Human-Centered Text Mining: A New Software System

Jonas Poelmans[1,5], Paul Elzinga[3], Alexei A. Neznanov[5], Guido Dedene[1,4], Stijn Viaene[1,2], and Sergei O. Kuznetsov[5]

[1] KU Leuven, Faculty of Business and Economics, Naamsestraat 69, 3000 Leuven, Belgium
[2] Vlerick Leuven Gent Management School, Vlamingenstraat 83, 3000 Leuven, Belgium
[3] Amsterdam-Amstelland Police, James Wattstraat 84, 1000 CG Amsterdam, The Netherlands
[4] Universiteit van Amsterdam Business School, Roetersstraat 11 1018 WB Amsterdam, The Netherlands
[5] National Research University Higher School of Economics (HSE), Moscow, Russia
{Jonas.Poelmans,Stijn.Viaene,Guido.Dedene}@econ.kuleuven.be,
{aneznanov,skuznetsov}@hse.ru,
Paul.Elzinga@amsterdam.politie.nl

Abstract. In this paper we introduce a novel human-centered data mining software system which was designed to gain intelligence from unstructured textual data. The architecture takes its roots in several case studies which were a collaboration between the Amsterdam-Amstelland Police, GasthuisZusters Antwerpen (GZA) hospitals and KU Leuven. It is currently being implemented by bachelor and master students of Moscow Higher School of Economics. At the core of the system are concept lattices which can be used to interactively explore the data. They are combined with several other complementary statistical data analysis techniques such as Emergent Self Organizing Maps and Hidden Markov Models.

Keywords: Formal Concept Analysis, Text Mining, Software System, Applications, Concept lattices.

1 Introduction

A crucial enabler for innovation in 21st century data-intensive organizations is being able to deal with massive amounts of textual information. Amongst others in the crime data mining framework of Chen et al. (2004) text mining was pointed out as a promising research field. Unfortunately till date only few successful applications have been reported on in the literature. Over the past years several papers have been published on applying Natural Language Processing (NLP) techniques and extracting key words from texts often with the aim of building an ontology and classifying some documents (e.g. Cimiano et al. 2005, Maio et al. 2012). In this paper we try to go further and focus on semi-automatically exploring textual data using visual models. More in particular we showcase the "COncept Relation Discovery and Innovation

P. Perner (Ed.): ICDM 2012, LNAI 7377, pp. 258–272, 2012.

Enabling Technology" (CORDIET) software system which takes its roots in several real-life case studies with the Amsterdam-Amstelland Police and GZA hospitals.

Before our research, no automated analyses were performed on the observational reports written by officers and filed in the Amsterdam-Amstelland police region. The reason was an absence of good instruments to detect the observations containing interesting information and to analyze the texts they contain. Only on the structured information stored in police databases, analyses were performed. These include the creation of management summaries using Cognos information cubes, geographical analysis of incidents with Polstat and data mining with Datadetective. A few projects were devoted to automatically identifying domestic violence in statements made by victims however the results were not convincing enough to make it into operational policing practice.

At the core of our system are concept lattices (Wille 1982, Ganter et al. 1999) which can be used to visualize and interactively gain insight in the underlying concepts of the data. The lattice-based analysis can be combined with Hidden Markov Models (HMM) (Rabiner 1989) and Emergent Self Organising Maps (ESOM) (Ultsch 2003, Ultsch 2005). We chose for a human-centered (Fayyad 2002) setup of the system. A domain expert, who is not a trained computer scientist or statistician, can apply a powerful arsenal of analysis methods to his particular problem and adapt them to his needs without having to deal constantly with technical details.

The remainder of this paper is composed as follows. Section 2 describes the setup of the project and the software system. In section 3, an overview of the datasets used during the research is given. In section 4 we describe the functionality of the CORDIET toolset. In section 5 we discuss the case studies and showcase the potential of our approach. In section 6 we elaborate on why a human-centered Knowledge Discovery in Databases (KDD) approach may be better suited for text mining than standard fully automated machine learning techniques.

2 CORDIET Project Setup

2.1 Student Groups

After several presentations of the text mining research we are doing and the software system we want to develop, over 20 bachelor and master students of Moscow Higher School of Economics (HSE) showed their interest to actively participate. After several plenary meetings where the overall specification of the system was discussed, each student chose a component he or she wanted to develop. The goal was not only to develop a working system but also to help students gain experience which is valuable for their future career. Each student was given the task to first collect relevant literature, existing open source implementations, etc. Then we discussed in detail specifications for the component, interoperability requirements with other student components, programming language and useful APIs, etc. Although for several components of our system open source implementations exist, we chose to let students re-implement them so they could fully master technical programming skills and the data analyses techniques they are working with.

2.2 Software Architecture

We chose for a three-layered client-server architecture. The majority of the computationally intensive tasks is performed by the server components which are made available through web services.

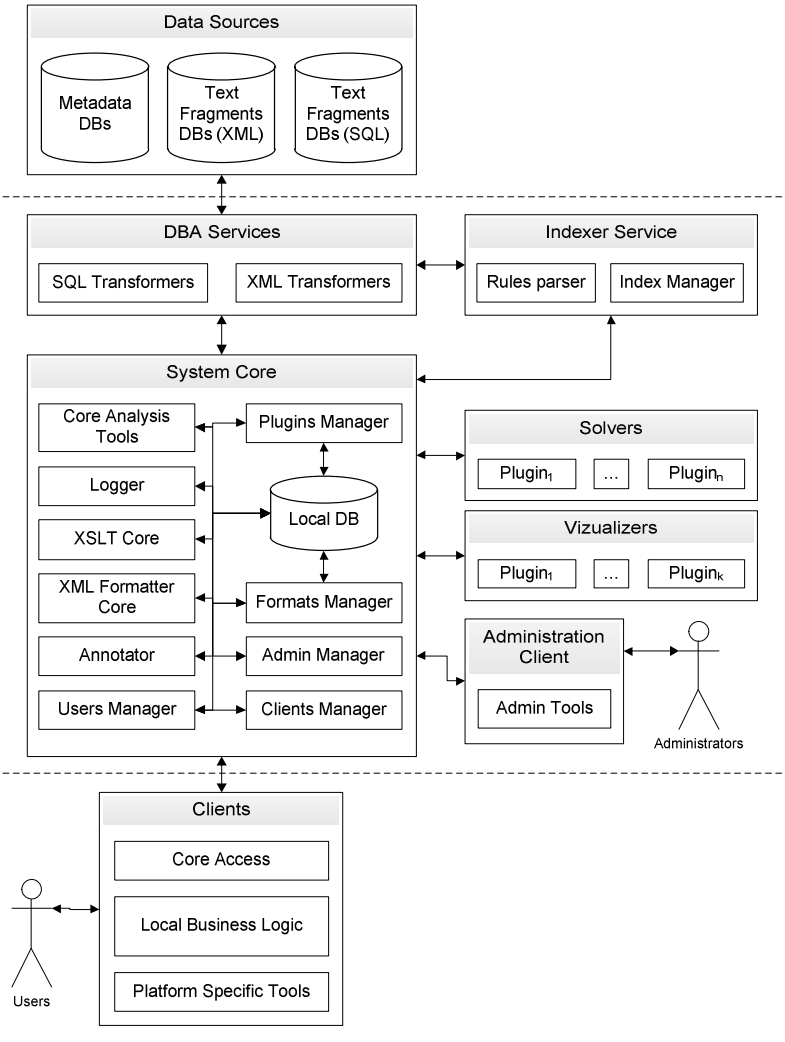

Fig. 1. A representation of the CORDIET architecture

The major components of the data layer are a relational database (at the moment we use PostgreSQL), XML input data files and the Lucene indexer. The data indexing component reads XML files from a selected dataset, parses these files into the SQL database and generates the Lucene index. The database content and Lucene index can be accessed and used by the business layer components through the data access layer. The

business layer will offer functionality to create business objects such as Hidden Markov Models, Emergent Self Organising Maps, concept lattices etc. (at the moment only the concept lattice module is fully implemented and working). The user can install the client module and remotely connect to the services which make functionality of the middle layer available. Figure 1 contains a high level overview of the software system.

3 Data Sources

In this research four main data sources have been used for empirical validation. The first data source was the police database "Basis Voorziening Handhaving" (BVH) of the Amsterdam-Amstelland police. Multiple datasets were extracted from this data source, including the domestic violence, human trafficking and terrorism dataset. The second data source was the World Wide Web, from which we collected 1072 scientific articles on Formal Concept Analysis (FCA). The third dataset consists of 148 breast cancer patients that were hospitalized in GZA hospital campus Sint-Augustinus during the period January 2008 till June 2008. The fourth dataset contains 533 chat conversations of pedophiles from the website www.perverted-justice.com.

Table 1. Overview of data sources used for empirically validating the CORDIET toolset

	textual	unstructured	data type	#data items	year
Domestic violence	X	X	incident reports	4814	2007
Human trafficking	X	X	observational reports	266157	2005-2010
Terrorism	X	X	observational reports	166577	2005-2009
Pedophiles	X	X	chat conversations	533	2004-2011
Clinical pathways	X		patient data	148	2008
Scientific articles	X	X	papers	1072	2003-2011

3.1 Data Source BVH

The database system BVH is used by all police forces of the Netherlands, the military police and the Royal Marechaussee. This database system contains both structured and unstructured textual information. The contents of the database are subdivided in two categories: incidents and activities. Incident reports describe events that took place which are in violation with the law. These include violence, environmental and financial crimes. During our first case study we analyzed 4814 incident reports describing violent incidents, filed in 2007, and we aimed at automatically recognizing the domestic violence cases.

Activities are often performed after certain incidents occurred and include interrogations, arrestments, etc., but activities can also be performed independent of any incident, such as motor vehicle inspections, an observation made by a police officer of a suspicious situation, etc. Each of these activities performed are described in a textual report by the responsible officer. In the year 2005, Intelligence Led Policing (Collier 2006) was introduced at the police of Amsterdam, resulting in a

sharp increase in the number of filed activity reports describing observations made by police officers, i.e. from 34817 in 2005 to 67584 in 2009. These observational reports contain a short textual description of what has been observed and may be of great importance for finding new criminals. In the second and third case study, we used the observations made by police officers to find indications for human trafficking in 266157 reports and indications for radicalizing behavior in 166577 reports. The involved persons and vehicles are stored in structured data fields in a separate database table and are linked to the unstructured report using relational tables. Therefore, we wrote an export program that automatically composes documents based on the most recently available information in the databases. These documents are stored in XML format and can be read by the CORDIET toolset.

3.2 Data Source Scientific Articles

Over 1000 pdf files containing articles about FCA research were downloaded from the WWW and analyzed with the CORDIET system. During the analysis, these pdf-files were converted to ordinary text and the abstract, title and keywords were extracted. Lucene was used to index the extracted parts of the papers using our thesaurus containing terms referring to interesting research topics. The result was a cross table describing the relationships between the papers and the research topics from the thesaurus. This cross table was used as a basis to generate the lattices.

We only used abstract, title and keywords because the full text of the paper may mention a number of concepts that are irrelevant to the paper. For example, if the author wrote an article on information retrieval but also gives an overview of related work mentioning papers on fuzzy FCA, rough FCA, etc., these concepts may be irrelevant however they are detected in the paper. If they are relevant to the entire paper we found they were typically also mentioned in title, abstract or keywords.

3.3 Data Source Clinical Pathways

The third dataset consists of 148 breast cancer patients that were hospitalized, in GZA hospital campus Sint-Augustinus, during the period from January 2008 till June 2008. They all followed the care trajectory determined by the clinical pathway Primary Operable Breast Cancer (POBC), which structures one of the most complex care processes in the hospital. Every activity or treatment step performed to a patient is logged in a database and in the dataset we included all the activities performed during the hospitalization of these patients. Each activity has a unique identifier and we have 469 identifiers in total for the clinical path POBC. We clustered activities with a similar semantic meaning to reduce the complexity of the lattices. The resulting dataset is a collection of XML files where each XML contains all activities performed to one patient.

3.4 Data Source Pedophiles

Because original chat data collected by the Dutch police force organizations is restricted by law, results may not be made public. To demonstrate our FCA based method we use

the chat data collected by a public American organization, Perverted Justice, which actively searches for pedophiles on the internet. We downloaded 533 chat files, i.e. one for each of the 533 different suspects. The victims in all chat files are adults playing the role of a young girl or boy in the age from 12 to 14. All these adults are members of the Perverted Justice organization and are trained to act as a youngster. The adults playing the victim try to lure the suspect by playing his or her role as good as possible. The behavior of the victims cannot be representative for young girls or boys, but the behavior of the suspects is realistic since they really believe to have contact with a young girl or boy and act in that way.

4 Functionality of the CORDIET Software

Figure 2 displays a screenshot of the CORDIET software which the user will see when he starts the system. First, the user can load a set of textual XML files (e.g. police reports) in the database (1). The structured part of these reports will be displayed together with their textual content on the right of the screen (2). One of the central components of our text analysis environment is the semantic network containing the collection of attributes used to index the data files. The initial semantic network is typically constructed based on expert prior knowledge and incrementally improved by analyzing the concept gaps and anomalies in the resulting lattices. The user can then edit the semantic network containing terms (e.g. my father, my mother, my sister), clusters of terms (e.g. family members), temporal (e.g. January till June 2009) and more complex compound attributes (e.g. “family members” and not “January till June 2009”) which will be used to analyze the texts with CORDIET. The semantic network contains multiple abstraction levels. The first level of granularity

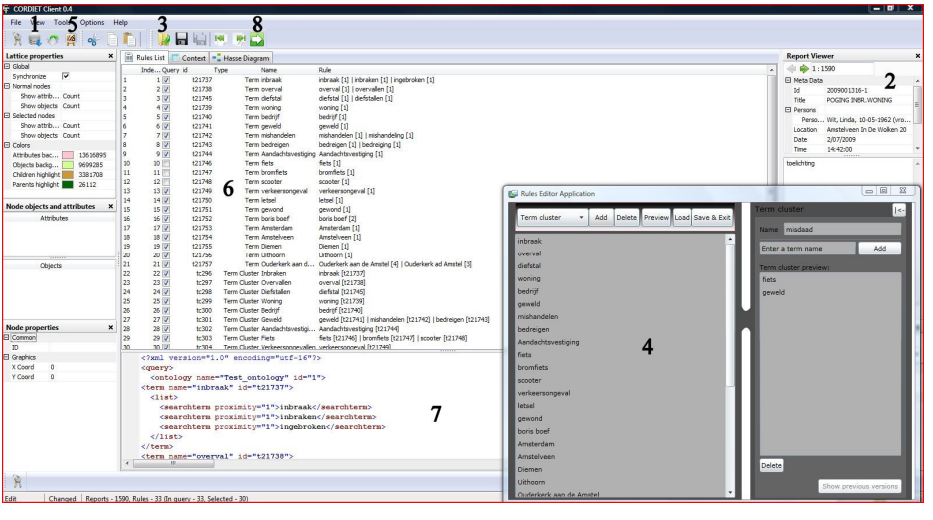

Fig. 2. CORDIET data preparation stage

contains the search terms of which most are grouped together based on their semantic meaning to form the term clusters at the second level of granularity. The compound attributes can be composed of simple attributes using first order logic operators. The user can open (3) a textual editor (4) or open (5) a separate graph based editor to work with this semantic network. Using the selection pane (6) he can select the attributes which will be used for analysis. The corresponding XML code used to store the ontology is shown below (7). By clicking on the green button (8) a formal context will be created from which other visual artifacts can be derived. In the current prototype only the concept lattice based algorithms have been implemented.

The user has several options to interact with the lattice (9) visualization in figure 3. He can choose to display the contents of objects and attributes but also to see their names in the lattice and use a condensed representation as in the lattice in Figure 3 (10). After clicking on a concept, the names of the texts (in this example police reports) in the extent of the concept are shown (12) together with the attribute names in the intent (11) on the left. After clicking on the name of a report, the contents will be visualized on the right (13). After clicking on the name of an attribute its components will also be visualize on the right (14). Clicking on a node will highlight all concepts on paths to the infimum and supremum of the lattice (see figure 5).

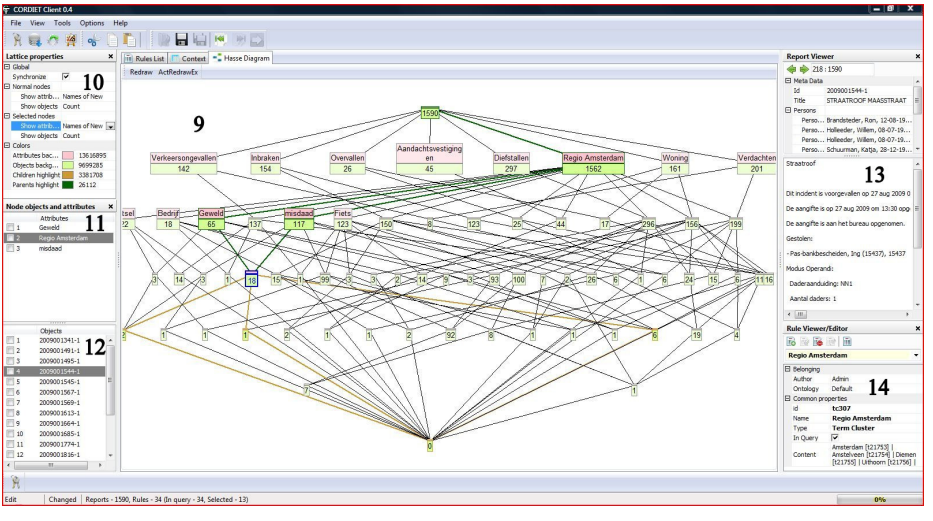

Fig. 3. CORDIET lattice interaction stage

5 Case Studies

In the healthcare case study (see section 5.1) in collaboration with GZA hospitals we worked on structured textual data, namely database logs of activities performed to patients to identify quality issues in care processes (Poelmans et al. 2010c). The other four case studies described in this section were a collaboration between KU Leuven and the Amsterdam-Amstelland police. In the domestic violence research (see section

5.2) we aimed at exploring and refining the concept and definition of domestic violence by analyzing statements made by victims to the police. An important spin-off of this exercise was the development of a highly accurate automated case labeling system (Poelmans et al. 2011a, Poelmans et al. 2010a). When we started analyzing observational reports the goal was to extract unknown suspects potentially involved in human trafficking, forced prostitution (see section 5.3 and 5.4) and terrorist activities (Poelmans et al. 2011b, Elzinga et al. 2010). We also investigated how we could offer police investigators quick but thorough insight in long chat conversations (see section 5.5) of potential pedophiles with children (Elzinga et al. 2012). Finally, we show how we used FCA as a meta-technique to analyze the literature on FCA (see section 5.6). In the remainder of this section we showcase briefly the potential of our software system for each of these cases. For a thorough description of these cases the reader is kindly referred to the papers mentioned.

5.1 Care Process Analysis

Care pathways are a methodology to structure multidisciplinary care processes of patients with a specific clinical problem. During auditing of the breast cancer care process in the GZA hospital group we obtained the concept lattice in Figure 4. The diagram shows that several mandatory key interventions were not always performed to hospitalized patients who underwent breast conserving surgery. For example, "physiotherapy", "emotional support" and "counseling by social service" were not performed to 15, 2 and 3 of the 60 cancer patients respectively. After presenting these results to the care process managers, we jointly looked for the root causes of this

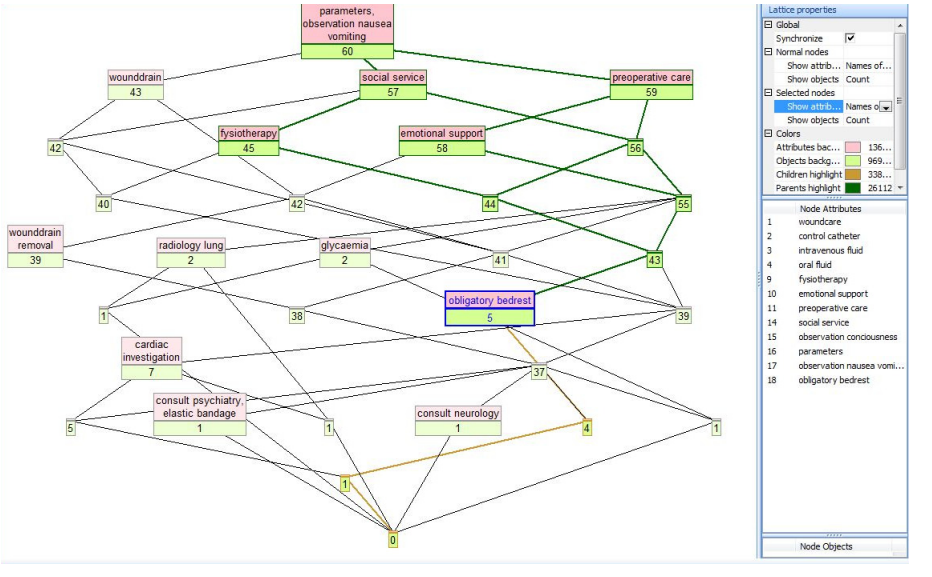

Fig. 4. Activities performed to breast cancer patients during hospitalization

problem. One of these causes turned out to be that over the past years the length of stay of the patients in the hospital was dramatically lowered without modifying the original care process model. After these findings the prescribed process model was rewritten to take into account this shorter length of stay.

5.2 Domestic Violence under Scrutiny

A definition of a problem is often inaccurate and incomplete due to the complex nature of the reality it was designed to deal with.

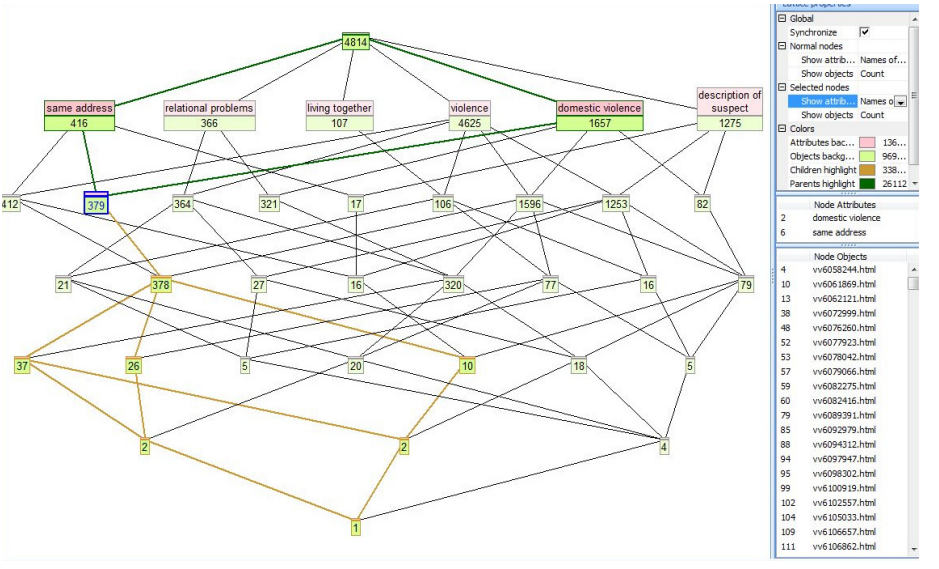

Fig. 5. Analyzing statements made by victims of a violent incident

An example is the domestic violence definition which was employed by the Amsterdam-Amstelland police (Keus et al. 2000): "*Domestic violence can be characterized as serious acts of violence committed by someone in the domestic sphere of the victim. Violence includes all forms of physical assault. The domestic sphere includes all partners, ex-partners, family members, relatives and family friends of the victim. The notion of family friend includes persons that have a friendly relationship with the victim and (regularly) meet with the victim in his/her home.*" The lattice in Figure 5 contains 4814 police reports of which 1657 were labeled as domestic violence by police officers.

With CORDIET, the user can visually represent the underlying concepts in the data, gain insight in the complexity of the domain under investigation and zoom in on interesting concepts. For example we clicked on the node with 379 reports where suspect and victim lived on the same address and labeled as domestic violence by officers. Domain experts assumed that a situation where perpetrator and victim live at the same address is always a case of domestic violence, since these persons are

probably family members, however this turned out not to be true. Analysis of the reports with attribute "same address" and not labeled as domestic violence revealed borderline cases such as violence in prisons, violence between a caretaker and inhabitant of an old folks home, etc. Each of these cases were presented to the steering board of the domestic violence policy. This resulted in an improved definition of domestic violence and an improved handling of domestic violence cases.

5.3 Identifying Human Trafficking Suspects

In the past, relevant suspects sometimes remained undetected in the overload of observational reports. Since human trafficking indications for observed persons are spread over multiple reports which are typically filed by different officers, a visual picture which summarizes these data and makes it accessible for exploration is an important instrument for investigators. With CORDIET, we present this picture to the user in the form of a concept lattice. For example Figure 6 shows potential human trafficking suspects, i.e. men who force girls to work in prostitution, with Eastern European nationality (Poelmans et al. 2011).

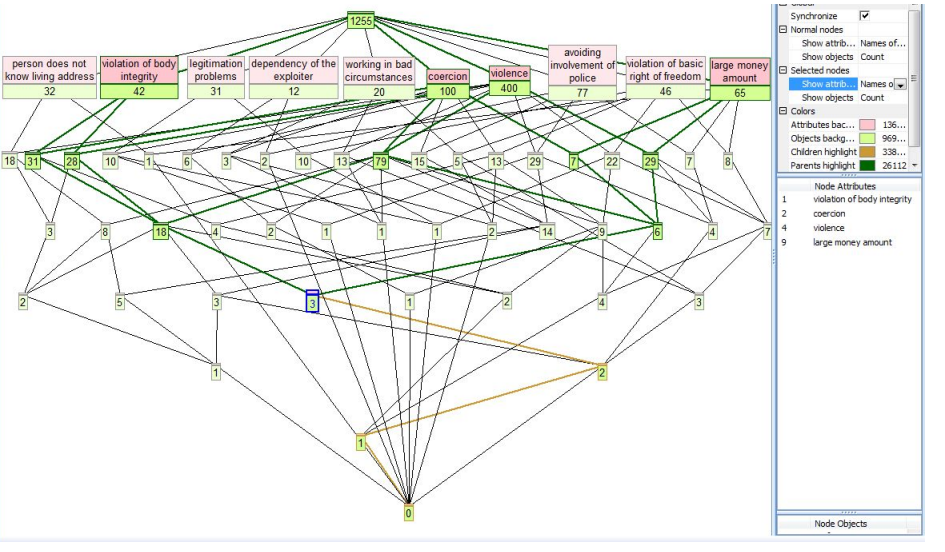

Fig. 6. Lattice of potential suspects and victims of human trafficking

We clicked on the concept with 3 persons in the extent and attributes "violation of body integrity", "coercion", "violence" and "large money amount" in the intent. The police officer can see the names of the persons in the "node objects" pane and double click on a name to create a detailed profile lattice for him or her (see also section 5.4). Please note that we made the node objects pane containing names of potential suspects and victims invisible. Similar results were achieved for terrorism data (Elzinga et al. 2010).

5.4 Profiling Human Trafficking Suspects

A profile of a selected potential suspect can be automatically generated using our system and displays all available information in a lattice together with the temporal evolution of this person (Poelmans et al. 2011b). Figure 7 shows a lattice profile of a loverboy suspect. The objects are names and birth dates of persons found in reports in which our main suspect was mentioned. After analyzing the lattice with domain experts it became clear that he used violence to make the 2 young girls (Sardientje and Hermina) consent to sexual exploitation.

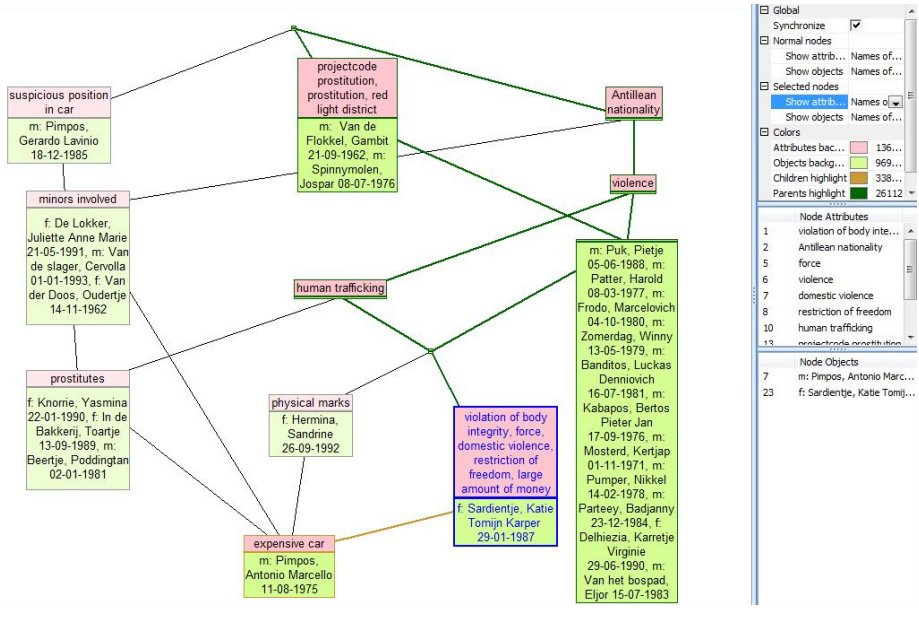

Fig. 7. Analysis of social network of suspect Pimpos Antonio Marcello (fictive name)

5.5 Analyzing Chat Conversations of Pedophiles

Chat conversations can be very long and time-consuming to read. A system which helps officers quickly identify those conversations posing a threat to a child's safety and understand what has been talked about may significantly speed up and improve the efficiency of their work (Elzinga et al. 2012).

The lattice in Figure 8 shows how a set of 533 chat conversations was analyzed with FCA. We defined 7 term clusters containing keywords which were used by pedophiles in their chat conversations. We numbered these 7 attributes according to the severity of the threat to the child's safety. We clicked on a concept with 96 conversations in the extent and attributes "asks", "asks about sex", "describes about sex" and "asks for address". In the "node objects" pane the user can click on the name of a conversation to display its contents. In Elzinga et al. (2012) we describe in detail

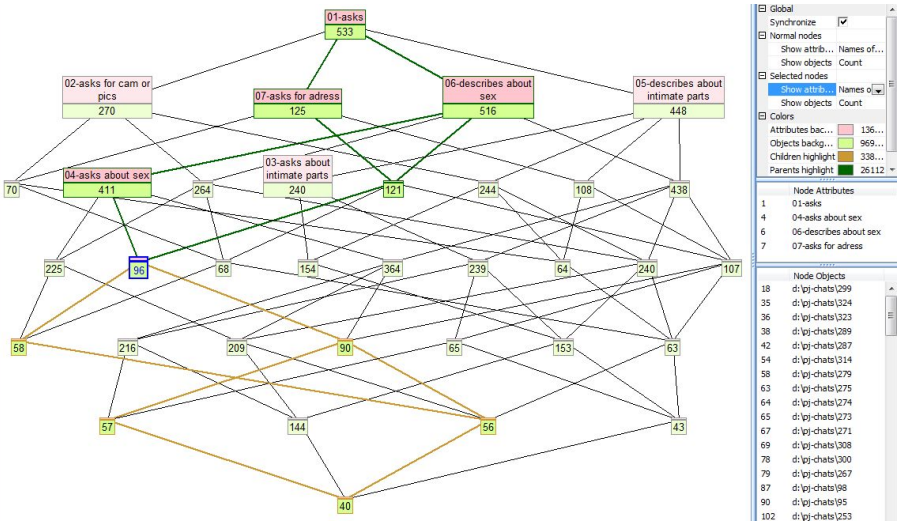

Fig. 8. Analyzing chat conversations of pedophiles with members of the perverted justice organization who pretend to be a young child

how we selected chats from such a concept lattice and analyze them in detail with temporal relational semantic systems.

5.6 FCA Literature Study

CORDIET was also used in an exploratory study for visually representing and exploring scientific papers (Poelmans et al. 2010b, Poelmans et al. 2012). The user can dynamically select and deselect attributes representing research topics and relevant papers will be shown. An author looking for relevant works in his/her filed may benefit from such a system.

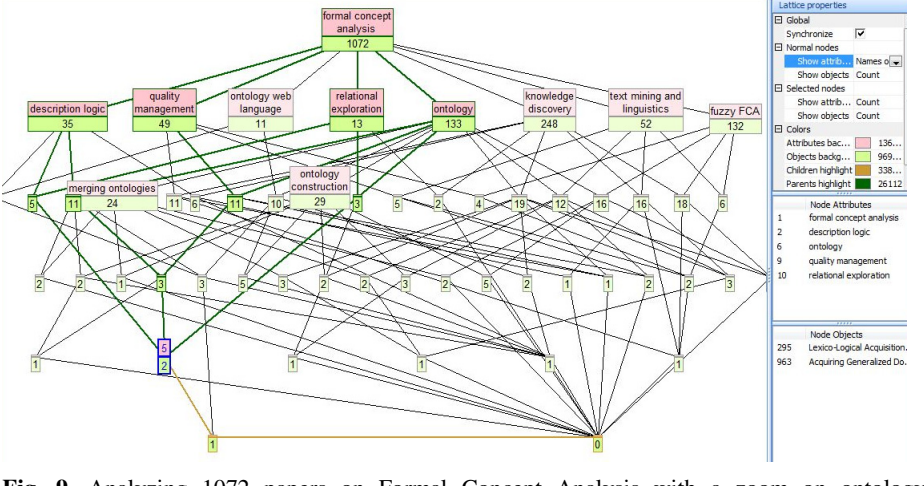

Fig. 9. Analyzing 1072 papers on Formal Concept Analysis with a zoom on ontology engineering papers

The lattice in Figure 9 displays 1072 papers on Formal Concept Analysis. The attributes which we chose to analyze these data are related to ontology engineering. The selected node contains 2 papers on quality management of ontologies using relational exploration, an algorithm which can be applied in the context of Formal Concept Analysis enriched with description logics.

6 Discussion, Conclusions and Future Work

Yearly more than 5000 statements are made by victims of a violent incident to the Amsterdam-Amstelland police and over 60000 observational reports are filed by officers. The need for a text analysis system became apparent since there is no capacity to deal with all this information manually. Existing fully automated text mining methods were found to be no good option. First, automated machine learning techniques such as decision trees, support vector machines, neural networks, automated keyword extraction techniques, etc. assume that the underlying concepts of the domain are clear, can be extracted from the data and used for classification. Unfortunately we found this is often not true, e.g. there was no consensus amongst police officers on whether certain borderline cases should be labeled as domestic violence or non-domestic violence, officers labeled several cases wrongly, prior knowledge was not always useful, etc. Second, in observational police reports, indications against a certain suspect are spread over multiple reports. In the case of human trafficking for example, an additional complicating factor is that only a few thousand reports in a dataset of over 250 000 reports are relevant. Third, often a human expert should stay in the loop because of the particular nature of the domain under investigation. Selecting suspects for in-depth investigation should be done by a human expert who can be held accountable for his decisions.

In each of the described case studies FCA was used to create intuitive and interactive visualizations which can easily be interpreted by domain experts. To cope with scalability issues, attributes can be clustered, segments of the data can be selected, objects can be grouped, etc. in the CORDIET system. In particular the data summarization capabilities of the lattice diagrams were found to be of interest to the users. Texts, their properties, connections with other pieces of text, etc. can be distilled from such diagrams with ease. We believe that the presented CORDIET system may significantly improve the efficiency of working with unstructured textual data. However, we are aware that a lot of work remains to be done. Avenues for future research include:

- Implementation of other (complementary) data visualization techniques such as Emergent SOM and HMMs and integrating these models with the existing concept lattice module and with each other such that data exploration can be done in an efficient yet thorough manner.
- Analyzing the possibilities of NLP techniques for Dutch language to speed up thesaurus building.
- Extending existing functionality to better analyze social structure of criminal communities.

Acknowledgments. Jonas Poelmans is Aspirant of the "Fonds voor Wetenschappelijk Onderzoek – Vlaanderen" (FWO) or Research Foundation – Flanders.

References

[1] Cimiano, P., Hotho, A., Staab, S.: Learning Concept Hierarchies from Text Corpora using Formal Concept Analysis. J. Artif. Intell. Res (JAIR) 24, 305–339 (2005)

[2] Collier, P.M.: Policing and the intelligent application of knowledge. Public Money & Management 26(2), 109–116 (2006)

[3] Elzinga, P., Poelmans, J., Viaene, S., Dedene, G., Morsing, S.: Terrorist threat assessment with Formal Concept Analysis. In: Proc. IEEE International Conference on Intelligence and Security Informatics, Vancouver, Canada, May 23-26, pp. 77–82 (2010)

[4] Elzinga, P., Wolff, K.E., Poelmans, J., Viaene, S., Dedene, G.: Analyzing chat conversations of arrested child abusers with temporal relational semantic systems. In: Contributions to 10th International Conference on Formal Concept Analysis, Leuven, Belgium, May 6-10 (2012)

[5] Keus, R., Kruijff, M.S.: Huiselijk geweld, draaiboek voor de aanpak. Directie Preventie, Jeugd en Sanctiebeleid van de Nederlandse justitie (2000)

[6] Maio, C.D., Fenza, G., Gaeta, M., Loia, V., Orciuoli, F., Senatore, S.: RSS-based e-learning recommendations exploiting fuzzy FCA for Knowledge Modeling. Applied Soft Computing 12(1), 113–124 (2012)

[7] Poelmans, J., Dedene, G., Verheyden, G., Van der Mussele, H., Viaene, S., Peters, E.: Combining Business Process and Data Discovery Techniques for Analyzing and Improving Integrated Care Pathways. In: Perner, P. (ed.) ICDM 2010. LNCS, vol. 6171, pp. 505–517. Springer, Heidelberg (2010c)

[8] Poelmans, J., Elzinga, P., Viaene, S., Dedene, G.: A Case of Using Formal Concept Analysis in Combination with Emergent Self Organizing Maps for Detecting Domestic Violence. In: Perner, P. (ed.) ICDM 2009. LNCS, vol. 5633, pp. 247–260. Springer, Heidelberg (2009)

[9] Poelmans, J., Elzinga, P., Viaene, S., Dedene, G.: Curbing domestic violence: Instantiating C-K theory with Formal Concept Analysis and Emergent Self Organizing Maps. Intelligent Systems in Accounting, Finance and Management 17(3-4), 167–191 (2010a)

[10] Poelmans, J., Elzinga, P., Viaene, S., Dedene, G.: Formal Concept Analysis in Knowledge Discovery: A Survey. In: Croitoru, M., Ferré, S., Lukose, D. (eds.) ICCS 2010. LNCS, vol. 6208, pp. 139–153. Springer, Heidelberg (2010b)

[11] Poelmans, J., Elzinga, P., Viaene, S., Dedene, G.: Formally Analyzing the Concepts of Domestic Violence. Expert Systems with Applications 38(4), 3116–3130 (2011a)

[12] Poelmans, J., Elzinga, P., Dedene, G., Viaene, S., Kuznetsov, S.O.: A Concept Discovery Approach for Fighting Human Trafficking and Forced Prostitution. In: Andrews, S., Polovina, S., Hill, R., Akhgar, B. (eds.) ICCS-ConceptStruct 2011. LNCS, vol. 6828, pp. 201–214. Springer, Heidelberg (2011b)

[13] Poelmans, J., Ignatov, D.I., Viaene, S., Dedene, G., Kuznetsov, S.: Text mining scientific papers: a survey on FCA-based information retrieval research. In: 12th Industrial Conference on Data Mining. LNCS, July 13-20, Berlin, Germany. Springer (2012)

[14] Rabiner, L.R.: A tutorial on Hidden Markov Models and selected applications in speech recognition. Proceedings IEEE 77(2), 257–286 (1989)

[15] Stumme, G., Wille, R., Wille, U.: Conceptual Knowledge Discovery in Databases using Formal Concept Analysis Methods. In: PKDD 1998. LNCS, vol. 1510, pp. 450–458. Springer, Heidelberg (1998)
[16] Ultsch, A.: Maps for visualization of high-dimensional Data Spaces. In: Proc. WSOM 2003, Kyushu, Japan, pp. 225–230 (2003)
[17] Ultsch, A., Hermann, L.: Architecture of emergent self-organizing maps to reduce projection errors. In: Proc. ESANN 2005, pp. 1–6 (2005)
[18] Wille, R.: Restructuring lattice theory: an approach based on hierarchies of concepts. In: Rival, I. (ed.) Ordered Sets, pp. 445–470. Reidel, Dordrecht-Boston (1982)
[19] Wolff, K.E.: States, Transitions, and Life Tracks in Temporal Concept Analysis. In: Ganter, B., Stumme, G., Wille, R. (eds.) Formal Concept Analysis. LNCS (LNAI), vol. 3626, pp. 127–148. Springer, Heidelberg (2005)