public authorities. Citizens' ability to discover OGD, understand it and exploit it ensures accountability for public authorities and the creation of enhanced public services and data. However, although a number of recently launched applications have been developed by reusing OGD[2], academic studies have shown that the full potential of this data has not yet been realized. One of the main reasons is that capacity-oriented education and training on OGD and relevant technologies such as Linked Open Data (LOD) in Europe is still limited.

Another important domain regards Project Management (PM), where students are required to create a project that solves a specific problem by actively participating in all the different PM phases (e.g. analysis, design, implementation, evaluation).

This research aims to perform pilots by applying the developed tools and strategies in these two domains, i.e. PM and OGD education and training. The current paper will focus on providing details on the OGD pilot. The incorporation of the work done in OGD courses will unleash many benefits for citizens who will be equipped to:

1. exploit OGD and create applications and services that address existing societal, economic or political needs

2. identify emerging societal, economic or political needs and participate actively in finding a solution

3. enrich and publish OGD for various domains

4. use ICT tools efficiently and participate in activities such as digital democracy, use of OGD in eParticipation initiatives etc.

This fusion of the "well tested but old" with the "new but hectic" will allow a more evidence-based approach when integrated with different educational domains.

---

[2] http://www.city-data.com/, http://www.civicplus.com/, http://energyiq.lbl.gov/, http://www.icitizen.com/

2. Method

To ensure the quality of the research work carried out in the context of this thesis as well as the rigor of its results, we draw upon the design science research methodology [11]. The steps of the methodology are defined as shown in Figure 1.
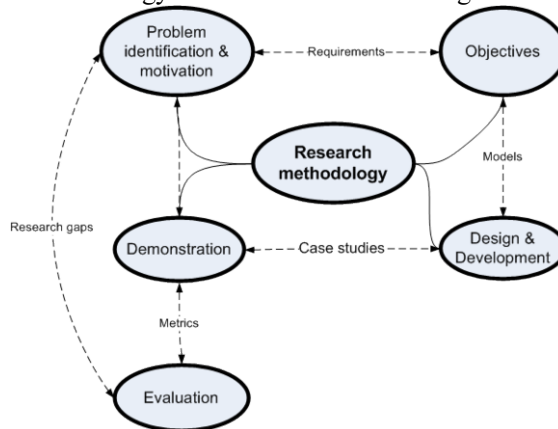
Figure 1 Research methodology

Each step regards a significant milestone in the thesis's work.

- Problem identification and motivation. This step regards the process of defining the problem that the thesis will address and the justification of the solution that is proposed to address it. This aims to motivate the research work to be carried out and to trust the results that will be produced.
- Objectives. This step structures the work's objectives, both qualitative and quantitative.
- Design and development. This step specifies and develops the different artifacts of the proposed solutions. These artifacts can be either models, frameworks, architectures etc.
- Demonstration. This step refers to the usage of the developed artifacts for solving the defined problem through a series of well organized case studies in specific domains, e.g. OGD or PM.
- Evaluation. This final step of the research methodology assesses to what degree the artifacts addressed and solved the initial problem and to what extent they managed to fulfill the set objectives.

It should be noted that the methodology follows an agile approach, where various interval rounds of assessment are carried out in order to improve the internal processes of each step as well as all operations.

3. Background work

3.1. Data-driven Problem based Learning

The PBL learning strategy shifts the focus from understanding common knowledge to developing new knowledge through "learning by doing" activities and accommodates active participation of students. A very commonly used PBL model was

developed by Aalborg University [8], which consists of 7 well-defined steps, as shown in Table 1. The table also depicts a mapping between each PBL step with the type of work required by the learners and with examples of supporting technology.

Table 1. PBL Aalborg model steps and supporting technology

PBL step	Work	Supporting technology
Clarification of terms and concepts	Group	Forum Resources Web research Glossary Self reflection questionnaire Interim quiz
Problem definition	Group	Thematic blocks Forum Self reflection questionnaire Interim quiz
Problem analysis	Group - Brainstorming	Forum Notes Mind map Tasks chain Self reflection questionnaire Interim quiz
Inventory of problems and solutions	Group – Discussion	Mind map Research Resources Self reflection questionnaire Interim quiz
Definition of self-study learning objectives	Group	Checklist Self reflection questionnaire Interim quiz
Private study	Individual	Journal Notes Self reflection questionnaire Interim quiz
Conclusions – Synthesis	Group	Forum Wiki – Collaborative report Self reflection questionnaire Final quiz

Additionally, the "Opening up Education" Agenda[3] highlights the benefits of using technologies in education, which supports learning with no cultural, social or geographical boundaries. The usage of ICTs however, should be structured in order to ensure the development of proper skills. This unleashes a number of benefits, such as:

- Enables learners to become more familiarized with the different types of ICT tools they can utilize during class but also in the future.
- Allows learners to develop transversal skills such as critical thinking, problem solving, communication of ideas, collaboration etc, that will transform them into participatory and engaging citizens.
- Enables learners to get an active foretaste on how their actions and participation in solving a problem or creating a project makes a difference.

Subsequently, the utilization of a pedagogical model that allows students to participate more actively as well as the incorporation of technologies that support

---

[3] European Commission, 'Opening up Education' - Making the 21th century classroom a reality, 2013. http://europa.eu/rapid/press-release_SPEECH-13-747_en.htm

students' activities lead to the generation of multiple types of data ([17], [18], [7]). The exploitation and analysis of the data can be carried out through a data-driven learning approach, where data is gathered and analyzed with specialized methods and technologies such as LA ([12], [3], [13]). This approach will take into consideration the whole LA lifecycle:

- *data gathering*, by identifying educational data generated in all steps of PBL
- *information processing*, by analyzing processes and techniques that transform educational data to meaningful, multi-modal information
- *knowledge application*, by identifying all intervention mechanisms that could be put into practice based on LA feedback during course design and execution.

3.2. Pilot: Open Government Data

A recent study by the McKinsey Global Institute estimated "the global annual economic potential value of Open Data to \$3 trillion" [9] whereas another study by Deloitte measures the value of Open Data to consumers, business and the public sector in terms of \$ billions[4]. The successful publication and re-use of Open Data is also one of EU's main priorities[5]. The inventor of the World Wide Web stated that "opening up data is fundamentally about more efficient use of resources and improving service delivery for citizens." [2]. This has led to national and European initiatives that have brought forth abundant datasets from different fields (government, statistics, finance, environment, health, transport etc)[6] paving the way for the creation of new services and the growth of new markets[7]. However, these opportunities require competent workforce with in depth knowledge on Open Data as well as high-level skills that will enable them to identify existing societal, political and financial needs as well as identify OD and create innovative applications that can address these needs. The ongoing and gradually increasing release of OGD has created significant added value, including:

- Transparency, accountability and democratic monitoring
- Citizens participation
- Improved and efficient new private and public products and services

One of the most important benefits of OD in general is the underpinning of the collective intelligence of the public. The main idea is that, under the right circumstances, groups can create better alternative solutions and result to better decisions [14]. Thus, when the public is equipped to collaboratively access and re-use OD as well as publish OD for different domains, it can significantly and fundamentally alter and improve the political, financial and societal ecosystems.

However, the openness of data faces various challenges [4]. Some of them are:

- Lack of skills on tracking and assessing the relevance of data
- Lack of knowledge of citizens on the different possible uses of data
- Lack of motivation and incentive to re-use or publish data

These challenges indicate a prominent need for education opportunities in the field, addressing both lifecycles of the OGD domain, i.e. re-use and publication of OGD. Currently, EU and expert organizations carry out training initiatives mostly as seminars

---

[4] Market Assessment of public sector information, Department for Business Innovation & Skills, 2013. https://goo.gl/bRJUwA

[5] Digital Agenda for Europe: Open Data*, 2011. http://ec.europa.eu/digital-agenda/en/open-data-0

[6] http://data.gov.uk, https://open-data.europa.eu/en/data, http://publicdata.eu/, http://data.gov.gr

[7] Open data An engine for innovation, growth and transparent governance, 2011. http://goo.gl/gEJ1XT

and workshops that usually include content in slides and occasionally quiz tests that allow self-reflection (e.g. Europa portal, OD Support, Open Knowledge Foundation etc). However, these are not structured based on pedagogical models that would foster participatory activities and usually only provide content that learners passively study. Thus, a new paradigm of innovative OGD education and training is required, based on novel pedagogies and advanced technologies.

4. Expected results

In the context of this thesis, issues will be tackled that together can solve one major problem; the integration of PBL with LA tools as well as the utilization of novel technologies such as semantics, LOD etc to connect education with the web of data. Finally, the thesis will launch the developed outputs in OGD and PM courses and assess the results.

4.1. Results so far

Currently, the OGD lifecycles have been thoroughly studied and approximately 300 slides of learning content have already been created to cover all aforementioned steps in depth. This content has been used in OGD higher education courses and has led to over 200 students discovering, transforming and linking OGD as well as creating a number of interesting and innovative visualizations and applications by re-using OGD. Figure 2 shows a representative example of a web-based graph created with OGD that depicts the number of students and employees in primary, secondary and tertiary education institutes in Thessaloniki, Greece. This visualization changes dynamically when the stored data is updated and thus provides a dynamic overview of the city's educational institutes. This could alert citizens of interesting phenomena, e.g. if a school with a large number of students does not employ the appropriate number of employees, and citizens can take appropriate measures based on this information. For now this is being carried out traditionally, with no support of PBL or any novel tools.
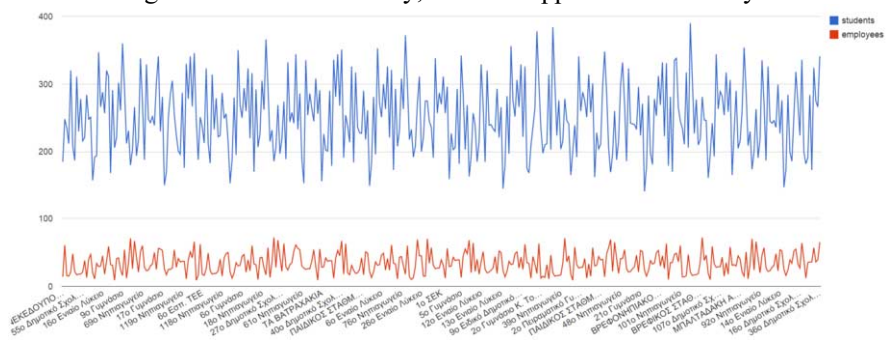
Figure 2 Example of visualization using OGD

On the other hand, the PBL model and ICT and LA tools have been utilized in other project-oriented higher education courses such as Project Management [10]. The following Figure shows feedback based on an LA tool integrated in the e-learning platform used for the PM course, showing statistical data on the access for each learning resource.
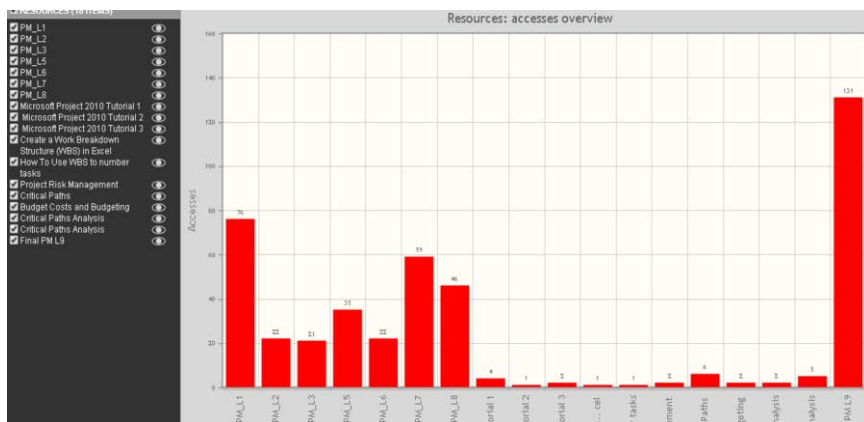
Figure 3 usage of LA tools in PM course

A list of quantitative and qualitative indicators have been formulated that depict the results of the work done so far. Each indicator serves as an evaluation metric to measure the success of the research.

Table 2 List of evaluation indicators

Thesis outcome	Quantitative indicators	Qualitative indicators
State of the art on PBL	- Number of PBL models studied - Number of concepts in chosen PBL model	-Results provide an overall overview of PBL steps and concepts -Results provide information on the transversal skills that are fostered with PBL
State of the art on OGD	- Number of scientific papers studied and stakeholders reached - Number of concepts identified	-Results identified provide an overall overview of OGD lifecycles steps and corresponding concepts
Educational materials for OD publication and re-use	- Number of units of learning - Number of content modes (e.g. online pages, slides, videos etc)	- Content covers all phases of both publication and re-use OD lifecycles - Content is provided in many modes
e-learning platform	- Modules installed - Number of registered participants - Collaboration tools installed	- Is active and online. - Is based on existing technologies - Includes LA functionalities
LA tools and techniques for course re-design	- LA tools installed - Data gathered. - Number of LA tools configured / developed	- Allow data gathering. Provide visualization and statistical analysis opportunities. - Guidelines for course re-design.
Educational activities	- Number of educational activities - Number of university students participating - Number of courses	- Allow participatory learning. - Provide OGD content. -Allow practical experimentation with existing OGD. -Increased interest of students in active participation. - Increased students' interest in OGD and its benefits.

4.2. Next steps

This thesis aims to combine the know-how of both cases and to result in a new paradigm of data-driven OD_PBL education. Towards this goal, the OGD re-use and publication lifecycle will be mapped across each PBL step during the learning process.

Each step will also be supported with the corresponding ICT tools and LA-specific tools in order to allow active participation and gather all generated data for analysis. This will provide informative insights on students' progress, support adaptive learning pathways and allow helpful content or suggestions.

Finally, the implementation and evaluation results will lead to the suggestion of future work that needs to be tackled in order to solve any issue that occurred and improve the identified benefits.

References

[1] Council of the European Union, *Messages from the EYC Council in the field of education as a contribution to the discussion on the post-2010 Lisbon Strategy*, 2009.
[2] European Public Sector Information Platform, *Open Data and EU Funding*, 2013. http://goo.gl/rfwKjS
[3] Ferguson, R. *The state of learning analytics in 2012: A review and future challenges*, 2012. Available at: http://kmi.open.ac.uk/publications/techreport/kmi-12-01.
[4] Jansen, M. *Benefits, Adoption Barriers and Myths of Open Data and Open Government*, Information System Management, 2012.
[5] Johnson, L., Smith, R., Willis, H., Levine, A. and Haywood, K. The 2011 Horizon Report. Austin, Texas.
[6] Kalampokis, E., Tambouris E. and Tarabanis K. *A Classification Scheme for Open Government Data: Towards Linking Decentralized Data*, International Journal of Web Engineering and Technology, Vol.6, No.3, pp.266-285, 2011.
[7] Lee, Y. J. *Developing an efficient computational method that estimates the ability of students in a Web – based learning environment*. Computers&Education,58(1),579-589, 2012.
[8] Mair, Michael; Halkier, Henrik. *Problem based-learning and life-long learning: Principles, applications, challenges*, report from Aalborg Universitet, 2013. Available at: http://goo.gl/VAiOUq
[9] Manyika, J. *Open data: Unlocking innovation and performance with liquid information*, McKinsey Global Institute, 2013.
[10] Panopoulou E., Tambouris E., Sanchez-Nielsen E., Zotou M. and Tarabanis K. *Learning from eParticipation initiatives of regional and local level authorities in Greece and Spain*. Journal of Balkan and Near Eastern Studies, 13:1, pp. 77-96. DOI: 10.1080/19448953.2011.550805, 2011.
[11] Peffers, K., Tuunanen, T., Rothenberger, M. & Chatterjee, S. *A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems*, 24, 2008.
[12] Phillips, R., Maor D., Preston G. and Cumming-Potvin, W. *Exploring Learning Analytics as Indicators of Study Behaviour. In: World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 2012.
[13] Siemens, G. *What are Learning Analytics?*, 2010. http://goo.gl/3iW7UI
[14] Surowiecki, J. *The Wisdom of Crowds*: *Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations,* Little, Brown ISBN 0-316-86173-1, 2004.
[15] Tambouris, E., Zotou, M and Tarabanis, K. *Towards designing cognitively-enriched project-oriented courses within a blended problem-based learning context*, Education and Information Technologies, p. 1-26, doi: 10.1007/s10639-012-9209-9, 2012.
[16] Tambouris, E., Zotou M. and Tarabanis, K. *Towards designing cognitively-enriched project-oriented courses within a blended problem-based learning context*, Education and Information Technologies, p. 1-26, doi: 10.1007/s10639-012-9209-9, 2012.
[17] Timonidou, V., Zotou, M., Tambouris, E. and Tarabanis. K. *A review on Facebook as a user-generated content environment in tertiary education*. Int. J. Media and Interactive Learning Environments, 2013.
[18] Wu, B. and Wang, M. *Integrating problem solving and knowledge construction through dual mapping*. Knowledge Management & E-Learning:An International Journal, 4(3), 248–257, 2012.
[19] Zotou, M., Papantoniou, A., Kremer, K. Peristeras, V. and Tambouris E. *Implementing "Rethinking Education": Matching Skills Profiles with Open Courses through Linked Open Data technologies*. Bulletin of the IEEE Technical Committee on Learning Technology, Volume 16, Number 4, 2014.

Posters

This page intentionally left blank

Investigation and Evaluation of the Usage of Social Media in Saudi Government 2.0

Reemiah ALOTAIBI

School of Computing and Creative Technologies, Leeds Beckett University, UK

Abstract. Introducing Government 2.0 which simply translates as employing social media platforms in the government communication system is a worldwide trend. Despite the fact that more than half of the Saudi population are users of social media platforms [1] this wide acceptance of social media networks and tools among Saudi population could be a great opportunity for governmental agencies to communicate with the citizens through these tools. Evidence from the literature [2; 3; 4] showed that Saudi governmental institutions are not fully adopting these platforms to engage the citizens. Little research efforts have been done to evaluate the progress of social media adoption in general. The overall aim of this research is to investigate and evaluate the usage of social media in Saudi government. To conduct this research project, mixed methods were used. The web analysis conducted by the researcher has shown that minority of Saudi Ministries are effectively employing social media platforms. Moreover, the results from the interviews conducted with the Saudi e -government staff used CATWOE Analysis tool to explain the state of Saudi Government 2.0 implementation. The main factor for the gap between usages of social media in Saudi government agencies, which is affecting social media adoption, is that each government agency has its vision and strategic plan for their e-government transformation to fulfil the government goals and initiatives. Future research will be done to explore more factors affecting social media adoption in Saudi government 2.0.

Keywords. Social media, adoption, Government 2.0, Government' agencies, Saudi e -government.

1. Introduction

Research problem and Research aim:

Despite the fact that more than half of the Saudi population are users of social media platforms [1] this wide acceptance of social media networks and tools among Saudi population .The literature [2; 3; 4] showed that Saudi government agencies are not fully adopting these platforms to engage the citizens. The overall aim for this research is to investigate the usage of social media in Saudi government 2.0.

Research Methodology:

Mixed methods approach was deemed appropriate for this study .The qualitative with semi-structured interviews allow us to explore in depth with Saudi e -government staff (CATWOE Analysis tool used to explain the state of Saudi Government 2.0 implementation) and the quantitative research with Web analysis conducted by the researcher to investigate the current usage (Saudi Ministries as a sample of government agencies). The researcher used several indictors for the 3 most popular social media tools among Saudi government agencies.

Results:

The web analysis conducted by the researcher has shown that there are three levels of social media usage among the different Saudi Ministries, some are actively used to communicate with citizens, and others use it for information distribution only or do not use it at all. According to the analysis of the collected data out of 23 ministries only 16 Ministries use some forms of social media tools.

Ministry	Social Media tool	Like/Followers/ subscribers		Up to date	Direction of communications	Citizen to government Interactivity	government to citizen Interactivity
		June 2014	December 2014				
Ministry of Interior	Facebook	286,478	502,009	Yes	Two ways	Yes	Yes
	Twitter	128,000	193,767	Yes	Two ways	Yes	Yes
	YouTube	9,147	12,467	Yes	One way	No	No
Ministry of Transport	Facebook	305	408	No	Two ways	No	No
	Twitter	1,956	3,104	No	Two ways	Yes	No
	YouTube	…….	30	No	One way	Yes	No

Table1: result to Ministry of Interior Transport ministries as example Saudi government ministries

CATWOE Analysis tool: is a Soft Systems Methodology tool used to set up a precise and comprehensive root definition to establish the base for solving issues with various readings [5]. The main reason for such differences in social media usages by Saudi government agencies is that each government agency has its own vision and strategic action plan for its e-government transformation to fulfil the e- government goals and initiatives.

Letter	Stands for	Saudi government
C	Clients	Citizens , Private sector and Other government agencies
A	Actors	The actors in the Saudi case can be divided into two classes: **The decision makers** are: Yasser (the main government body responsible for e-Government), Ministry of communications, and Ministry of planning. **Policy executers**: these are the actual government agencies responsible for choosing specific strategies to fulfil the government's goals and initiatives.
T	Transformation	The process of moving towards government 2.0 requires a change in the mentality, processes, and tools of both the public and the staff in the different governmental agencies .The style of communication with the citizens should also be changed to a more democratic one. Based on the initial investigation on achieving the full outcome of Government 2.0, many of them have not adopted any social media platforms yet. Only a minority have fully facilitated social media platforms in their communications with the citizens.
W	Weltanschauung (or worldview)	The empowerment that people in the Middle East area have gained, especially after the Arab Spring events, has also pushed governments in the region to move towards adopting social media to communicate with their citizens.
O	Owner	The government as a whole and the separate government agencies individually.
E	Environmental constraints	There are several constraints that can be considered in this case: the Saudi culture, regulations, speed of growth of technology, Saudi organization's regulations and policy, technological infrastructure and security constraints.

Table2: CATWOE Analysis for Saudi Government 2.0

References

[1] Wearesocial , *Global Digital Statistics 2014: WE ARE SOCIAL'S SNAPSHOT OF KEY DIGITAL INDICATORS*, 2014(online)Available from:http://etonpreneurs.com/uploads/Global%20Social,%20Digital%20&%20Mobile%20Statistics,%20Jan%202014.pdf> [accessed on 20 January2015].

[2] Winterstein, D., Yagan, G., Ataya, L., Trabulsi, P., Al-Assi, R., Shaaban, H., Bukattara, S. *Citizen Engagement and Public Services in the Arab World, The potential of Social Media*, 4th Edition, Mohammeed Bin Rashid School of Government, 2014.

[3] Alshaer, S., Mourtada, R., Salem, F, *the Arab Government Services Outlook 2014. The 2nd Government Summit. Dubai*: The Governance and Innovation Program, Mohammed Bin Rashid School of Government, 2014.

[4] Al-Khalifa, H. S., Al-Razgan, M. S., Al-Rajebah, N. I., & Almasoud, A. M, *Exploring social media usage in Saudi e-government websites*. In Proceedings of the 6th International Conference on Theory and Practice of Electronic Governance (pp. 243-247). ACM, 2012.

[5] Checkland, P., & Poulter, J, Learning for action: a short definitive account of soft systems methodology and its use for practitioner, teachers, and students*, Chichester*: (Vol. 26). Wiley, 2006.

Time Divide, Money Divide and Attitudes towards Globalization, and eGovernment

Helena LINDSKOG[a,1], Alicja SENEJKO[b] and Zbigniew LOS[b]

[a] *Department of Management and Engineering, Linköping University, Sweden*
[b] *Institute of Psychology, Wroclaw University, Poland*

Abstract. . In order to achieve this enhanced eGovernment, it was necessary to provide adequate infrastructure to the citizens. This has been achieved in many countries. The digital divide has become quite narrow. One aspect that has not been covered to the same extent is time and money divide and attitudes towards globalisation. The Department of Management and Engineering, Linkoping University, Sweden, and the Institute of Psychology, Wroclaw University, Poland, have started a joint project in order to fill this gap by investigating different profiles based on two models;
- Socio-economical (time-rich/time-poor and money-rich/money-poor) by Helena Lindskog
- Attitudes towards globalisation by Alicja Senejko et al.

Keywords. Socio-economical profiles, time divide, attitudes towards globalization, electronic government

Extended abstract

eGovernment has become a mainstream feature in many countries and for contacts between government authorities and citizens. Governments use Internet to provide information about public administration, laws, rules, instructions and forms to fill in. Internet is also more and more a vehicle to get citizens' feedback and opinions, as well as in web-dialogues engage and involve them in the democratic processes and political decisions.

In order to achieve this enhanced eGovernment, it was necessary to provide adequate infrastructure to the citizens. This has been achieved in many countries. Today, the digital divide is quite narrow. One aspect that has not been covered to the same extent is time and money divide and attitudes towards globalisation. The Department of Management and Engineering, Linkoping University, Sweden, and the Institute of Psychology, Wroclaw University, Poland, have started a joint project in order to fill this gap by investigating different profiles based on two models:

Attitudes towards globalisation by Alicja Senejko et al

Attitudes towards globalisation are defined as a set of personal biases towards various dimensions of the globalization reality that influence how people function in that reality. The survey included estimated answers to 35 items concerning:

---

[1] helena.lindskog@liu.se.

- **Accepting attitude (**11 items**)** – openness to and exploration of the globalized reality, trusting the changes
- **Critical attitude (**10 items**)** – opposition and lack of consent to globalization-related abuse, condemning most aspects of globalization
- **Fearful attitude (**9 items**)** – sensitivity to globalization-related threats, feelings of uncertainty about the future, which are associated with the lack of trust

Socio-economical (time-rich/time-poor and money-rich/money-poor) developed by Helena Lindskog

The survey included estimated answers to 38 questions concerning:

- **A/ Material resources** (22 questions):
- Type of employment and income
- Additional sources of income (letting of an apartment, scholarship, donations)
- Standard of housing (space, standard and type of residence)

B/ Temporal resources (16 questions):

- Time for everyday responsibilities (professional, school, family, commuting)
- Time for rest and relaxation (sleep, rest, personal hygiene)
- Time for leisure activities (sport, entertainment, hobby, social meetings)

Both the Polish and Swedish sides are currently in parallel collecting information through the same questionnaire in both countries. The collected information will be analysed in each country and then compared. We hope to get answers to questions such as:

- Are there differences between time-rich/money-poor and time-poor/money-rich as well as time-rich/money-rich and time-poor/money-poor towards globalisation and usage of Internet for contacts with companies and government and for private usage?
- Are there differences between the sexes, age and education together with the divisions presented above?
- What is common between the two investigated countries and what differs?

The answers from this study can then be applied for marketing, product development and also for further analysis of the impact of time and money divides and attitudes towards globalisation on eGovernment. In this way, programs can be developed which are inclusive for all citizens independently of their income, disposable time and attitudes towards Internet and globalisation.

From data to data's transparency

Maria Sokhn, Alexandre Cotting, Eliane Maalouf, Arnaud Zufferey

University of applied Sciences of Western Switzerland

Abstract. Decades ago governments, businesses and other actors used to make benefit of generated data in order to manage and monitor their activities. However traditionally data cleaning and analysis have been time consuming and thus used to be kept limited of access. Making the governmental data available and easily accessible and reusable increases the information visibility and decision makers, thus it improves the transparency and efficiency. In this paper we present an ongoing project, which aims to enhance the transparency and efficiency in Swiss public administration, by making use of dashboards communication power.

Keywords. Democracy 2.0, transparency, open data, dashboards

1. Introduction

Since 2009, a revolution over the open data approach has gained importance. This revolution is based on the information and communication technology and has been catalyzed by open government directives that have emerged. Making the governmental data available and easily accessible and reusable will increase the information visibility to the citizens and decision makers and thus improve the transparency and efficiency. Nowadays, Swiss public administrations do not make full use of data power in order to enhance their communication services and their strategy of transparency towards the citizens. This paper presents a data transparency based approach conducted over the Swiss public administrations, which aims to enhance the efficiency and decision making in Swiss public administration, by making use of dashboards.

2. Approach

In the current structure of a Swiss public administration, the different members of the public body at both strategic and operational levels will hardly share the information. This situation lead to: No transparency, Subjective Vision, Decision-making "feeling-driven", and Inefficiency. We propose to make use of dashboards, which will allow different stakeholders to share the same picture of the situation. This sharing will facilitate the dialogue between the different politics party, and it will help the convergence of objectives. In a first phase these dashboards will be available at public administration to target only the efficiency of internal operations. The actors of these public bodies could then decide to include, for example, these dashboards in the reports presented at the primary assemblies. In a third step, data and dashboards could be offered to citizens before meetings so that they can take full knowledge and better prepare their questions. In a later stage, the citizens would have the means to contribute

to the political implementation by taking advantage of the open data in order to propose open innovation services. In our project we work with a set of public administration, which play the role of pilot administrations. We adopted the following methodology: **(1) Analysis of the current situation:** The analysis of public administration needs is defined through iterative workshops. These workshops include different actors and public bodies in order to make a list of potential needs of public bodies. **(2) Data aggregation:** The collection of data necessary to dashboards is primarily with external partners. Then we format the data for integration, processing and sharing. **(3) Data analysis:** A phase of validation and clustering of data is required depending on the chosen dashboards, potentially heterogeneous sources of data and data mining algorithms to be selected. **(4) Data visualization:** Depending on the areas of activity of public organizations and target players, we conduct analysis visualizations of the most appropriate data. **(5) Validation:** A session with specifically chosen public bodies is organized to provide the results for their field and harvest their feedback.

A first study has been conducted over 10 public administrations. It was focused on the energy domain. The analyses were conducted over 10 different communes. A set of different questions was addressed to the public administration. A set of dashboards was designed by combining energetic and financial data revealing hence an overview of the energetic status. In a glance, the dashboards allow to identify the large consumers, the most expensive resources and their evolution over time (Figure below).

These dashboards helped identifying critical points and consequently undertake appropriate decision to investigate in depth the cause of the raised issue. They were a base for discussion to stakeholders and decision makers. These discussions helped enhancing the comprehension of the critical issues. The dashboards have been realized by extracting from the annual account the rows involving water or energy (electricity and heat) using the KNIME software. Then the data are mashed-up in a single view.
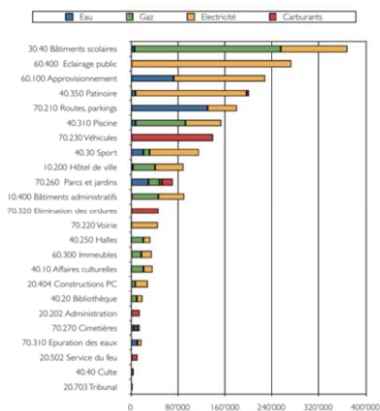
3. Conclusion

These first results pointed out the issues in public administrations in order to provide dashboards. It confirmed that dashboards seem to meet the needs related to visualization and communication; it also revealed a clear lack of operational tools that public administrations can use everyday for monitoring and driving their decisions.

In a Search for Open Budget Effects: The Context of Russia

Radomir Bolgov,[a] Andrei Chugunov,[b] Yuri Misnikov,[b] Vitalina Karachay[b]
[a]St. Petersburg State University, St. Petersburg, Russian Federation
[b]ITMO University, St. Petersburg, Russian Federation

Abstract. The paper seeks to understand how ICT tools are used in Russia to obtain information about government budgets, how to participate online in discussing government spending, and what the effects of open budget initiatives are.

Keywords.eGovernance, eDemocracy, eParticipation, Open Government Data, Open Budgets, Russia, Participatory Budgeting

1. Motivation

We are interested in better understanding democratic effects of eParticipation on decision-making outside western liberal democracies. More specifically, our research interests are focused on the use of open government data for participatory budgeting realized online.

2. Approach

eParticipation is context-dependent [1]. Conceptually, openness, participation, decision-making are interconnected. Fiscal transparency and democracyare already closely intertwined [1].However, it is not always clear how to interlink these seemingly disparate processes into one whole, and encourage different actors work in concert? Brazil's rich experience indicates that participatory budgeting can lead to stronger government transparency and accountability, more equitable and democratic decisions on public spending, higher levels of public participation, stronger sense of democratic socialization and citizenship [2]. We also know that participatory budgeting implies public dialogue and deliberation to decide on government spending priorities. Yet it is not sufficient to make public sector information open and accessible. It is paramount that authorities deploy dedicated information and communication technology (ICT) instruments touse the availability of open data to engage citizens in public debate and increase the effectiveness of managing public resources, including budgets, and eventually raise the legitimacy of taken decisions [3].

3. Current model

As elsewhere, eGovernance landscape in Russia includes two main parts: eGovernment and eDemocracy (see Figure 1 below). Services, Openness and eParticipation dimensions constitute the foundation of eGovernance map. Each dimension includes additional branches in the form of more specialized web platforms and resources. Viewed from that

perspective, eDemocracy looks quite comprehensive, especially the Openness dimension represented by several resources, including those related to Open Government Data and Open Budgets (yellow marked).

Figure 1: Open Government Data and Open Budgets on eGovernance map in Russia

Despite scepticism regarding the effectiveness of these platforms in Russia [4, 5, 6], Open Budgets are part of the overall eParticipation concept. However, in practice there is a clear disconnect between eGovernment and Open Government developments. The existing portals are not interlinked. The Open Budget web page does not contain or refer to eConsultation tools for participatory budgeting tools. Moreover, the budgetary information is not aggregated enough reflect upon policy priorities which is essential for citizens' comprehension of financial information. There have been a lot of participatory budgeting experimentations during 2013-2014 but mostly without a using ICT tools.

4. Conclusion

There is no consolidated eParticipation platform in Russia which would interlink different resources into one whole. The current model is incomplete. It lacks distinctive eConsultation tools that could be used for public deliberation across the government in general and to discuss government spending online in particular. The future research will investigate the use of ICTs for participatory budgeting at regional and municipal levels.

References

[1] A. Macintosh, S. Coleman, A. Schneeberger, eParticipation: The Research Gaps, *Proceedings of the 1st International Conference on Electronic Participation*, ePart'09 (2009), Springer-Verlag Berlin, Heidelberg, 1-11.

[2] S. Khagram, A. Fung, P. De Renzio, P. (Eds.), *Open Budgets: The Political Economy of Transparency, Participation, and Accountability*, Washington, D.C, Brookings Institution Press, 2013.

[3] B. Wampler, *Participatory Budgeting in Brazil: Contestation, Cooperation, and Accountability*, The Pennsylvania State University Press, 2009.

[4] Guidelines on Open Government Data for Citizen Engagement, New York, USA, UNDESA, 2013.

[5] L. Bershadskaya, *A. Chugunov, D. Trutnev, E-participation development: A comparative study of t*he Russian, USA and UK E-petition initiatives, *Proceedings of the 7th International Conference on Theory and Practice of Electronic Governance, ICEGOV* (2013), 73-76.

[6] E. Styrin, N. Dmitrieva, Models for Assessing Openness and Transparency of Executive Authorities, *Proceedings of the 2014 Conference on Electronic Governance and Open Society: Challenges in Eurasia Pages, EGOSE'14* (2014), ACM, New York, USA, 84-87.

[7] L. Smorgunov, Institutional Learning for Government-Society Relations: Russian Case of 'Open Government', *Proceedings of the 2014 Conference on Electronic Governance and Open Society: Challenges in Eurasia Pages, EGOSE'14* (2014), ACM, New York, USA, 101-108.

Evaluation of Open Data maintenance models with Public Administrations and SMEs

Quynh Nhu NGUYEN[a,1], Antonella FRISIELLO[a], Antonio LOTITO[a], Federico MORANDO[b]

[a] *Mobile Solutions, ISMB, Torino, Italy (www.ismb.it/en)*
[b] *NEXA Center for Internet and Society, Torino, Italy (http://nexa.polito.it/)*

Abstract • The Open Data exploitation is strictly related with the quality of data and by consequence with the maintenance of dataset. The paper presents a qualitative study aimed at investigating open data maintenance process (including sharing and re-use issues). Participatory activities and visual tools, involving public, private and societal stakeholders have been applied. The outcome consists in the design of an Open Data Maintenance Life Cycle and operational proposals raised by groups. The study highlights that innovation by design framework is an effective way to manage complexity and plurality of actors, adoptable by any organizations to improve their innovation capability and gain strategic perspectives.

Keywords. Open Data; Innovation by Design; Open Data maintenance, Open Data quality, Open Data validation.

1. Context and aim of the study

How to design an effective and sustainable open-data maintenance process in favor both of public administrations (PA) and novel actors? This question has driven a qualitative study, part of a wider project aimed at studying Open Data in mobility. According to the national observatory (www.dati.gov.it), the Italian PAs have so far made available online over 3000 dataset, 70% of which classified by three out of five stars on the Tim Berners Lee scale [1]. At present the data are published online and in open formats, but there are no structured links between them nor a working flow between data-holders and re-users. Aim of the study is to raise and to assess, through a participatory process involving data holders (PA) and re-users (SMEs), innovative solutions to optimize the process of collection, management, maintenance and provision of Open Data.

2. Methodology: innovation by design

The European Commission has recognized service design as a key driver of service, social and user-centred innovation. In a growing number of countries, governments adopt techniques of design innovation to policy-making and design of services and processes that affect the public [2], [3]. A qualitative study based on this approach was performed in January 2015 through three focus groups organized according to pre-selected application scenarios: Urban Mobility, Culture and Tourism, Hiking and Mountain Tourism. It involved 25 participants from: 11 private enterprises, 8 public institutions, 4 non-profit organizations, 3 Digital Champions [4]. The working path consisted of two main phases: first, the collection and sharing of current practices,

---

[1] Corresponding Author: Quynh Nhu NGUYEN, ISMB, nguyen@ismb.it.

needs and demands, motivations and strengths related with the process of collection and use of open data as is; then, the translation of requirements and inputs in a re-design process of the open-data management cycle.

3. Results of the Study: the Maintenance Life Cycle and Actions, Results, Benefits

As a result of the participatory activities, the focus groups allowed to reconstruct a shared hypothesis of an Open Data Maintenance Life Cycle (Fig. 1), in five phases. New roles and functions have been envisaged by the stakeholders (PA, SMEs, citizens), putting more attention on the open dataset re-use and feedback on them. It came out that the maintenance activities pertains every actor at every stage (orange blocks of the Fig.1), and therefore it should be a distributed task to trigger a virtuous cycle. The heterogeneity of the groups (PA and SMEs) allowed observing and recording the negotiation of needs and the path bringing them to the definition of shared proposals, presented as operational suggestions regarding:
• Actions enabling a sustainable open data maintenance cycle;
• Specifications concerning the quality criteria to reach the 5 stars data;
• Benefits accruing to the involved actors, both public and private.
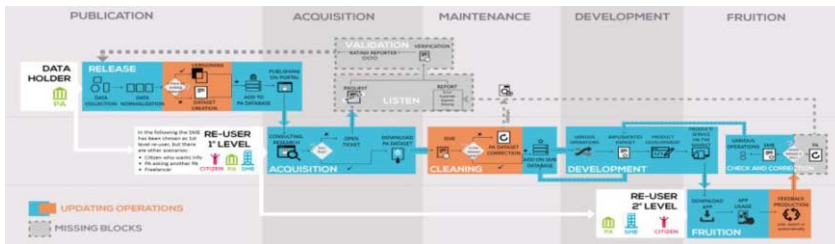
Figure 1. Maintenance Open Data Life Cycle

4. Benefits experienced from the methodology

The application of participatory design methods and hands-on tools has allowed to tackle complex challenges and to gain strategic vision: the involvement of different actors of the process allowed to assume different points of view and trigger new insight about the key-question. Furthermore, visual tools (such as the Platform Model Canvas [5] Issues board, Journey Map[6]) allowed to map the flow of information, touchpoints and actors involved, providing a quick overview of tangible and intangible aspects. On the base of underlying patterns (e.g. good practices, recurring problems and needs) several proposals have been outlined and composed into an 'ideal open data management process'. The study highlighted that new ways of working and dealing with plurality of stakeholders bring transparency and visibility and enable participants' engagement and shared awareness, moving from an abstract plan to the local and pragmatic problem setting and addressing solutions aimed at simplifying complex processes.

References
[1] http://www.w3.org/DesignIssues/LinkedData.html
[2] C. Bason, *Design for Policy*, Gower Publishing, England, 2014
[3] SEE Platform, Policy Booklet 7, *An Overview of Service Design for the Private and Public Sectors*, 2013
[4] http://ec.europa.eu/digital-agenda/en/digital-champions
[5] A. Osterwalder & Y. Pigneur, *Business model generation*, John Wiley & Sons, New Jersey, 2010
[6] Stickdorn, Schneider, *This is Service Design Thinking*, John Wiley & Sons, New Jersey, 2011, p.158

This page intentionally left blank

Workshops

This page intentionally left blank

Enabling Effective Policy Making - Coupling the Power of the Data with the Wisdom of the Crowd

Sotirios KOUSSOURIS [a,1], Panagiotis KOKKINAKOS [a], Konstantinos TSERPES [b], Antonis RAMFOS [c], Vasilis KOULOLIAS [d], Efthimios MPOTHOS [e] Gregoris MENTZAS [e], Yuri GLICKMANN [f] and Peter PHAROW [g]

[a] *Decision Support Systems Laboratory, National Technical University of Athens*
[b] *Knowledge, Media and Distributed Systems Lab, National Technical University of Athens*
[c] *INTRASOFT International*
[d] *Government To You*
[e] *Information Management Unit, National Technical University of Athens,*
[f] *Fraunhofer-Institut for Open Communication Systems FOKUS*
[g] *Fraunhofer-Institut for Digital Media Technologie IDMT*

Keywords. Policy Making, Global Systems, Policy Modelling, System Science

1. Workshop's Objectives

It can be taken for granted that present socio-economic systems are more complex than ever before, as the involved stakeholders have dissimilar (and sometimes contradicting) agendas and, most important of all, might even understand the success (or failure) of specific initiatives in a different way. Thus, the concept of including all interested parties in a structured and active dialogue is constantly gaining momentum and can heavily aid the policy making process. In addition, the wealth of available data of the Web2.0 era can also constitute a very effective weapon towards coping with the aforementioned problem; however, the non-unified nature of the data, accompanied by the different approaches of their publishing and exploitation have not yet revealed their full potential.

Along the above lines, having a set of proper instruments able to aggregate and interpret this dispersed information in an aggregated and timely manner, can be of great help in supporting prompt and effective policy decisions, offering institutions and organizations in charge of policy modelling and policy making a considerable advantage in their policy design and governance. Related methods and systems either leverage the social web and mine information from discussions in social networks, blogs and news streams while inferring sentiments and opinions, or directly involve stakeholders through online participatory and crowdsourcing platforms.

2. Workshop Structure

The purpose of this workshop is to present and discuss the up-to-date methodologies and findings of five of the projects constituting members of the Global Systems Science (GSS) cluster and promote active dialogue among the participants. In more detail, findings and breakthroughs of the following projects will be presented:

- CONSENSUS: Identifying publicly acceptable policy implementations. Even though it is unrealistic to expect citizens to pinpoint the policy implementation that they prefer from the set of alternatives, it is still possible to infer such information through an exercise of ranking the importance of policy objectives according to them. Assuming that the mapping between policy options and objective evaluations is a priori known (through models and simulations), this can be achieved either implicitly through appropriate analysis of social media content related to the policy objective in question or explicitly through the direct feedback provided in the frame of a game. The session will focus on the demonstration of the latter, discussing about the impact of such indications to policy making as well as to the education of citizens about policy decisions.

- EU Community. EU Community will present its approach to provide decision makers with better policy options by combining social media interactions, qualified contributors, document curation, visual analysis plus online and offline trust-building tools. The discussion will focus also on the innovations brought by project through the open source platforms developed, alongside with re-usable data that could serve third party developers. Moreover, EU Community will present early results of the three pilots selected

- Policy Compass: Open Data Driven Policy Analysis and Impact Evaluation. The presentation will demonstrate the underlying concepts and the technical infrastructure of a research prototype put in place that enables citizens and public officials to easily create, apply, share, embed, annotate and discuss causal models, charts and graphs of historical data from trusted open data sources. The aim is to develop methods and tools that facilitate more factual, evidence-based, transparent and accountable policy evaluation and analysis.

- SENSE4US: Data Insights for Policy Makers and Citizens. The presentation will provide an overview of the SENS4US project and the components focusing on its added value to users, starting from the point of analysing social media to inform policy making and talking about modelling and simulating public policy problems of our age.

- SYMPHONY: Nowcasting Policy Indices with Information Markets. The availability of appropriate methods and tools to aggregate scattered pieces of information and stakeholders' expectations is becoming increasingly important for policy making officials. This presentation will show how to use 'Information Markets' as a tool to collect, aggregate and analyse stakeholders' expectations on policy indices in order to support public policy design and implementation. We will discuss various types of Information Markets, as well as cases where both human and computational agents participate in such markets.

- GRACeFUL "Global systems Rapid Assessment tools through Constraint FUnctional Languages. GRACeFUL is working towards a universal language to communicate about the dynamics of global systems and the complex decision making necessary to produce good policy advice. Such formal domain-specific language will be employed by Rapid Assessment Tools supporting a visual and participatory way of modelling, which can be used in situations involving the participation of different stakeholders in decision making and policy design. Constraint Technology will support active, goal-directed problem solving.

The Smart Cities and Smart Government Research-Practice (SCSGRP) Consortium

Theresa A. PARDO [a], Luis LUNA-REYES [b] and Hans Jochen SCHOLL [b]

[a] Center for Technology in Government, University at Albany, NY, USA,
tpardo@ctg.albany.edu

[b] University at Albany, NY, USA, lluna-reyes@albany.edu and University of Washington,
The Information School Seattle, WA, USA, jscholl@uw.edu

Abstract. This workshop is a working meeting of the Smart Cities and Smart Government Research-Practice Consortium. The main goal of the workshop is to continue conversations started since 2012 among teams of researchers interested in doing collaborative work by sharing protocols and data with the purpose of advancing knowledge in the area of smart cities and smart government. The workshop will welcome current and future members of the Consortium and build on SCSGRP Consortium meeting held at dg.o 2015.

Keywords Smart City, Smart Government, Smart Initiative, Smart Governance, Research Collaboration

1. Introduction

The Smart Cities and Smart Government Research-Practice Consortium (SCSGRP) is a robust global research community focused on innovations in technology, management and policy that change the fabric of the world's cities. Created in 2012, the Consortium now includes more than 15 teams of researchers from around the world, sharing ideas, new knowledge, and research and practice innovations in the interest of increasing opportunity for all those who live in and work in these cities. The main purposes of the consortium are

- To formally connect those engaged in cutting edge research on smart cities.
- To support the development of a robust, global and well-connected smart cities research community.
- To create a foundation for the development of multi-institution, multi-national research teams focused on the use of ICTs in cities.

2. Potential Discussion Points

The purpose of the workshop will be to advance the vision of the consortium, exploring frameworks of reference and potential funding opportunities and collaborations among members build on SCSGRP Consortium meeting held at dg.o 2015. Potential topics for discussion include some pending topics and next steps agreed on the ongoing conversations of Consortium members, such as:

- Organize a series of face to face meetings at conferences and through teleconferences.

- Develop strategies and agree on practices that maximize opportunity for data and results sharing among the members.
- Define a research framework to support efforts to conduct comparative studies.

3. Workshop Description

The workshop will be organized as a 4-hour session, including a series of brief presentations and facilitated discussions related to the main points included in the preceding section. The proposed agenda for the workshop will include the following items.

1. Welcome and brief introductions of participants.
2. Short presentation of SCSGRP, and report on progress of main agreements from the dg.o workshop and from the quarterly management conference calls.
3. Discussion on tools and procedures for document, protocols and data sharing.
    i. Dropbox
    ii. GoToMeeting
    iii. Reference management tools (Zotero, Endnote, etc.)
4. Discussion on current research frameworks and projects by members and other interested parties.
5. Outreach plans and Consortium Meeting.

4. About Workshop Facilitators

Theresa A. Pardo is the Director of the Center for Technology in Government, University at Albany, State University of New York (SUNY) and a Research Associate Professor at the Rockefeller College of Public Affairs and Policy and the College of Computing and Information. Dr. Pardo serves as Open NY Policy Adviser to New York State's Governor Andrew Cuomo, a member of the National Advisory Committee for the U.S. Environmental Protection Agency and President of the Digital Government Society.

Luis Felipe Luna-Reyes is an Associate Professor of Informatics at the University at Albany, NY. He holds a Ph.D. in Information Science from the University at Albany, and he is also a member of the Mexican National Research System. His research focuses on electronic government and on modeling collaboration processes in the development of information technologies across functional and organizational boundaries. He is the author or co-author of articles published in Government Information Quarterly, European Journal of Information Systems, Information Polity, Gestión y Política Pública, and System Dynamics Review, among others.

Hans Jochen Scholl is an Associate Professor with tenure in the University of Washington's Information School (Seattle, WA). In his public sector-related research Dr. Scholl's special interests include disaster science, interoperability, smart government, and information artifact evaluation in government. He served as President of the Digital Government Society from 2010 to 2011 and serves as Chair of the IFIP Working Group 8.5 (Information Systems in Public Administration). Dr. Scholl chairs the Electronic Government Track at the Hawaii International Conference on System Sciences (HICSS) and co-organizes the European IFIP EGOV conferences.

Create, Expand and Exploit Linked Open Statistical Data

Efthimios TAMBOURIS [a,b], Evangelos KALAMPOKIS [a,b] and Konstantinos
TARABANIS [a,b]

[a] *Univeristy of Macedonia, Thessaloniki, Greece*
[b] *Centre for Research & Technology – Hellas, Greece*

Abstract. A major part of Open Data concerns statistics that is structured as multi-dimensional data cubes. Linked Data technologies have the potential to realise the vision of combining and performing analytics on top of previously isolated cubes on the Web. Although a few recently launched tools focus on handling linked data cubes, they fall sort to support the whole lifecycle including the integration of multiple cubes and performing analytics on top of integrated and expanded cubes. In this workshop, we will present the OpenCube Toolkit, a set of open source tools that cover the whole linked data cubes lifecycle in an integrated manner. The participants will be guided to a hands-on tour on using and evaluating some of the tools with real datasets e.g. from Eurostat or Digital Agenda.

Keywords. Open data, linked data, statistics, cube, data analytics.

1. Introduction

During the last years, an increasing number of public agencies and organisations have started to open up their data following the so-called Open Government Data movement [1]. A major part of this data concerns statistics such as financial and social indicators. Statistical data is often structured as data cubes where a measured fact is described based on a number of dimensions.

Linked Data has been introduced as a promising paradigm for opening up data as it facilitates the integration of datasets across the Web. In the case of cubes, Linked Data has the potential to realise the vision of performing data analytics on top of integrated but previously isolated cubes across the Web [2]. A fundamental step towards this vision is the RDF data cube (QB) vocabulary, which enables modelling cubes as RDF graphs [3].

2. Dealing with Linked Open Statistical Data

In the case of linked data cubes we consider that relevant raw data go through a process that creates, expands, and exploits linked data cubes. Stakeholders interested in this process include: (a) data publishers (e.g. public agencies, international organisations, research institutes, etc.) that create linked data from raw statistical datasets in an easy way, and (b) data consumers (e.g. policy makers, researchers, citizens, businesses, etc.) that combine and exploit statistics e.g. in decision-making.

The *create* phase deals with transforming raw data into linked data cubes and stakeholders are able to:

- Create RDF data adhering to the Data Cube vocabulary
- Manage and re-use controlled vocabularies (concept schemes, code lists etc.)
- Publish cubes through different interfaces i.e. Linked Data, SPARQL endpoint etc.

The *expand* phase deals with expanding linked data cubes by joining them with other cubes on the Web and stakeholders are able to:

- Discover compatible to join linked data cubes from (a) existing collections of cubes, and (b) an initial cube by computing aggregations across dimensions.
- Establish typed links between compatible to join cubes.
- Create expanded cubes by increasing the size of one of the sets that define a cube i.e. measures, objects of a dimension's level, levels of a dimension, or dimensions.

The *exploit* phase deals with exploiting linked data cubes in data analytics and visualisations and stakeholders are able to:

- Perform OLAP operations on linked data cubes.
- Perform statistical analyses on linked data cubes e.g. compute descriptive statistics, calculate statistics such as correlation coefficient, and create learning models.
- Communicate results through visualisations.

3. Workshop Outline

The aim of the workshop is to present the OpenCube Toolkit [4] comprising software tools that support the three phases of the linked data cube lifecycle. After that, the participants will be guided to a hands-on tour on using and evaluating some of the OpenCube tools with real datasets e.g. from Eurostat or Digital Agenda. These include tools for visualising statistical data and for identifying and merging data residing in different Web locations.

Acknowledgements

References

[1] Kalampokis, E., Tambouris, E., Tarabanis, K.: A Classification Scheme for Open Government Data: Towards Linking Decentralized Data. International Journal of Web Engineering and Technology, 6(3), 266-285 (2011)

[2] Kalampokis, E., Tambouris, E., Tarabanis, K. (2013): Linked Open Government Data Analytics. M.A. Wimmer, M. Janssen, and H.J. Scholl (eds.): EGOV2013, LNCS 8074, pp.99-110. IFIP International Federation for Information Processing

[3] Cyganiak, R., Reynolds, D. (2014): The RDF Data Cube Vocabulary: W3C Recommendation. W3C, Tech. Rep., January 2014.

[4] Kalampokis, E., Nikolov, A., Haase, P., Cyganiak, R., Stasiewicz, A., Karamanou, A., Zotou, M., Zeginis, D., Tambouris, E., Tarabanis, K. Exploiting Linked Data Cubes with OpenCube Toolkit, Proc. of the ISWC 2014 Posters and Demos Track a track within 13th International Semantic Web Conference (ISWC2014), 19-23 October 2014, Riva del Garda, Italy, CEUR-WS Vol.1272 (2014).

Subject Index

Author Index

This page intentionally left blank

This page intentionally left blank