Expert Systems with Applications 40 (2013) 6538–6560

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Review

Formal concept analysis in knowledge processing: A survey on applications

Jonas Poelmans [a,c,*], Dmitry I. Ignatov [c], Sergei O. Kuznetsov [c], Guido Dedene [a,b]

[a] KU Leuven, Faculty of Business and Economics, Naamsestraat 69, 3000 Leuven, Belgium
[b] Universiteit van Amsterdam Business School, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands
[c] National Research University Higher School of Economics, Pokrovsky boulevard, 11, 109028 Moscow, Russia

ARTICLE INFO

Keywords:
Formal concept analysis (FCA)
Knowledge discovery in databases
Text mining
Applications
Systematic literature overview

ABSTRACT

This is the second part of a large survey paper in which we analyze recent literature on Formal Concept Analysis (FCA) and some closely related disciplines using FCA. We collected 1072 papers published between 2003 and 2011 mentioning terms related to Formal Concept Analysis in the title, abstract and keywords. We developed a knowledge browsing environment to support our literature analysis process. We use the visualization capabilities of FCA to explore the literature, to discover and conceptually represent the main research topics in the FCA community. In this second part, we zoom in on and give an extensive overview of the papers published between 2003 and 2011 which applied FCA-based methods for knowledge discovery and ontology engineering in various application domains. These domains include software mining, web analytics, medicine, biology and chemistry data.

1. Introduction

Formal Concept Analysis (FCA) was introduced in the early 1980s by Rudolf Wille as a mathematical theory (Wille, 1982) taking its roots in the works of Barbut and Monjardet (1970), Birkhoff (1973) and others for the formalization of concepts and conceptual thinking. FCA has been applied in many disciplines such as software engineering, knowledge discovery and information retrieval during the last 15 years. The reader is kindly referred to part one of this survey (Poelmans, Kuznetsov, Ignatov, & Dedene, 2013) for an overview of the FCA-based methods for data analysis. In this paper, we zoom in on and give an extensive overview of the papers published between 2003 and 2011 on using FCA for knowledge discovery and ontology engineering in various application domains.

The applications of FCA are very diverse and the same underlying models were regularly used in different application areas. To structure the FCA research domain from an application point of view we tried to devide the large amount of papers into a limited number of coarse-grained domains. While making this division we tried to take into account the amount of attention (amongst others measured by the number of qualitative publications on the subject) paid by researchers to the application area. The first application domain we chose to survey in this paper is FCA-based software mining, i.e. gaining insight into source code with FCA. One of the first papers applying concept lattices to software analysis (Krone & Snelting, 1994) analyzed the relationships between source code pieces and preprocessor variables in Unix system software. Later on multiple papers appeared on identifying modules or classes in legacy system code (e.g. Siff & Reps, 1997). More recently dynamic code analysis gained interest, e.g. Ammons, Mandelin, Bodík, and Larus (2003) analyzed execution traces which they clustered with FCA to debug specifications in temporal logic. In Snelting (2005), Hesse and Tilley (2005) an overview of FCA applications in software engineering published in 2003 or earlier can be found. In Tilley and Eklund (2007), an overview of 47 FCA-based software engineering papers is given. The authors categorized these papers according to the 10 categories as defined in the ISO 12207 software engineering standard and visualized them in a concept lattice.

The second domain we surveyed is FCA-based web mining. One of the first papers applying FCA to internet data was Krohn, Davies, and Weeks (1999) who analyzed the relationships between keywords used by communities of users to retrieve domain specific documents. A bit later, in Cole and Eklund (2001) another approach was presented in which FCA was applied for browsing web-documents. Cole and Stumme (2000) presented the Conceptual Email Manager to browse through email collections. Later on the Conceptual Email Manager was further developed to become Mail-Sleuth (Eklund, Ducrou, & Brawn 2004). Around the same time Carpineto and Romano (2005, 2004b) also highlighted the potential of FCA to be used as a meta search engine resulting in the CREDO system.

* Corresponding author at: KU Leuven, Faculty of Business and Economics, Naamsestraat 69, 3000 Leuven, Belgium.

E-mail addresses: Jonas.Poelmans@econ.kuleuven.be (J. Poelmans), dignatov@hse.ru (D.I. Ignatov), skuznetsov@hse.ru (S.O. Kuznetsov), Guido.Dedene@econ.kuleuven.be (G. Dedene).

0957-4174/$ - see front matter
http://dx.doi.org/10.1016/j.eswa.2013.05.009

The third domain we surveyed is FCA-based text mining and linguistics. Priss (1998) was one of the first works applying formal concept analysis in the domain of computational linguistics including the analysis of lexical databases and faceted thesauri structures. Mackensen and Wille (1999) used conceptual data systems for content analysis of unstructured texts. Hotho, Maedche, and Staab (2002) first clustered texts based on a thesaurus and k-means to reduce the number of documents and then analyzed these clusters with FCA.

The fourth domain we surveyed is FCA-based knowledge discovery in biology, chemistry and medicine. Cole and Eklund (1996) wrote one of the first papers applying formal concept analysis in combination with SNOMED (medical nomenclature system) for analyzing a large collection of medical discharge summaries. Schnabel (2002) used FCA to analyze diseases, treatments and symptoms to find implications and relationships hidden in the data. The first paper on applying FCA in biological research was Ganter and Wille (1989) where the authors proposed their conceptual scaling method. Brüggemann et al. (1997), Brüggemann and Bartel (1999) was one of the first papers which applied FCA in chemistry and more in particular for the analysis of environmental databases. Bartel and Brüggemann (1998) applied FCA to the analysis of structure activity relationships to predict the toxicity of chemical compounds.

Finally, we surveyed FCA-based research on ontology engineering. Richards and Compton (1997) was one of the first papers which used FCA in combination with ripple down rules to extract ontological vocabulary from knowledge bases. Stumme and Maedche (2001) proposed the first FCA based ontology merging algorithm. Jiang, Ogasawara, Endoh, and Sakurai (2003) used FCA in combination with natural language processing for building an ontology in the cardiovascular medicine domain.

The remainder of this paper is composed as follows. In Section 2 we describe recent papers on using FCA in text mining and linguistics. In Section 3 we survey recent papers on web mining applications using FCA. In Section 4 we discuss recent papers on applying FCA in software source code mining. In Section 5, we discuss recent papers on applying FCA in bio-informatics, chemistry and medicine. In Section 6 we survey the recent papers on using FCA in ontology construction and merging. In Section 7 a concise overview of other related application domains is given. Section 8 concludes the paper.

2. Text mining and linguistics

One of the most popular topics within the FCA research community is ontology engineering. Multiple techniques have been devised to semi-automatically build ontologies from unstructured or semi-structured texts and to improve their quality (see Section 6 of this paper). In this section however we zoom in on the FCA-based research in text mining which did not have the primary goal to construct an ontology but to gain insight in textual documents. We also give an overview of FCA-based research in linguistics. Tables 1 and 2 summarize the data and the underlying FCA-based models used by the authors during analysis.

The paper of Maille, Statler, and Chaudron (2005) describes initial steps in the development of a Kontex system for identifying causal factors in aeronautical incident reports. They use information, captured in a relational database, about Outcome and Initial Situation to cluster these reports. A generalized context (see cube lattice Section 5.3.2, Poelmans et al., 2013) which can be transformed in a regular FCA context and lattice is used to analyze this information expressed by first-order literals. Further research will focus on analyzing the free text part of the report. Roth, Obiedkov, and Kourie (2008a), Roth, Obiedkov, and Kourie (2008b) use FCA in combination with nested line diagrams and concept stability to identify communities of researchers in scientific papers on a well-defined topic. They use authors as objects and the notions which they mentioned in one or more of their papers as attributes. As a case study they looked at the zebrafish research community from which they analyzed 25 authors and 18 words. To retain the most important subcommunities they used concept stability (Kuznetsov, 2007) as a pruning technique and to improve and to make the visualization scalable for large communities they also used nested line diagrams. Kuznetsov, Obiedkov, and Roth (2007a) built further on this work and look more closely at the peculiarities of intensional and extensional concept stability and their practical relevance for analyzing knowledge communities. Intensional stability, as it was also used by Roth et al. (2008a), Roth et al. (2008b), measures how much a group of attributes depends on some of its individual members and is thus particularly useful for detecting unstable intents. The author shows the application to a context where the objects are scientists and the topics on which they work are attributes and the resulting concepts represent knowledge communities as groups of topics representative of a field along with corresponding scientists. The first case study zooms in on attendees of a conference and the second example is again the zebrafish context. The author also looks at extensional stability which indicates how a concept extent depends on particular attributes. The latter may be helpful for measuring how durable links between people within a community are. Empirical validation of the usefulness of extensional stability was performed on a dataset where objects are people and attributes indicate attendance to social events. Finally the author shows how concept lattices which were pruned using stability measures can be compared to identify emerging and declining topics in a community. Girault (2008) presents an unsupervised method for named entity annotation based on concept lattices, where FCA is used to analyze the relations between named entities and their syntactic dependencies observed in a training corpus. In Pan and Fang (2009) FCA was used to produce different levels of ontological concepts from radiology reports. The authors compared the ontology concept lattice of the radiology report's content before and after the adoption of the "picture archiving and communication system" and observed a delicate change in radiology report terms before and after the adoption of the system. Boutari, Carpineto, and Nicolussi (2010) use concept lattices from a document – term matrix to expand short texts, which pose difficulties to many traditional text mining algorithms. To overcome sparseness and the lack of shared terms, the authors analyze the relationships between term concepts in the document lattice and expand the texts with similar relevant words. They experiment with five similarity measures (proximity, concept similarity, connection strength, damping-weighted proximity, proximity and strength) and two classification techniques (k-NN classification and k-means clustering) and were able to improve classification performance. Dufour-Lussier, Lieber, Nauer, and Toussaint (2010) use FCA for adapting culinary recipes. The user enters an ingredient which he or she would like to use to replace another ingredient in a given recipe. Their system tries to find an ingredient preparation prototype, which is a sequence of culinary actions applied to the user-given ingredient, to replace the ingredient which the user specified in the given recipe. The authors first preprocess the given recipes into tree structures using amongst others part-of speech tagging and regular expressions. In case the user would like to replace an ingredient in a recipe by some other ingredient the system first identifies the subtree containing the ingredient to be replaced and then build a concept lattice to identify recipes containing a sequence of culinary actions performed to the ingredient which will replace the deleted one. This concept lattice will have as objects those recipes containing the user specified ingredient and can be used to identify the recipe

Table 1
Papers describing an application of FCA in text mining. Note that ontology-related research papers are described in Section 6.

Paper	Data	Methods	Research goal
Maille et al. (2005)	O = {incident reports} A = {Outcome + Initial situation attributes} I = {(o,a) ∈ I|incident report o has outcome or initial situation attribute a}	• Generalized cube lattice context (see Section 5.3.2 of part one of this survey) • FCA	Cluster incident reports based on their outcome and initial situation.
Kuznetsov et al. (2007)	Two single-valued contexts: 1. O = {scientists} A = {terms mentioned in article title and abstract} I = {(o,a) ∈ I|scientist o mentioned term a in the title or abstract of his paper related to the conference} 2. O = {persons} A = {social events} I = {(o,a) ∈ I|person o attended social event a }	• FCA • Intensional stability • Extensional stability	Analyze the structure of scientific communities
Roth et al. (2008a), Roth et al. (2008b)	O = {authors} A = {notions} I = {(o,a) ∈ I|author o used notion a in one of his papers}	• FCA • Concept stability • Nested line diagrams	Represent the structure of a knowledge community
Girault (2008)	O = {named entities} A = {syntactic co-texts + internal components} I = {(o,a) ∈ I|named entity o has syntactic co-text or internal component a}	• FCA • Curvilinear component analysis for visualization	Annotate named entities to make them less ambiguous
Carpineto et al. (2009)	O = {articles} A = {phrases} I = {(o,a) ∈ I|phrase a appears in article o}	• SVM with lattice based kernel	Automated text classification
Poelmans, Elzinga, Viaene, and Dedene (2008, 2009), Poelmans, Elzinga, Viaene, and Dedene (2010a)	O = {police reports} A = {phrases} I = {(o,a) ∈ I|phrase a appears in article o}	• FCA • Emergent Self Organizing Maps	• Explore textual information • Automated text classification • Improve expert knowledge about domestic violence
Pan et al. (2009)	O = {radiology reports} A = {phrases} I = {(o,a) ∈ I|report o contains phrase a}	• FCA	Comparison of radiology report content before and after adoption of PACS system
Boutari et al. (2010)	O = {short texts} A = {terms} I = {(o,a) ∈ I|term a appears in short text o}	• FCA • Concept similarity, proximity, connection strength, damping-weighted proximity, proximity and strength • k-NN classification, k-means clustering	Expanding short texts with additional terms to reduce context sparseness
Dufour-Lussier et al. (2010)	O = {recipes containing the query ingredient} A = {actions performed in the recipe to the query ingredient} I = {(o,a) ∈ I|culinary action a is performed in recipe o}	• FCA • Part-of speech tagging	Identify ingredients which can be used to replace an ingredient in an other recipe
Poelmans et al. (2010c, 2011b, 2012a), Elzinga et al. (2010)	O = {criminals}, A = {observed indications} I = {(o,a) ∈ I|criminal o has indication a}	• FCA • Temporal concept analysis	Identify human trafficking and terrorism suspects from observational police reports

Table 2
Papers describing an application of FCA in linguistics.

Paper	Data	Methods	Research goal
Priss (2004)	Single-valued context: O = {meanings}, A = {words} I = {(o,a) ∈ I|word a has meaning o}	• FCA • Neighborhood lattices	Exploration into possibilities of lattices for visually representing lexical databases
Old (2006)	O = {homographs}, A = {meaning} I = {(o,a) ∈ I|homograph o is associated to meaning a if o has meaning a}	• Neighborhood lattices • Type-10 chain components	Disambiguating homographs
Priss and Old, 2010	Single-valued context: O = {words}, A = {hypernyms} I = {(o,a) ∈ I|word o has hypernym a}	• FCA • Neighborhood lattices	Develop an FCA-based interface for browsing WordNet
Falk et al. (2010), Falk et al. (2011)	O = {French verbs}, A = {subcategorization frames} I = {(o,a) ∈ I|verb o is associated to frame a in an existing subcategorization lexicon}	• FCA • Concept stability, separation and probability	Extract most relevant verb-frame associations

containing the closest matching culinary action sequence to the one which was applied to the deleted ingredient. In a final step the text of that recipe is automatically modified to contain the new information.

FCA was used in several security text mining projects. The goal in each of these papers was to make an overload of information available in an intuitive visual format that may speed up and improve decision making by police investigators on where and when to act. In the first case study, with the Amsterdam-Amstelland police which started in 2007, FCA was used to analyze statements made by victims to the police. The concept of domestic violence was iteratively enriched and refined, resulting in an improved definition and highly accurate automated labeling of new incoming cases (Poelmans, Elzinga, Viaene, & Dedene, 2011a). Later on the authors made a shift to the millions of observational and very short police reports from which persons involved in human trafficking and terrorism were extracted (Elzinga, Poelmans, Viaene, Dedene, & Morsing 2010, Poelmans et al., 2012a). In these security applications, suspects were mentioned in multiple reports and a detailed profile of one suspect (and persons in his social network) depicted as a lattice, with timestamps of the observations as objects and indications as attributes helped to gain insight into his (their) threat to society (Poelmans et al., 2011b). Recently Temporal Concept Analysis (see Section 5.3.6 of part 1 of this survey) and its relational counterpart Temporal Relational Semantic Systems (TRSS, Wolff 2010) were successfully applied to the analysis of chat conversations (Elzinga, Wolff, Poelmans, Viaene, & Dedene 2012, see Section 3 of this paper).

Priss (2005b) sketched the possibilities of FCA in linguistics. Her paper discussed how FCA could be used for analyzing linguistic features, word sense disambiguation, ontology engineering, formalizing and merging lexical databases. Priss (2004) discusses how FCA can be used to visually represent lexical databases, i.e. organized collections of words in electronic form. She performed an in depth study of the structure of Roget's thesaurus and how the underlying tree-like structure and lexical relations can be optimally represented in FCA. The author starts from a context which has senses of words as objects and which has the words itself as attributes and crosses indicate if a word has a certain sense. Since the resulting concept lattice can become very large for Roget's thesaurus the author proposes to use neighborhood lattices. Starting from a given word she applies several times the so-called plus-operator which is used to obtain the semantic neighborhood of a word from the thesaurus. This semantic neighborhood consists of all words that share senses with the original word. This plus-operator can be applied one time or multiple times. In case it is applied once the attributes are all senses of the original word, in case it is applied multiple times the attributes will be all senses of all the words in the neighborhood. An additional result of this research is an online lattice-based system which can be used for exploring Roget's thesaurus. Priss and Old (2005) discuss how FCA can make the Semantic Mirrors Method (Tyrik, 1998) simpler to understand. The Semantic Mirrors Method is a method for automatic derivation of thesaurus entries from a word-aligned parallel corpus. The method is based on the construction of lattices of linguistic features. The authors also use FCA for conceptual exploration of a medium quality bilingual dictionary. Priss and Old (2006b) combined FCA with relational algebra to further formalize her approach for visually representing lexical databases. She shows how her approach can be used amongst others for analyzing words and their antonyms using concept lattices. Old (2006) uses FCA and in particular neighborhood lattices (Priss, 1996) to extract and visualize homographs (objects), and their meaning (attributes) with the aim of disambiguating them. Homographs are words with identical spellings but different origins and meanings. The author extracted 373 homographs from Roget's International Thesaurus (Berrey, 1962) and compared the performance of neighborhood lattices with Type-10 chain components. Priss and Old (2010) and Choi et al. (2008) build further on this work and investigate how FCA can be used not only for Roget's thesaurus but also for visualizing WordNet. WordNet has several structural differences from Roget's thesaurus which make it more difficult to apply neighborhood lattices directly. WordNet groups words into sets of synonyms and each set of synonyms belongs to a part-of speech namely a noun, a verb, an adjective or an adverb. Based on this part-of speech the synonyms set can participate in several semantic and lexical relationships. In contrast to Roget's thesaurus, in WordNet the synonym sets tend to be smaller and tend to intersect only in one or sometimes two words. Since the developers tried to deliberately avoid too much overlap between synonym sets. Therefore building neighborhood lattices which apply the plus-operator several times does not result in interesting lattices. Therefore the authors took into account the semantic relationships from WordNet while building the neighborhood lattices. They chose to use the hypernym relation which they used to build the hypernym neighborhood for a word. They selected a subset of about 17000 individual nouns and verbs and used these words as objects and their hypernyms as attributes. The results were concept lattices which could also be used to explore the WordNet lexical database. Zhang, Pei, and Chen (2007b) discuss the extraction of fuzzy linguistic summaries from a continuous information system. They use FCA in combination with degree theory to obtain these fuzzy linguistic summaries. Falk and Gardent (2011) and Falk et al. (2010) use French verbs as objects and subcategorization frames as attributes. A subcategorization frame characterizes the number and type of the syntactic arguments expected by a verb. The authors use FCA to cluster these verbs and used concept stability, separation and probability to extract the most relevant verb-frame associations.

3. Web mining

In this section we discuss FCA-based web mining and improving the quality of web search. Note that some of the papers reviewed in this section are closely related to information retrieval or text mining. The FCA based research on information retrieval has been summarized in Poelmans et al. (2012b). The papers on web mining and web search result improvement which we judged to be more related to the KDD field are summarized here. For text mining the reader is referred to Section 2 of this survey. While analyzing the papers on this topic, we were able to discover some popular subtopics. Tables 3–8 and Fig. 1 summarize the papers of this section together with their research topics and descriptions of their contents.

The first subtopic is periodic web personalization. Periodic web personalization aims at recommending the most relevant resources to a user during a specific time period by analyzing the periodic access patterns of the user from web usage logs. Cho and Richards (2004) developed a domain specific web search system which reuses keywords and web pages previously entered and visited by other persons. These keywords and web pages are stored in a cache and are used to build a Concept Tree Map in the form of a lattice. The query entered by a later user is matched against this lattice and relevant web pages belonging to the most relevant concept are proposed. If none of them is acceptable, neighboring concepts are proposed or the user can search herself for new information using a meta search engine and store new concepts for later use. Zhou, Hui, and Chang (2005) use an FCA-based model to mine association rules from web usage logs. The rules can be used by a web page recommendation engine which matches them with the user's recent browsing history. The authors compared the performance of FCA and a Apriori-based algorithm on two session datasets from Microsoft's Anonymous Web Data,

Table 3
FCA papers on web personalization.

Paper	Data	Methods	Research goal
Cho et al. (2004)	Single-valued context: O = {web documents} A = {user's query keywords} $I = \{(o, a) \in I$|web document o contains keyword $a\}$	• FCA • Caching	Search engine for a specific domain which reuses keywords and web pages previously used or visited by users
Zhou et al. (2005)	Web usage data, single-valued context: O = {browsing session} A = {accessed web pages} $I = \{(o,a) \in I$|during session o the page a was accessed}	• FCA • Association rules	Recent browsing history is mined to recommend web pages
Beydoun et al. (2007), (2008)	Web usage data, single-valued context: O = {web page URLs} A = {keywords of session names} $I = \{(o,a) \in I$|URL o was visited during a surfing trail labeled by the user as $a\}$	• FCA • Association rules	Surfing trails are used to build a lattice to recommend web pages

Table 4
FCA papers on web search result optimization.

Paper	Data	Methods	Research goal
Du et al. (2009)	Many-valued context: O = {web pages} A = {keywords} I = interval $\in$ [0,1] based on the term frequency-inverse document frequency (TF-IDF) score calculated for each keyword in the web page	FCA	Mining association rule for search result optimization
Ignatov et al. (2009)	Single-valued context: O = {documents} A = {shingles or fingerprints} $I = \{(o,a) \in I$|document o contains the fingerprint $a\}$	• FCA • Shingling	Near duplicate detection
Carpineto et al. (2004)	Web search results: O = {URLs} A = {terms in URL title and summary} $I = \{(o,a) \in I$|term appears in the title or summary of URL $o\}$	Iceberg lattice	CREDO system for browsing search results
Koester (2005), Koester (2006)	Web search results: O = {URLs} A = {terms in URL title and summary} $I = \{(o,a) \in I$|term appears in the title or summary of URL $o\}$	• FCA • Attribute ranking	FooCA system for browsing search results
Ducrou et al (2007), Dau et al. (2008)	Web search results: O = {URLs} A = {terms in URL title and summary} $I = \{(o,a) \in I$|term appears in the title or summary of URL $o\}$	FCA	SearchSleuth system for browsing search results
Kim et al. (2004), Kim et al. (2006)	O = {home pages of staff and research students} A = {keywords referring research topics} $I = \{(o,a) \in I$|user annotated home page o with keyword $a\}$	• FCA • Conceptual scales	KANavigator system for small web communities in specialized domains who wants to browse and annotate documents.

Table 5
FCA papers on mining emails.

Paper	Data	Methods	Research goal
Cole et al. (2003), Eklund et al. (2004)	Email archives: O = {emails} A = {virtual folders} $I = \{(o,a) \in I$|email o contains the attributes of virtual folder $a\}$	FCA	Conceptual Email Manager (CEM) and Mail-Sleuth software
Geng et al. (2008)	Single-valued context: O = {emails} A = {keywords} $I = \{(o,a) \in I$|email o contains the keyword $a\}$	• FCA • Concept selection using fuzzy membership value	Identifying topics in emails

which contains the pages accessed by users in the period of a week. They claim that with FCA tools they generated 60% fewer rules while maintaining comparable quality. Zhou, Hui, and Fong (2006) use fuzzy FCA for a similar purpose. Beydoun, Kultchitsky, and Manasseh (2007) introduce a system KAPUST which captures user trails as they search the internet. Such a trail consists of a sequence of URLs which were visited during a browsing session. This browsing session was annotated at the start by the user with keywords referring to the search topic. The authors then construct a semantic web structure from the trails and this semantic web structure is expressed as a concept lattice. It is used to give user recommendations in the form of categorized web page links which may lead to more focused search results. Beydoun (2008) further investigates the possibilities of applying this system for processing

Table 6
FCA papers on (web) service mining.

Paper	Data	Methods	Research goal
Bruno et al. (2005), Aversano et al. (2006)	O = {web services} A = {keywords} I = keyword a appears in the service interface of service o	• FCA • Natural language processing	Identify services which are more generic or specific than others Identifying service operations requiring common inputs or producing common outputs Understanding the relationships between operations of a complex service
Fenza et al. (2008), 2009)	Fuzzy context, web service descriptions: O = {web services} A = {ontological concepts} I = relevance of ontological concept a in the description of the capabilities of o	• Fuzzy C-Means clustering • Fuzzy FCA	Identifying relevant web services for a service request
Azmeh et al. (2010)	Single-valued context: O = {web services} A = {groups of similar operations} I = web service o offers the functionality represented by the corresponding group of similar operations a	•FCA •Similarity measures	Classify web services based on the operations they offer
Chollet et al. (2012)	Single-valued context: O = {services} A = {service technology, service functionalities and other non-functional properties required and / or provided by the service} I = service o has feature a	•FCA •Concept selection based on relevance	Fulfill a dynamic service request and select the most appropriate services

Table 7
Social media mining with FCA.

Paper	Data	Methods	Research goal
Ebner et al. (2010)	Single valued context :O = {twitter users} A = {keywords in tweets } I = {(o,a) ∈ I|keyword a appears in a tweet of user o}	• FCA • Keyword extraction	Analyzing the usefulness of twitter for external participants who want to follow a conference
Cuvelier et al. (2011)	Single valued context: O = {tweets} A = {keywords} I = {(o,a) ∈ I|tweet o contains keyword a}	• FCA • Topographic network of tags	Analyze content of tweets
Elzinga et al. (2012)	O = {chat conversations}, A = {pedophile keywords} I = {(o,a) ∈ I|pedophile keyword a is used in chat conversation o}	• FCA • Temporal Concept Analysis • Temporal Relational Semantic Systems • Nested line diagrams	Estimate threat level of a chat conversation

Table 8
Other papers on FCA-based web document clustering.

Paper	Data	Methods	Research goal
Okubo et al. (2006)	Single-valued context: O = {web documents} A = {feature terms} I = {(o,a) ∈ I|feature term a appears in web document o}	FCA	Web document clustering
Ignatov et al. (2008)	Single-valued context: O = {firms} A = {advertising terms} I = {(o,a) ∈ I|firm o bought the term a}	• FCA • Frequent itemset mining	FCA based recommender system

student's virtual surfing trails. In a case study the web pages visited by 12 students during 12 consecutive weeks were logged together with their browsing session names. The final lattice contained 225 web pages, 98 keywords and 109 concepts. Huang and Li (2008b) apply the method of Beydoun to recommend web pages to employees of an enterprise. An ontology containing information relevant to this enterprise is constructed from keywords entered while users were surfing.

The second closely related subtopic is search result optimization. The results returned by web search engines for a given query are typically formatted as a list of URLs accompanied by a document title, a short summary of the document. Several FCA based systems were developed for analyzing and exploring these search results. CREDO (Carpineto and Romano, 2004), FooCA (Koester 2005; Koester 2006) and SearchSleuth (Dau, Ducrou, & Eklund 2008; Ducrou & Eklund, 2007) build a context for each individual query which contains the result of the query as objects and the terms found in the title and summary of each result as attributes. The CREDO system then builds the top of the lattice digram which is represented as a tree and can be interactively explored by the user. FooCA shows the entire formal context to the user which gives flexibility in selecting objects and attributes (e.g. according to their ranking), applying stemming and stop word removal, etc. SearchSleuth does not display the entire lattice but focuses on the search concept, i.e. the concept derived from the query terms. The user can easily navigate to its upper and lower neighbors and

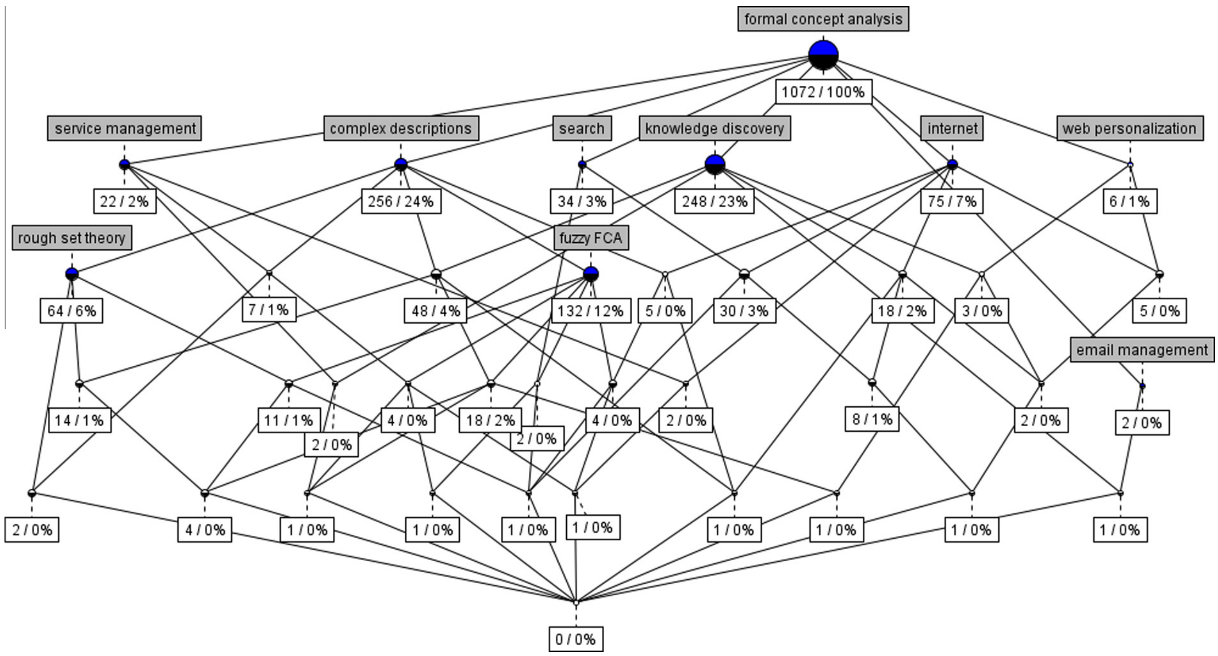

Fig. 1. Lattice diagram containing papers on FCA and web mining.

siblings. Kim and Compton (2004) presented FCA-based document navigation system for small web communities in specialized domains. Relevant documents can be annotated with keywords by the users. Kim and Compton (2006) extended the search functionality by combing lattice-based browsing with conceptual scales to reduce the complexity of the visualization.

Yang et al. (2008b) developed a topic-specific web crawler which only gathers particular pages related to a specific topic and traverses the web graph through a topic-relevant ordering instead of breadth- or depth-first. Assigning a proper prediction score to unvisited URLs is done using FCA and the semantic similarity of concepts. Starting from a concept lattice which has visited pages as objects, the core concepts which reflect the query topic are extracted and a so-called concept similarity graph depicts semantic similarity between core concepts and other concepts. Du et al. (2009) present a method based on FCA for mining association rules that can be used to match user queries with web pages to avoid returning irrelevant web pages for search engine results. The authors first divide webpages as follows: the so-called low-adjacence set consists of all web pages that contain all user-query words, the high-adjacence set consists of web pages which contain only part of the user-query words. To identify related and unrelated web pages in the high adjacence set they use a 3-stage process: constructing a concept lattice for the low-adjacence set, mining association rules between non-user query words and user-query words from this lattice and matching the user-query with web pages using association rules. Ignatov and Kuznetsov (2009) developed a method to detect near-duplicates in web search results where sets of duplicates (documents) are given by concept extents and similarity (common attributes) of documents are given by concept intents.

For mining large email archives, the Conceptual Email Manager (Cole et al., 2000, 2001, 2003) and Mail-Sleuth software (Eklund 2004) were developed. These plugins take emails as objects and "virtual folders" as attributes resulting in a lattice which is interactively explored. More recently Geng et al. (2008) took emails as objects, keywords as attributes and built a lattice to identify email topics. A fuzzy membership value was used during post-processing to extract the concepts which best represent a single topic.

A web service is a software entity that provides a set of functionalities described in a service description. Web services can be composed to form a software application. However, retrieving those web services meeting specifications given by a user is a challenging task. Several authors tried to use FCA for this purpose. Bruno, Canfora, Penta, and Scognamiglio (2005) and Aversano et al. (2006) classified web services using FCA. The service lattices which they generated represent web services and relationships between them using keywords which were extracted from the web services interfaces. This lattice can be used to identify services which perform the same type of task. The authors also use FCA for understanding complex services by analyzing service operations in the service description. During more detailed analysis, parameter names were extracted from individual service descriptions to understand the service behavior. Fenza, Loia, and Senatore (2008, 2009) present a system which uses fuzzy FCA for supporting the user in the discovery of semantic web services. This system is divided into lower and upper layers. In the lower layer, the semantic descriptions of the web services are transformed into fuzzy multisets. This description is an OWL-S document that sketches the capabilities of the service. Using Fuzzy C-Means clustering the web services are grouped in fuzzy clusters. Fuzzy matchmaking is performed to retrieve those services that are approximate replies to the input request. In the upper layer a fuzzy formal context is used to represent the prototypes (i.e. representative semantic web services for the clusters retrieved in the lower layer) and the given ontological concepts which are present or not. Navigating in the lattice, the user can then discover terminology associated to the web resources and may use it to generate an appropriate service request. Azmeh et al. (2010) generate three different types of concept lattices for classifying web services. The first concept lattice maps web services against the operations which they offer. In a second step they create a similarity matrix where objects are operations and attributes are also operations and the values between 0 and 1 indicate how similar the operations are to each other. Then they impose a threshold value to scale this many valued context into a binary context. Afterwards a concept lattice is created based on which operations which are similar are grouped together. Finally, a concept lattice containing web services as objects and these groups of operations as attributes can be used to make a final classification. This lattice shows which web services offer which operations but also show interesting links between different web services and their operations. Chollet et al. (2012) start from services which are the objects, these services can be of diverse nature and include web services. As attributes they use the service functionalities and also other non-functional properties which are required or provided by the service. The user makes a request which is in the form of a set of mandatory features and a set of optional features. After such a request is received a concept lattice is created however only the relevant concepts will be generated, namely concepts which only have non-functional features in the intent are not relevant. This results in improved efficiency for handling the request in real time.

Another emerging research topic is social media mining. Some FCA researchers worked on topics such as blog mining, twitter mining, etc. where the relationships between these web resources and semantic tags were analyzed. Kim, Hwang, and Kim (2007b) use FCA for mining internet blogs. They manually extracted tags which become attributes of the context and bloggers which are the objects of the context. From this lattice they extract bloggers with similar interests. Ebner et al. (2010) used FCA to analyze the content generated on twitter during and after a scientific conference. They first extracted keywords using Yahoo term extraction web service from the tweets and then used the twitter users as objects and the keywords in their tweets as attributes. FCA allowed them to categorize twitter users and to understand the usefulness of twitter for external participants who want to follow the topics of the conference. Using FCA they were able to identify four groups of tweets namely irrelevant, administrative and topical tweets as well as topical discussions. They however found out that only a very small percentage of the tweets was useful for external participants. Cuvelier et al. (2011) used FCA in combination with a topographic network consisting of keywords to analyze tweets on e-reputation posted on Twitter. In their analysis tweets were used as objects and keywords which they contained as attributes. To be able to explore this dataset with FCA they only visualized concepts with a sufficient support and complemented their analysis with a variation of a topographic map containing keywords and links between them if they were mentioned together. In both visualizations, keywords which appeared significantly more than others were given more weight in the visualization. Elzinga et al. (2012) applied FCA in combination with Temporal Concept Analysis and Temporal Relational Semantic Systems in the analysis of chat conversations between a pedophile and his victim. First the dataset was obtained from a public organization named Perverted Justice and was indexed using a thesaurus containing keywords specified by domain experts. From this initial concept lattice suspicious chat conversations were selected and Temporal Concept Analysis was used to gain insight in the evolution over time of these suspicious chat conversations. Temporal Concept Analysis made the different phases in the chat conversations visible and allowed for quickly gaining insight into their contents. To deal with possibly large lattice visualizations the authors also experimented with nested line diagrams.

Other papers applying FCA in web document clustering can be briefly described as follows. Myat and Hla (2005) use FCA to cluster textual web documents based on the terms they contain. The authors first extract these terms and then apply tf-idf index to determine their importance. Only terms with a score higher than a given threshold are used to build a concept lattice. Okubo and Haraguchi (2006) also use FCA for clustering web documents and providing a conceptual meaning for each document cluster. Ducrou (2007) presented DVD-Sleuth (based on Image-Sleuth, Ducrou 2006) which builds a concept lattice with DVD's as objects and description of these DVD's as attributes. Hsieh et al. (2007) parse user questions in natural language about basketball and use an FCA-based model to return relevant Chinese web pages. Ignatov and Kuznetsov (2008) used FCA to develop a recommender system for internet advertisement which suggests potentially interesting advertisement terms that can be bought. Chou and Mei (2008) used web APIs as objects and tags describing these APIs as attributes. They use fuzzy FCA to identify the best candidates for fulfilling a service request of a user. Kirchberg et al. (2012) performed an in-depth comparison of all performance aspects of analyzing large amounts of sematic web data, obtained from the Internet, in real-time. These aspects include preprocessing semantic web data into an FCA-readable format and concept generation algorithm scalability.

4. Software mining

35 of the 248 KDD papers are related to software mining and describe how FCA can be used to gain insight in amongst others software source code. The underlying model used by the majority of authors working in this field is FCA's single-valued formal context and the concept lattice derived from it. The software elements they used to build their FCA lattice differed however significantly and in Table 9 and Fig. 2 we try to give an overview of the FCA contexts authors used in their research.

In Eisenbarth, Koschke, and Simon (2003), a semi-automatic technique is presented for reconstructing the mapping of features that are triggered by the user to the source code of the system. A feature is a realized functional requirement of a system that makes use of multiple computational units (e.g. classes, routines, subsystems, etc.). Based on scenarios (corresponding to use case) which are sequences of user inputs that trigger actions of a system with observable results and the computational units they invoke, concept lattices are derived which can be used to reconstruct the feature-unit map. Mens and Tourwé (2005) use FCA to delve a system's source code for relevant concepts of interest. They use substrings of the names of source code entities to cluster them and were able to discover code duplication, which concerns are addressed in the code, which patterns, coding idioms and conventions were adopted and where and how are they implemented. In Cole, Tilley, and Ducrou (2005), FCA is used to conceptually analyse relational structures in software source code and to detect unnecessary dependencies between software parts. The lattices are based on the call graph that indicates which classes contain calls into which other classes. This information can for example help. Pfaltz (2006) uses FCA to identify casual dependencies from data containing execution traces as objects, executed operations as attributes and crosses indicating which operations are executed during which traces. They retain only logical implications for which the generator precedes the remainder of the consequent in all supporting trace sequences. In Cellier et al. (2008), FCA is used in combination with association rules for fault localization in software source code. The fault localization process starts with a trace context which has execution traces as objects add all the source code lines of the program as attributes. A cross in the formal context indicates a line of code is executed during a trace. Two additional attributes indicate whether or not execution of the trace failed. The author then searches for association rules based on closed itemsets which have a set of executed lines in the premise and the attribute FAIL in the conclusion. The algorithm used is described in Cellier (2007). In the second step the author uses a rule context in which objects are association rules and source code lines are attributes. Each association rule is described by the lines of its premise. The rules which are too specific to explain the error are at the bottom of the resulting lattice, more general rules are higher in the lattice. Wermelinger, Yu, and Strohmaier (2009) use FCA

Table 9
Formal Contexts and identified data structures in software mining papers.

Publication	Objects O	Attributes A	Incidence relation I	Identified Structures by means of concept lattices
Eisenbarth et al. (2003)	Computational units: classes, instructions, modules, subsystems, etc	Software scenarios	A pair (computational unit u, scenario s) is in relation I if u is executed when s is performed	• Relationships between scenarios and computational units • Relationships between scenarios and features and between features and computational units
Tonella et al. (2004)	Execution traces associated with use cases	Computational unit: class methods	A pair (execution trace e, class method u) is in relation I if u is executed by e.	• Refactoring options • Detection of crosscutting functionality
Mens et al. (2005)	Source code entities: classes, packages, methods, parameters, etc.	Substrings of names of chosen source code entities	A pair (source code entity c, name n) is in relation I if the name of c contains n	• Programming idioms: polymorphic methods, chained messages, delegating methods, • Code duplication • Design patterns • Relevant domain concepts• Opportunities for refactoring
Cole and Becker (2005), Cole et al. (2005)	1. Source code entities: classes, packages 2. Packages	1. Source code entities: classes, packages 2. Names in the path name of a package	1. A pair (source code entity a, source code entity b) is in relation I if b is directly or indirectly called by a during execution 2. A pair (package p, name n) is in relation I if the name of p contains n.	1. Class layers and cyclic dependencies 2. Insights in arrangement of packages
Pfaltz (2006)	Execution traces	Executed operations	A pair (execution trace e, operation o) is in relation I if o was executed during e	Causal dependencies between operations
Breu et al. (2006)	Locations of method calls	Methods	A pair (location l, method m) is in relation I if m is called at location l	Detection of cross-cutting concerns and aspect candidates
Del Grosso et al. (2007)	SQL queries	Fields selected from the database tables and conjuncts of query conditions	A pair (query Q, field f) is in relation I if f is part of the SELECT clause and a pair (query q, conjunct c) is in relation I if c is part of the WHERE clause of q	Possible features that can be exported as services in database-oriented applications.
Cellier et al. (2008)	1. Execution traces 2. Association rules	1. Lines of code and pass / fail attribute for the trace 2. Lines of code	1. A pair (trace t, line l) is in relation I if l is executed by trace t. 2. A pair (association rule a, line l) is in relation I if line l is in the premise of a.	Lines of code which cause execution of a software program to fail
Molloy et al. (2008)	Users of the systems	Permissions	The pair (user u, permission p) is in relation I if user u has permission p in the software system	User roles and role hierarchies
Wermelinger et al. (2009)	Source code files	Developers	A pair (source code file f, developer d) is in relation I if d worked on f during a specified period.	• Which developers have widest knowledge of system • Software parts that may be at risk of becoming legacy • Which developers can replace persons leaving the project

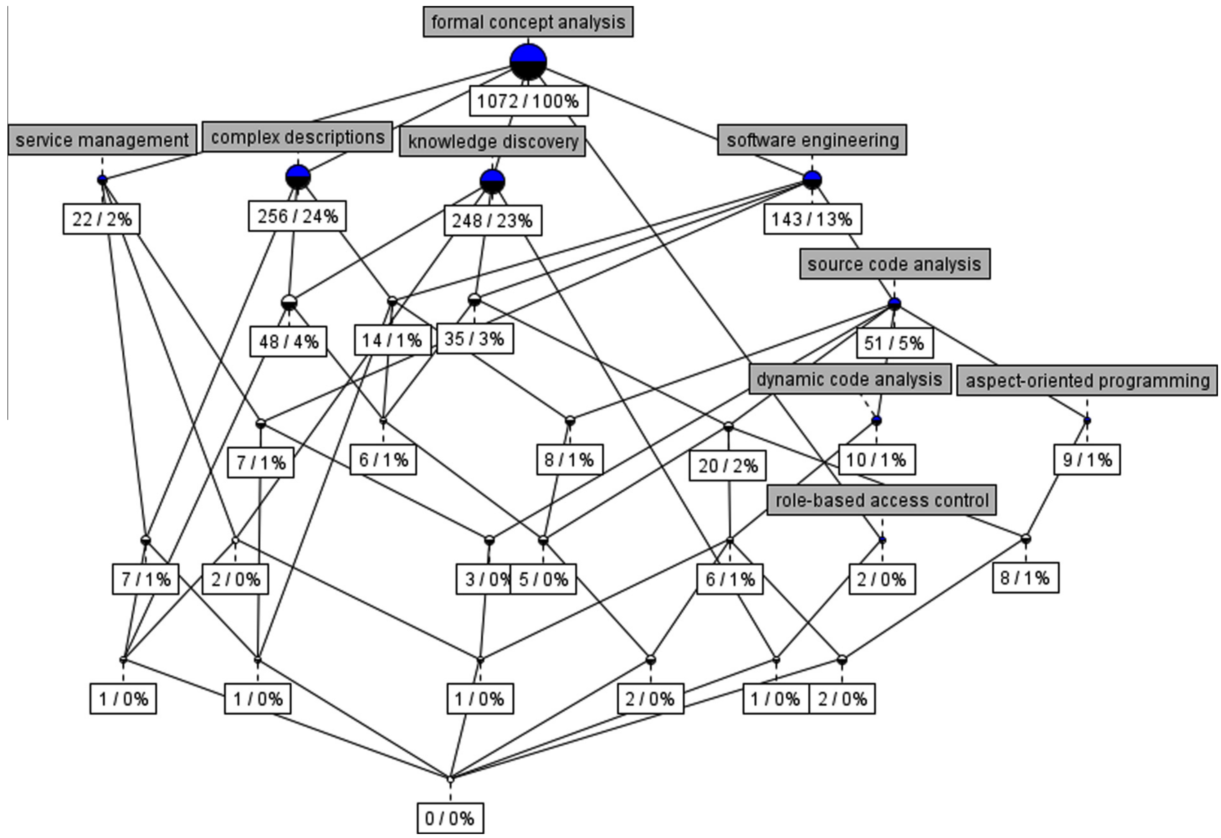

Fig. 2. Lattice diagram containing papers on FCA and software engineering.

lattices to visualize the relations between the software artifacts and to indicate developers who should fix the bugs in them. For example, objects of the context can be source code files, attributes are developers and the crosses indicate which developers worked on which files for a given period of analysis. The authors experimented with their method on the Bugzilla database for Eclipse and argue for the advantages of FCA with respect to other techniques on bi-partite and nested graphs.

Crosscutting concerns, i.e. functionalities that are not assigned to a single modular unit in the implementation, is one of the major problems in software evolution. Aspect Oriented Programming offers mechanisms to factor them out into a modular unit, called an aspect. In Tonella and Ceccato (2004), aspect identification in existing code is supported by means of dynamic code analysis. The work of the authors builds further on the work on feature location in source code in Eisenbarth et al. (2003). Execution traces are generated for the use cases that exercise the main functionalities of a given application. The relationship between execution traces and executed computational units is subjected to concept analysis. The invoked computational units that traverse system's use case models can be abstracted into potential aspects for re-engineering of the legacy system with Aspect Oriented System Design (AOSD). This improves the system's comprehensibility and enables migration of existing (object-oriented) programs to aspect-oriented ones. Breu, Zimmermann, and Lindig (2006) mined aspects from Eclipse by analyzing where developers added code to the program over time. A so-called transaction collects all code changes between two versions made by a programmer. Newly inserted method calls may result in cross-cutting concerns and aspect candidates. An aspect candidate consists of one or more calls to certain methods that are spread throughout the source code across several method locations. Other authors who worked on FCA-based aspect mining are Su, Zhou, and Zhang (2008) and Qu and Liu (2007). In Del Grosso, Penta, and Guzman (2007), an approach is proposed to identify pieces of functionality to be potentially exported as services from database-oriented applications. Database queries, dynamically extracted during the execution of the application over its main scenario, are clustered with FCA in concepts, potential services aiming to insert or retrieve information to or from the database are identified.

Role Based Access Control (RBAC) is a methodology for providing user roles in an IT system with specific permissions like read or write. Molloy et al. (2008) use FCA for mining roles from user-permission and user-attribute information to complement the costly top-down approaches for RBAC. The human-intensive top-down approach for constructing an RBAC system consists of performing a detailed analysis of business processes and deriving roles from them. Role mining is a bottom up approach that tries to discover roles from existing system configuration data. Dau and Knechtel (2009) apply FCA in combination with Description Logics to capture the RBAC constraints and for deriving additional constraints. The RBAC matrix is formalized as a triadic context from which dyadic contexts are derived. Using attribute exploration on these contexts, unintended implications between different roles, document types or permissions are derived and additional constraints are expressed with Description Logics.

5. Bioinformatics, chemistry and medicine

In this section we discuss applications of FCA in biology, chemistry or medicine. Tables 10–12 and Fig. 3 summarize the FCA papers with applications in these fields. The most popular topic in the bioinformatics domain is gene expression data analysis. In chemistry this is the analysis of the relationship between molecular

Table 10
Papers describing an application of FCA in bioinformatics.

Paper	Data	Methods	Research goal
Besson et al. (2004, 2005), Pensa et al. (2004a), Pensa et al. (2004b)	1. Gene expression data, many-valued context: O = {genes}, A = {biological situations }, I = expression value of gene o in biological situation ∈ [0…1] 2. Transcription factor regulation of genes, single-valued context: O = {genes}, A = {transcription factors }, I ={(o,a) ∈ I|transcription factor a regulates gene o}	• Concept mining • Constraints • Discretization	Mining concepts under constraints
Choi et al. (2006, 2008)	Gene expression data, 24 single-valued contexts: O = {genes}, A = {8 discretized gene expression values + 21 biological attributes}, I = {(o,a) ∈ I|gene o has attribute a}	• Discretization of continuous expression value • FCA • Lattice comparison based on common subgraphs	Identify biological relationships in gene expression data of influenza infected and healthy mouse lung tissue
Motameny et al (2008), Gebert et al. (2008)	Gene expression data, Many-valued context: O = {genes}, A = {situations}, I = Genes expression value ∈{1 …65535}	FCA with interordinal scaling	Identify biomarkers for breast cancer
Kaytoue et al. (2009)	Gene expression data, Many-valued context: O = {genes}, A = {situations}, I = Genes expression value ∈{1 …65535}	• FCA with interordinal scaling • Interval pattern structures	Identify co-expressed genes
Bertaux et al. (2009, 2011)	Macroscopic plant species living in water bodies data, fuzzy many-valued context: O = {species}, A = {traits}, I = affinity ∈{0…100}	• Histogram scaling: transformation to binary context • FCA	Identify ecological traits to assess water quality
Wollbold et al. (2009)	1. State context, single valued: O = {state labels} A = {gene regulators} I = {(o,a) ∈ I|gene regulator a is active in state o} 2. Transition context single valued: O = {transition between states} A = {gene regulators} I = {(o,a) ∈ I|gene regulator a is activated during state transition o}	• FCA • Attribute exploration	Build a knowledge base about the states and transitions of a gene regulatory network of a bacterium
Keller et al. (2012)	Single valued context: O = {gene sets}, A = {diseases}, I = {(o,a) ∈ I|gene o is significantly up regulated in biopsy tissue from subjects having disease a relative to tissue from living donors who do not have the disease}	• FCA	Determine disease similarity

structure and biological activity of chemical compounds (quantitative structure–activity relationship). In the medicine papers we see an emerging trend, namely applying FCA (in combination with some other techniques) to the analysis of time series data. Note that the papers on ontology engineering in the medical domain are described in Section 7. (See Fig. 4 and Table 13).

The mechanism that produces a protein from its gene is called gene expression and makes use of messenger RNA. A gene expression dataset is a many-valued context in which each row corresponds to a gene, each column corresponds to a sample and the attribute (expression) values indicate the abundance of mRNA in a sample. The first papers published on the application of FCA to the analysis of gene expression data were Besson et al. (2004, 2005). They introduced the D-miner algorithm which can be used to compute concepts under certain constraints (e.g. minimal and maximal frequency). Compared to other closed itemset mining techniques it performs better on dense Boolean datasets with large dimensions. Their method was empirically validated on UCI benchmark and gene expression datasets. The authors analyzed the expression values of 20000 genes before and after a perfusion of insulin in the skeletal muscle, 5 microarrays were recorded for healthy persons and 5 for type 2 diabetic patients. After discretization of these data and applying feature selection, the context with 8171 genes was merged with a Boolean context which indicates the transcription factors (proteins which specifically recognize and bind DNA regions of the genes) that regulate certain genes. In Potter (2005) a FCA-based method was developed for micro array data comparison. Choi, Laubenbacher, Duca, Lam, and Huang (2006) and Choi et al. (2008) build further on this work and start from micro array data representing the expression value of 11051 genes in the lung tissue of mice that were placed in 4 different situations (normal, infected by flu, in cigarette smoke, flu and in cigarette smoke). For each situation expression values were measured at 6 time points. The authors discretized these expression values in 8 Boolean attributes which were combined with 21 Boolean biological attributes in this case protein motif families. Concept lattices were created for each of the 24 (6×4) samples and were compared to each other using a distance metric based on common subgraphs. Motameny, Versmold, and Schmutzler (2008) and Gebert, Motameny, Faigle, Forst, and Schrader (2008)

Table 11
Papers describing an application of FCA in chemistry.

Paper	Data	Methods	Research goal
Richards et al. (2003)	Classification rules in knowledge base, single valued context: O = {classification rules' conclusions} A = {classification rules' conditions} I = {(o,a) ∈ I|the knowledge base contains a rule with a in the antecedent and o in the consequent}	• Ripple-Down Rules (RDR) • FCA	Restructure chemical pathology KBs by removing redundancy, generating higher level rules, etc.
Blinova et al. (2003), Ganter et al. (2004), Kuznetsov and Samokhin (2005),	Descriptions of molecular graphs Pattern structures (G, (D,), δ) on graphs: O = {names of compound} A = {molecular subgraphs} I = {(o,a) ∈ I|molecular subgraph a occurs in compound o}	• JSM-method • Pattern-based learning	Predict biological activity of chemical compounds
Lounkine et al. (2008)	Descriptions of molecular graphs and their activity O = {molecular fragments and combinations of fragments} A = {biological activity} I = {(o,a) ∈ I|molecular fragment o has activity a}	• FCA • Conceptual Scaling	Identify molecular fragments and fragment combinations which are specific for compound activity classes
Stumpfe et al. (2011)	O = {chemical compounds} A = {selectivity measured by potency ratio} I = {(o,a) ∈ I|compound o has a potency ratio larger than a certain threshold in the absence of a particular target a}	• FCA • Conceptual scaling	Analyze the relationships between the structure of chemical compounds and their selectivity towards certain targets

Table 12
Papers describing an application of FCA in medicine.

Paper	Data	Methods	Research goal
Sato et al. (2007)	Time series of test results single valued context: O = {patients} A = {daily text data} I = {(o,a) ∈ I|daily test data a appears in the test results of patient o}	• FCA • Weighted undirected graphs	Identify clusters of patients with similar test result data
Jay et al. (2008a), Jay et al. (2008b)	Network data, single valued context: O = {patients} A = {hospital} I = {(o,a) ∈ I|patient o has been treated in hospital a}	• Iceberg lattices • Stability indexes	Gain insight in cancer patient flows between hospitals
Poelmans et al. (2010)	Time series of activities performed to patients, single – valued context: O = {patients} A ={activities} I = {(o,a) ∈ I|activity a was performed to patient o}	• FCA • Hidden Markov Models	Identify quality of care issues for breast cancer patients
Egho et al. (2011)	Network data, single valued context: O = {patients} A ={sequences of hospitals} I = {(o,a) ∈ I|patient o was hospitalized sequentially in the hospitals in sequence a}	• FCA • Sequential pattern mining	Gain insight in cancer patient flows between hospitals
Sklenar et al. (2005), Sigmund et al. (2005), Belohlavek et al. (2007), (2011)	Physical activity questionnaire data, many-valued context: O = {respondents} A = {questions} I = respondent's answers to questions	• FCA with interordinal scaling • Fuzzy FCA	Identify dependencies between demographic data and degree of physical activity
Rouane-Hacene et al. (2009), Villerd et al. (2010)	Single valued context: O = {case reports} A ={taken drugs or adverse drug reaction} I = {(o,a) ∈ I|case report o describes a patient who took drugs a or had adverse drug reaction a}	• FCA • Iceberg lattice • Informative generic basis	Identify drug reaction combinations and drug interactions
Messai et al. (2011)	Two single valued contexts: O = {patient cases}A = {clinical information from the patient's profile and conformity of the decision with CPG} I = {(o,a) I|patient case o has clinical characteristic a }	• FCA • Conceptual scaling	Identifying patient characteristics which lead to noncompliance with CPG of physician in his decision

use an FCA-based model to identify combinatorial biomarkers of breast cancer from the expression values of 22215 genes, measured in 50 issue samples (20 metastasis, 28 primary tumors and 2 healthy). Combinatorial biomarkers are sets of genes that can distinguish healthy from cancer tissue or metastasis from primary tumor. In Kaytoue, Duplessis, Kuznetsov, and Napoli (2009), Kaytoue, Kuznetsov, Napoli, and Duplessis (2011), FCA in combination with interordinal scaling is compared to pattern structures (Ganter & Kuznetsov, 2001) based on interval vectors for mining and clustering gene expression data and extracting co-expressed genes. Using interval-based pattern structures is shown to be computationally efficient and to bring better interpretable results. Wollbold, Guthke, and Ganter (2008) start from time-series of mRNA and protein concentrations and derive using FCA a knowledge base consisting of a set of transition rules between states. Gene regulatory networks describe the interplay between the concentrations of different mRNA molecules and using the proposed system reasoning over temporal dependencies within gene regulatory

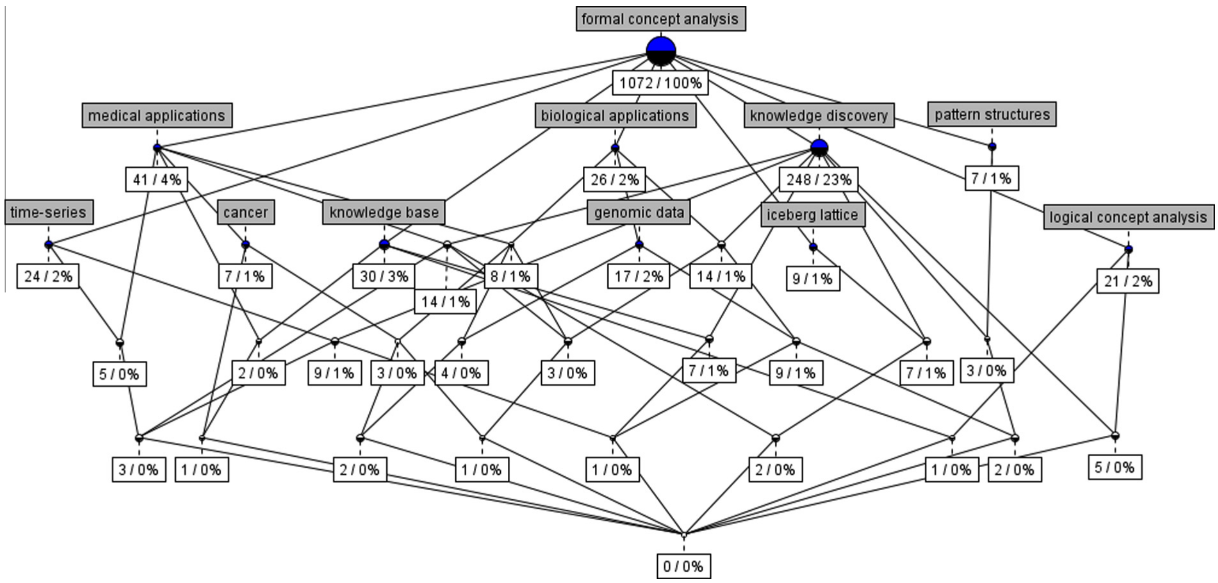

Fig. 3. Lattice diagram containing papers on FCA and applications in biology, chemistry or medicine.

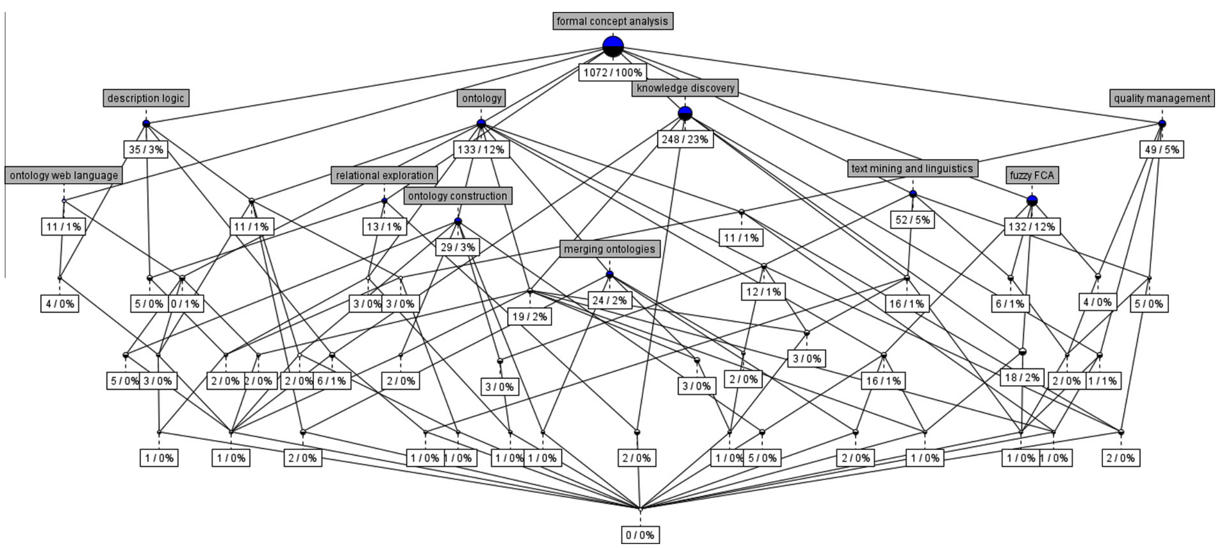

Fig. 4. Lattice diagram containing papers on FCA and ontologies.

networks becomes possible. Wollbold et al. (2009) use FCA and attribute exploration for building a knowledge base about a gene regulatory network of a bacterium. This bacterium has several states in which certain genes and gene regulators are active (represented by the author in the state context). Different exogenous factors may trigger a transition between two states and the transition context of the authors indicates which genes and gene regulators are activated during a transition. The transitive context derived from the transition context contains pairs of states such that the output state emerges from the input state by some transition sequence. On this context attribute exploration is performed to obtain a complete knowledge base. Wollbold et al. (2011) studied gene expression data of arthritic patients using Temporal CA. Time series of mRNA concentration levels in synovial cells from arthritic patients were measured and for small sets of interesting genes these data were represented as life tracks in transition diagrams. Bertaux, Le Ber, Braud, and Tremolieres (2009) describe a method to identify ecological traits of species based on the analysis of their biological characteristics. The complex structure of the dataset is first formalized by a fuzzy many-valued context. The concepts of the scaled context were interpreted by a hydrobiologist as sets of ecological traits. Keller, Eichinger, and Kretzler (2012) use FCA to determine disease similarity through analyzing the extent to which the gene sets which are significantly up regulated are overlapping. In their first analysis they used genes as objects and seven chronic renal diseases as attributes and the concept lattice showed which diseases shared which up regulated genes compared to healthy tissue. In their second experiment they used FCA to find genes whose protein products interact with those of the renal disease genes.

Richards and Malik (2003a) use their method described in Section 5.1 to restructure knowledge bases containing classification rules in the domain of chemical pathology. In Blinova, Dobrynin, Finn, Kuznetsov, and Pankratova (2003), Ganter, Grigoriev, Kuznetsov, and Samokhin (2004) and Kuznetsov and Samokhin (2005) biological activity of chemical compounds is studied with FCA tools. In Blinova et al. (2003) molecular graphs of the compounds are represented by attributes sets, where attributes are units

Table 13
Papers in which FCA is used for constructing an ontology.

Publication	Ontology language	Input	Method	Output
Jiang et al. (2003)	Japanese	• 368 textual discharge summaries from the cardiovascular medicine domain • Standard dictionary of Japanese diagnostic terms	1. NLP subsystem extracts diagnostic terms and medical compound phrases 2. FCA lattice is used to identify implications between medical concepts 3. validation by 5 clinicians of the results	1. 4724 compound medical phrases corresponding to medical concepts 2. 7666 semantic relations between medical concepts 3. 73% of compound medical phrases were meaningful, 57,7% of attribute implication pairs were relevant
Soon et al. (2004)	English	Appendix of European Water Framework Directive describing surface water monitoring domain	1. manual extraction of verbs and corresponding direct objects (nouns) 2. filtering to retain only verbs whose subjects refer to actors performing the monitoring activity 3. cross table with verbs as attribute and corresponding nouns as objects 4. action lattice which can be used to analyse the surface water monitoring domain 5. Entailment theory is used to capture additional semantic relations between verbs (actions)	Action lattice and additional entailments which can be used for creating task-oriented ontologies
Cimiano et al. (2005)	English and German	• Tourism domain: texts were acquired from 2 web sites and the British National Corpus, 118 million tokens • Finance domain: Reuters news from 1987, 185 million tokens	1. Part-of-speech tagging of corpus using TreeTagger and parsing of corpus with LoPar 2. Extraction of verb / subject, verb/object and verb/prepositional phrase dependencies from parse trees 3. FCA context has extracted nouns as objects and the corresponding verbs as attributes 4. Addressing data sparseness, removal of verb/ argument pairs with weight below a threshold 5. Pruning of lattice resulting in concept hierarchy 6. Comparison of generated concept hierarchy with results of hierarchical agglomerative clustering and Bi-Section-*k*-means	1. Two automatically acquired concept hierarchies from texts in finance and tourism domain 2. FCA has higher recall and similar precision than hierarchical agglomerative clustering and Bi-Section-*k*-means
Xu et al. (2006)	English1	DUC 2004 multi-document summarization dataset containing 50 sets with on average 10 documents, 149 event terms and 76 named entities each.	1. Preprocessing with GATE to extract named entities and verbs 2. Building a cross table in which event terms are objects and named entities are attributes. Event term a is related to named entity o if o is associated to a and to another named entity in the text 3. Event ontology is built from the lattice and relevance of events is determined using various metrics	1. Event ontology for the input texts 2. Event relevance measures
Gamallo et al. (2007)	English	• 25 million word tokens, ACL Anthology corpus consisting of articles published in the field of computational linguistics• Initial list of 175 terms	1. Clustering by abstraction: starting from some very specific classes relevant in the domain, i.e. few terms in extent and many lexico-syntactic contexts with which they co-occur in the intent, 2 clustering methods are subsequently applied to generate generic classes 2. Clustering by specification starting from very generic classes, a few lexico-syntactic contexts are now the extent and the words shared by those contexts the intent, again clustering is applied but this time to make the classes more specific 3. Validation by 3 human experts on the quality of the generic clusters and classification of new terms into their clusters	1. 201 generic classes containing 803 different terms, 600 new terms were extracted 2. 803 specific classes with 297 different terms 3. Only 5% of classification were judged as senseless
Bendaoud et al. (2008)	English	• SIMBAD database containing celestial objects and their properties• Collection of 11591 abstracts published between 1994 and 2002 in the Astronomy and Astrophysics journal.	1. NLP is used to extract celestial objects and their properties from paper abstracts 2. A concept lattice is built from the extracted object-attribute pairs. 3. A concept lattice with 470 objects and 92 properties is built from the information in the SIMBAD database. 4. Apposition of the text-based and SIMBAD -based lattices 5. Transformation of this final lattice into FLE-encoded ontology.	Enriched ontology represented in
Quan et al. (2006b)	English	9000 records of machine fault conditions reported by clients which are stored in a customer service	1. Extracting keywords and phrases from the service records	Machine service ontology containing information on machine faults and which web services may be used to

(*continued on next page*)

Table 13 (*continued*)

Publication	Ontology language	Input	Method	Output
		database of machine manufacturing company	2. Creating a fuzzy formal context which has problems reported by customers (i.e. machine faults) as objects, keywords (e.g. machine parts causing the fault) as attributes and the membership value reflects the degree of relatedness between reported problems and keywords 3. Creating a fuzzy fault concept lattice in which concepts have reported problems in extent and keywords in intent 4. Clustering of concepts into conceptual clusters using the author's own fuzzy clustering algorithm introduced in Quan et al. (2006a) resulting in fault concept hierarchy 5. Generation of machine service ontology to represent knowledge on machine faults 6. Comparison of recall, precision and F-measure of their FFCA-based method with 2 ANN and 2 kNN-based retrieval techniques for customer service support databases.	resolve their problems through automated semantic web-based customer service support system
Maio et al. (2009), Maio et al. (2012)	English	443 RSS feeds from OpenLearn project	1. RSS feed parsing to extract relevant keywords 2. Creation of fuzzy context indicating for each RSS feed the relevance of these keywords 3. Build a concept lattice using fuzzy FCA theory 4. Generate OWL- based ontology representation from this lattice 5. Validation and editing of resulting ontology by human expert	1. Concept lattice consisting of 193 concepts used for automatic categorization of OpenLearn RSS feeds 2. Comparison with classification performed by OpenLearn: 87% of feeds is classified coherently with OpenLearn's manual classification

of a specially designed Fragmentary Code of Substructure Superposition (FCSS) descriptor language. Hypotheses about structural classes of biological activity (toxicity) were generated as JSM-hypotheses. In Ganter et al. (2004) and Kuznetsov and Samokhin (2005) the authors use graph pattern structures in combination with several machine learning techniques such as JSM-hypotheses, decision trees, etc. to understand structural causes of biological activity and make predictions of activity. Lounkine, Auer, and Bajorath (2008) use FCA to analyze structure-activity relationships between compounds with overlapping biological activities. They use molecular subgraphs, also known as fragments, as objects and the activities which they have as attributes. The molecular subgraphs are obtained after applying hierarchical fragmentation of the original compound. Then conceptual scaling is used to reduce the complexity of the lattices. The set of fragments is subdivided into different subsets based on potency levels, the number of molecules in which a fragment combination occurs and based on their activity i.e. whether they are only present in active compounds or in inactive molecules. The goal of this research is to identify molecular fragments and fragment combinations which are specific for certain compound activity classes. Their method can also be applied to distinguish compound classes from one another and to search for active and highly potent compounds via simple substructure queries. Stumpfe, Lounkine, and Bajorath (2011) used FCA to analyze the relationships between the structure of chemical compounds and their selectivity towards certain targets. The authors used molecules as objects and bioactivity annotations as attributes. They subdivided the dataset consisting of compounds into several subsets based on their selectivity which is measured by their potency ratio in absence or presence of several targets. The goal of their research is to develop a method which can be used to query for compounds having a certain selectivity profile which allows for the detection of molecules which are not necessarily structurally similar to each other.

Sklenar, Zacpal, and Sigmund (2005) used FCA to evaluate epidemiological questionnaire physical activity data to find dependencies between demographic data and degree of physical activity. Belohlavek, Sklenar, Zacpal, and Sigmund (2007) build further on the work of Sklenar et al. (2005) and Sigmund, Zacpal, Sklenar, and Fromel (2005) by aggregating respondents and using fuzzy values to indicate the relative frequency of the attributes in the aggregated objects. Sato, Okubo, Haraguchi, and Kunifuji (2007) use FCA to find and explain similarities in the time series data related to medical test results. Case search systems may help diagnose a new patient by identifying past examples who had the same or similar set of symptoms, daily test results data etc. The formal concepts they extract from a weighted undirected graph, have patients in the extent and the daily test data they share in the intent. Jay et al. (2008) apply the stability index to choose better medical treatment trajectories for cancer patients. Their analyses also allow them to identify collaboration and flows of patients between hospitals. Egho, Jay, Raissi, and Napoli (2011) build further on this work and apply sequential pattern mining to this data resulting in sequences of hospitals (represented by attributes) where patients (represented by objects) were hospitalized during their cancer treatment. Poelmans et al. (2010) combine FCA with Hidden Markov Models to analyze patient – care activity data. Hidden Markov Models are used to identify the most frequent standard care pathway as well as care pathway deviations and exceptional situations. After identifying anomalies in these process models the authors use FCA to identify and explain the root causes for these anomalies. The authors found multiple quality of care issues and key interventions not performed to breast cancer patients. The resulting information formed the basis for optimization of the delivered care. Rouane-Hacene, Toussaint, and Valtchev (2009) and Villerd, Toussaint, and Lillo-Le Louët (2010) used FCA for analyzing case reports on adverse reactions to medicines. These reports capture patient characteristics including demographic data, a description of the observed adverse reaction and the suspected drugs. The goal of their analysis is two-fold namely to identify safety signals, i.e. drug reaction combinations and drug interactions. In particular, combinations of several drugs leading to a single adverse drug reaction are of interest. The objects in their study are case reports and the attributes are the taken drugs before the observed reactions. From this formal context they derive an iceberg lattice and the informative generic basis containing exact association rules. They validated their study on the spontaneous reporting system database of the French Medicine Agency. Messai, Bouaud, Aufaure, Zelek, and Séroussi (2011) use FCA to identify patient-related characteristics which may lead to noncompliance with Clinical Practice Guidelines (CPG) of the physician in his decisions. The authors used two datasets: one where decisions were made during multidisciplinary staff meetings where no clinical decision support system was used, a second containing cases where decisions were made during multidisciplinary staff meetings where a clinical decision support system was used. The objects were patient cases the attributes were clinical characteristics of these cases as well as information on the conformity of the decision of the physicians with respect to CPG. They first used conceptual scaling to transform this information into single valued contexts and then reduced the set of attributes in these single valued contexts to keep only those attributes which are related only to nonconformity or conformity. In these reduced lattices the authors looked for concepts which would cause nonconformity of multidisciplinary staff meetings with CPG in the domain of breast cancer management. They found that in case no clinical decision support system was in use, compliance with CPG is only the case for easy patient cases. For patients with a more serious condition there was typically no compliance with CPG. After the introduction of a clinical decision support system the number of cases for which there were no compliant decisions drastically reduced.

6. FCA in ontology engineering

Ontologies were introduced as a means of formally representing systems of concepts constituting human knowledge (Gruber 1995). Their purpose is to model a shared understanding of the reality as perceived by some individuals in order to support knowledge intensive applications (Gruber 2009). An ontology typically consists of individuals or objects, classes, attributes, relations between individuals and classes or other individuals, function terms, rules, axioms, restrictions and events. The set of objects that can be represented is called the universe of discourse. The axioms are assertions in a logical form that together comprise the overall theory that the ontology describes in its domain of application. Ontologies are typically encoded using ontology languages, such as the Ontology Web Language (OWL). Whereas ontologies often use hierarchical representations for modeling the world, FCA has the benefit of a non-tree hierarchical partial order representation which has a larger expressive power (Christopher 1965). A key objective of the semantic web is to provide machine interpretable descriptions of web services so that other software agents can use them without having any prior "built-in" knowledge about how to invoke them. Ontologies play a prominent role in the semantic web where they provide semantic information for assisting communication among heterogeneous information repositories.

As we can see in Fig. 4, 12% of the FCA papers discuss ontology-related research topics. In Section 7.1 we zoom in on the construction of ontologies using FCA, covering 29 of the 133 papers. 11 of the papers are about improving the quality of ontologies and are discussed in Section 7.2. In Section 7.3 we describe the 24 papers using FCA in ontology mapping and merging. 16 of the papers use fuzzy theory in combination with FCA for ontology construction or merging.

6.1. Ontology construction

An important topic in the FCA literature is how ontologies can be designed in an efficient manner. The unifying theme across ontology acquisition approaches is the considerable effort associated with developing, validating and connecting ontologies. The authors used FCA mostly as an ontological concept as an engine for extraction of concept hierarchies. The majority of them work with unstructured texts such as medical discharge summaries, RSS feeds, scientific papers, etc. Also semi-structured and structured information is used such as the information available in digital archive of a museum. If authors analyse unstructured texts they often use NLP tools for data preprocessing. With NLP they extract terms, phrases, lexico-syntactic contexts etc. from a corpus of texts. Sometimes they just start with a given initial set of terms. A concept lattice can be derived from this data and used for extracting ontological classes of terms, a hierarchical ordering of these concepts, implications between classes etc. Finally, this new ontological knowledge can be stored in e.g. OWL and new incoming texts can be categorized using this ontology. DL can be used to encode domain axioms and for reasoning purposes. Since the initial publications by amongst others Debbie Richards (University of Macquarie Australia), Philipp Cimiano (University of Karlsruhe) and Andreas Hotho (University of Kassel), FCA became a popular instrument for building ontologies all over the world. Cimiano, Hotho, Stumme, and Tane (2004), discuss and present several examples on how FCA can be used to support ontology engineering and how ontologies can be exploited in FCA applications. Richards (2006) promotes using FCA for building personal and ad hoc ontologies which may help gaining understanding of the research domain. Just like in software engineering such a prototype can be rapidly developed at a relatively low cost and stimulate exploration of domain and individual specific concepts.

Jiang et al. (2003) use FCA in combination with NLP for semi-automatically building a Japanese ontology in the cardiovascular medicine domain. As input the authors used textual discharge summaries and a standard dictionary of Japanese diagnostic terms MEDIS. Their NLP subsystem is based on the Japanese morphological analysis system ChaSen which they used for domain dependent term recognition. On top of ChaSen they introduced a heuristic algorithm for extracting compound nouns, also called compound medical phrases consisting of more than one noun. With FCA they extracted attribute implications and semantic relations between medical concepts and compound medical phrases. The relevance of these medical concepts and implications was validated by a medical expert panel. Soon and Kuhn (2004) use FCA for producing task-oriented ontologies. Verbs and nouns are extracted from a document that depicts user actions during a surface water monitoring process. The resulting so-called action context, has verbs as attributes, the nouns as objects and a cross in the context indicates the co-occurrence of respective verb and noun in the text. Entailment theory (Fellbaum 1990) is used to model additional semantic relations between verbs (actions). The action lattice and these entailments may be used for building an ontology describing the actions which are to be undertaken in the surface water monitoring domain. Cimiano, Hotho, and Staab (2005) applied FCA and NLP to textual data from the tourism and finance domain to automatically learn ontological concept hierarchies. First these texts were Part-of-Speech tagged, i.e. each word is assigned its syntactic category (noun, adjective, verb etc.), and then parsed, resulting in a parse tree for each sentence. The syntactic dependencies verb/object, verb/subject and verb/prepositional phrase are derived from these trees and used to create a context in which the verbs are attributes and the nouns are objects. A weighting and pruning of pairs below a fixed weight threshold is performed to address data sparseness issues. The result is a lattice from which a concept hierarchy can be automatically derived. The authors found FCA had a better recall and similar precision than the hierarchical agglomerative clustering and Bi-Section-*k*-means methods. In Xu, Li, Wu, Li, and Yuan (2006) FCA is used to build an event ontology from a set of textual documents. An event is defined as a triple consisting of an event term (a verb or action-denoting noun) and two named entities which it relates. These event components are extracted from texts using GATE and related to each other in a context. From the resulting lattice, an event ontology for the input data is derived. The authors explore various relevance measures to derive event relevance. Event relevance is important in event-based document summarization, which attempts to select and organize the sentences in a summary with respect to the events that the sentences describe. Gamallo, Lopes, and Agustini (2007) start from 25 million word tokens from technical articles in the field of computational linguistics and an initial list of 175 terms. The authors extract lexico-syntactic contexts of these terms and the formal context indicates which terms co-occur with which lexico-syntactic contexts. Using their so-called abstraction and specification operators inspired by FCA theory these terms are grouped in semantic classes. The resulting clusters can be represented by a concept lattice where a concept intent of this lattice contains lexico-syntactic contexts which co-occur in the corpus with all the terms in the extent. Bendaoud, Toussaint, and Napoli (2008) propose an FCA-based system for semi-automatically enriching an initial ontology from a collection of texts in the astronomy domain. The authors encoded existing domain knowledge in a concept formal context with celestial objects as objects and types of celestial objects as attributes. Using text mining techniques they constructed a context with celestial bodies as objects and phrases in the texts as attributes and then merged the two contexts by apposition. From the concept lattice of this merged context, an enriched ontology represented in Description Logics was derived. In (Richards 2004) the author showcases her method based on Ripple-Down Rules and FCA (described in Section 5) in the biology domain on four knowledge bases, about a plant which she first merges and then analyzes using FCA. From the concept lattice an ontology based on RDR knowledge base is reverse engineered.

Other attempts using FCA for building an ontology include the following papers which presented some interesting examples and background knowledge about various application domains. Chang (2007) start from 525 documents from the Electronic Theses and Dissertation System and use the keywords assigned by the authors as attributes. Using FCA they derive an ontology which they combine in Chang and Huang (2008) with Naïve Bayes for classifying unseen documents. Fang, Chang, and Chi (2007) integrate FCA with Protégé to build an ontology-based knowledge sharing platform, containing information about the acupuncture points from traditional Chinese medicine, for patients and physicians. Xu and Xiao (2009) contemplate on how FCA can be used to build a computer network management information specification ontology. Bao, Zhou, and He (2005) discuss how an iterative ontology building process using FCA for the construction of a pressure component design ontology would look like. Chi, Hsu, and Yang (2005) want to construct ontological knowledge bases for digital archive systems. The authors intend to use FCA for concept extraction, and OWL in combination with DL for presenting knowledge and reasoning.

Another recent research direction which is gaining interest (especially in China) is using fuzzy and rough FCA for ontology engineering. About 14% of papers belong to this direction. Quan, Hui, Fong, and Cao (2004b) use fuzzy FCA for the automatic generation of ontologies. These ontologies are used to support the Scholarly Semantic Web, which is used for sharing, reuse and management of scholarly information. Quan, Hui, and Fong (2006a) propose a method based on Fuzzy FCA, which they call Fuzzy Ontology Generation Framework, to automatically generate an ontology. In Quan et al. (2006b) the authors apply this method to build an ontology which can be used in a web-based help-desk application. In a case study they analyzed 9000 records stored in the customer service support database of a machine manufacturing company. Each of these records contains the description of a machine failure reported by a customer and proposed remedies to resolve the problem. Keywords such as certain machine parts can be extracted from this piece of text. A fuzzy formal context will then relate these machine failure cases (objects) with keywords explaining the nature of the failure (attributes) through a membership value indicating for example the possibility of a machine part being involved in a failure case. From this context a fuzzy fault concept lattice is created and a fault concept hierarchy is built from it by applying fuzzy clustering of the concepts. This hierarchy is then translated into an ontology which can be used to automatically suggest actions which can be undertaken to resolve a failure. Zhou, Liu, and Zhao (2007a) present an approach which is similar to this work. The authors derive a concept hierarchy from a fuzzy concept lattice which they applied as classification instrument on 13 datasets of UCI Machine Learning Repository. Maio, Fenza, Loia, and Senatore (2009), Maio et al. (2012) use fuzzy FCA to build an ontology for automatically classifying RSS feeds. From these feeds relevant keywords are extracted and a fuzzy context is built. The concept lattice constructed from this context is then automatically translated to an OWL ontology. A validation experiment was performed with 443 feeds from the OpenLearn project, containing the Open University's course materials, which were manually arranged in categories based on educational subjects. The authors achieved good performance with their method, 87% of feeds were categorized coherently with OpenLearn's manual classification. There are only a few researchers investigating the possibilities of applying Rough FCA to construct ontologies. Huang and Zhu (2008a) proposed rough FCA for semi-automatically constructing a marine domain ontology.

Hwang, Kim, and Yang (2005) uses FCA for the construction of ontologies in the domain of software engineering. There are many conceptual similarities between the design of a class hierarchy in OO software design and ontology. An OO software designer can design ontology by organizing classes in a class hierarchy and creating relationships among classes. UML classes can then be generated from the ontology. Kiu and Lee (2008) use FCA for managing existing ontological knowledge instead of building an ontology. They present a tool which allows the user to add, update and delete ontological concepts in 3 distinct ways. For small ontologies, the formal context itself can be edited. In the second case a SOM is used to cluster semantically similar ontological concepts and the user can choose the desired cluster from this visual interface. Clustering the ontology into clusters using SOM and k-means effectively reduces the search space for creating ontological concepts and querying documents. A concept lattice for this cluster is created and can be used to add, update or delete ontological concepts. The user interface OntoVis was introduced earlier in Lim and Lee (2005). Finally the user can also supply a keyword as a query for finding the nearest matching clusters in the SOM and again a lattice is created for the desired cluster.

Some researchers focused on various aspects of ontology engineering which may help future FCA based systems. For example Nazri, Shamsudin, and Bakar (2008) were inspired by the work of Cimiano et al. (2005), Cimiano (2006) to learn concept hierarchies and ontologies from texts. The authors compared the performance of 3 existing NLP tools which could be used for preprocessing Malay texts for FCA-based analysis.

6.2. Ontology quality management

Quality management is another important topic in ontology engineering and is covered by 10% of the papers. Rudolph (2004) proposes an incremental method based on FCA which uses empirical data to systematically generate hypothetical axioms about the domain of interest, which are presented to an ontology engineer for decision. The author focused on the axioms that can be expressed as entailment statements in description logic. In the position paper Rudolph, Völker, and Hitzler (2007), the author build further on this work and proposes Relational Exploration, a technique based on attribute exploration (Ganter & Wille, 1999), in combination with Text2Onto (NLP) for semi-automatically building an ontology. Text2Onto is a tool for lexical ontology learning which suggests the extraction of potential axioms and facts from texts. During the extraction some valid and relevant pieces of knowledge present in the text may be missed and it is also possible that the corpus itself does not contain all valid knowledge. Relational Exploration can be used to improve the precision and completeness of the learned ontology. Völker and Rudolph (2008b) combine the LExO approach for learning expressive ontology axioms from textual definitions with Relational Exploration to interactively clarify underspecified logical dependencies. The exploration guarantees completeness with respect to a certain logical fragment and increases the overall quality of the ontology. Rudolph (2008) proposes logical completeness as an important quality criterion for ontologies. The author defines a class of OWL axioms called "generalized domain range restrictions" which generalize domain and range statements commonly known from diverse ontology modeling approaches. FCA and role exploration are used to interactively specify all axioms of this form valid in a domain of interest. Kim et al. (2007a) propose to extract ontological elements from OWL source code and create a context family composed of five kinds of contexts. Using FCA they derive lattices which can be used to analyze the relations between e.g. classes and properties, classes and individuals, etc. and identify structural problems in the ontology. Sertkaya (2009) describes OntoComp which supports ontology engineers in checking whether an OWL ontology contains all the relevant information about the application domain and in extending the ontology if this is not the case. Using FCA, it acquires completer knowledge about the application domain by asking successive questions to the ontology engineer. Jiang, Pathak, and Chute (2009) audit the completeness and correctness of the International Classification of Disease (ICD) codes. They used ICD codes within a certain domain as objects for their analysis and atomic terms derived from the index entries as attributes. They used amongst others so-called anonymous nodes (nodes without own attributes) to identify incompletenesses in the ICD formalization. Afterwards they also transformed these concept lattices into Semantic Web Rule Language (SWRL) rules. Jiang and Chute (2009b) also used FCA to analyze the completeness and correctness of SNOMED which is a controlled vocabulary for the medical domain. Again they suggest to track down anonymous nodes appearing in the concept lattices which can be used to identify concept gaps in the original controlled vocabulary. The authors argue they may point to a missing concept.

6.3. Ontology mapping and merging

Ontology mapping is a key technology to resolve interoperability issues between heterogeneous and distributed ontologies. Ontologies in the same domain or overlapping fields can be built with different representations or different names for the same concept or different structures for the same domain. This topic is discussed in 17% of the papers on FCA and ontologies.

One of the first papers on this topic was Stumme and Maedche (2001) who presented the FCA-Merge algorithm. De Souza and Davis (2004a,b) use FCA for merging ontologies that cover overlapping domains, paying particular attention to the design of good similarity measures for the identification of cross-ontology related concepts. In De Souza, Davis, Evangelista, and (2006), the authors try to apply their alignment method to ontologies developed for completely different domains. Fan and Xiao (2007) propose an FCA-based method for ontology mapping, which can not only perform mapping by computing similarity measures between entities of different ontologies but can also perform subclass mapping by computing inclusion measures. Formica (2006) also discusses how FCA can promote reuse of independently developed domain ontologies. The author proposes an ontology-based method for assessing similarity between FCA concepts to support the ontology engineer in ontology merging and mapping. This work is further refined in Formica (2008) where the similarity measure is made independent of the domain expert knowledge. The author uses the information content approach (Resnik 1995) to automatically obtain attribute similarity scores.

Curé and Jeansoulin (2008) present an automated approach using FCA for creating a merged ontology from two source ontologies. Wang, Du, and Chen (2009) use FCA to compute the Concept-Concept similarity, the Concept-Ontology similarity and the Ontology-Ontology similarity for coordinating two Agent Crawlers and deducing the level of understanding between them, to guide them as parts of a search engine. Lu and Zhang (2008) present a Description Logic (DL)-based approach for conflict detection and elimination between two equivalent concepts in different source ontologies which may have different definitions of value and cardinality restriction. Le Grand et al. (2009) use FCA for complex systems analysis and compare different topic maps with each other both in terms of content and structure. Significant concepts and relationships can be identified and this method can also be used to compare the underlying ontologies or datasets.

Krötzsch et al. (2005) discuss morphisms, a general tool for modeling complex relationships between mathematical objects, in FCA and propose approaches in ontology research where morphisms help formalize the interplay among distributed knowledge bases. Zhao, Halang, and Wang (2007) propose an ontology mapping method based on rough FCA and a rough similarity measure which they introduced. The two source ontologies are first transformed into formal contexts and are then merged to obtain a concept lattice.

7. Other application domains

FCA has been used in several other interesting application domains. Busch and Richards (2004) use FCA for the analysis of psychological data. The authors developed a tacit knowledge inventory based on the measurement of responses to IT work-place scenarios, which is part of a questionnaire given to experts and non-experts in three IT organizations. Using FCA they were able to identify important groups of individuals that responded similarly to the peer-identified experts and the organisation was alerted of the important role these individuals potentially play. Ducrou et al. (2005a) developed a conceptual information system to determine surfing conditions on the South Coast of New South Wales in Australia. In Colton and Wagner (2007), FCA is used in combination with mathematical discovery tools to better facilitate mathematical discovery. Missaoui and Kwuida (2009) investigate the possibilities of using FCA for extracting actionable knowledge from data warehouses containing multidimensional data. In Hauff and Deogun (2007) concept lattices are used for disjoint clustering of transactional databases and several heuristics are developed to tune the support parameters used in the algorithm. The algorithm is applied to location learning to estimate the location of an electronic tag (in an RFID for example) given the signal strengths that can be heard. Solesvik and Encheva (2009) use FCA as a quantitative instrument for partner selection in the context of collaborative ship design. Ignatov, Mamedova, Romashkin, and Shamshurin (2011) built lattice-based taxonomies to represent the structure of student assessment data to identify the most stable student groups w.r.t the students achievements (and dually for course marks) at certain periods of time and tracked the changes in the states of these groups over time. Romashkin, Ignatov, and Kolotova (2011) analyzed university applications using lattice-based taxonomies derived from entrants' decisions about undergraduate programs. Admission data as well as formalized survey data were used to reveal possibly significant factors of entrants' decisions. Priss, Riegler, and Jensen (2012) used FCA to gain insight into the conceptual structure of difficulties students may encounter in their learning processes. Endres, Adam, Giese, and Noppeney (2012) analyzed the results of fMRI scans obtained from a human subject while he was viewing 72 gray-scale pictures of animate and inanimate objects during a target detection task. The pictures obtained were used as FCA objects. Attributes were obtained by learning a hierarchical Bayesian classifier, which maps BOLD (an indirect measure for neural activity) responses onto binary features, and these features onto object labels. The connectivity matrix between the binary features and the object labels then served as the formal context.

8. Conclusions

Since the introduction of Galois lattices in 1970s and FCA in 1982 concept lattices became a well-known instrument in computer science. Over 1000 papers have been published on FCA during the last 9 years on FCA and many of them contained case studies showing the method's usefulness in real-life practice. This paper showcased the possibilities of FCA as a meta technique for categorizing the literature on concept analysis. The intuitive visual interface of the concept lattices allowed for an in-depth exploration of the main topics in FCA research. In particular, its combination with text mining methods resulted in a powerful synergy of automated text analysis and human control over the discovery process.

The first domain we surveyed was text mining and linguistics using FCA methods. Over the past years FCA has been applied in several text mining projects ranging from the adaptation of cooking recipes to identifying criminals in hundreds of thousands unstructured textual police reports. Although the majority of the papers surveyed in Section 2 on text mining describe a proof of concept for an FCA-based text mining system, they clearly showed the potential of FCA in this area. FCA also had several applications in linguistics where the theory was used to visualize and gain insight in lexical databases such as Roget's thesaurus and WordNet. The second domain which we chose to survey is web mining to which a lot of attention has been devoted also. In particular web search result optimization and web personalization received considerable interest by the FCA community. Other topics which received slightly less attention but became popular quite recently are web service mining and social media mining. Several of these papers described high quality research that resulted in practical systems which are available to end users. Also software mining is a very popular topic in the community. Initially attention was mostly devoted to static source code mining but over the years more papers on applying FCA to dynamic code analysis appeared in the literature. The fourth domain which we chose to survey are the applications of FCA in life sciences research. In biology we found many applications of FCA to the analysis of gene expression data. Authors not only used basic FCA structures such as a concept lattice but also more complex descriptions of FCA including pattern structures. In medicine FCA has been applied amongst others to time series data, to questionnaire data and to textual data often in combination with other techniques such as sequential pattern mining, stability-based pruning, Hidden Markov Models, etc. Several of these papers described real-life applications with real outcome affecting daily healthcare practice. In chemistry FCA was mostly used for the analysis of structure-activity relationships of chemical compounds and in particular for identifying molecular subgraphs which have a certain activity profile. The final research area which we surveyed where FCA had a significant number of applications is ontology engineering, ontology merging and ontology quality management. We surveyed several mature applications of FCA for constructing ontologies in different domains. Recently authors who are active in this research domain gained a significant interest in Fuzzy FCA for ontology construction and rough FCA for ontology mapping. In this survey, we chose to subdivide the publications on FCA in some very coarse grained research domains. We are conscious that other subdivisions were possible however we believe that our current approach gives an adequate overview of the development of FCA applications over the past 10 years. There were also many other applications on which we didn't zoom in that much. They include applications on analyzing data in the context of ship building, student learning curves, etc. We decided to mention several of them in a final concluding section.

In the future, we will host the references and links to the articles on a public interface and hope that this compendium may serve to guide both practitioners and researchers to new and improved avenues for FCA.

Acknowledgements

We are grateful to Paul Elzinga for his help in composing the concept lattices.

References

Ammons, G., Mandelin, D., Bodík, R., & Larus, J. R. (2003). Debugging temporal specifications with concept analysis. In *PLDI 2003* (pp. 182–195).

Bao, S., Zhou, Y., & He, S. (2005). Research on pressure component design ontology building based on knowledge sharing and reusing. In *Proc. 9th int. conf. on computer supported cooperative work in design*.

Barbut, M., & Monjardet, B. (1970). *Ordre et classification, algèbre et combinatoire*. Paris: Hachette.

Bartel, H.-G., & Brüggemann, R. (1998). Application of formal concept analysis to structure-activity relationships. *Fresenius Journal of Analytical Chemistry, 361*, 23–28.

Belohlavek, R., Sigmund, E., & Zacpal, J. (2011). Evaluation of IPAQ questionnaires supported by formal concept analysis. *Information Sciences, 181*(10), 1774–1786.

Belohlavek, R., Sklenar, V., Zacpal, J., & Sigmund, E. (2007). Evaluation of questionnaires by means of formal concept analysis. In J. Diatta, P. Eklund, & M. Liquiere (Eds.), *CLA 2007, int. conference on concept lattices and their applications* (pp. 100–111). October 24–26, Montpellier, France.

Bendaoud, R., Toussaint, Y., & Napoli, A. (2008). PACTOLE: A methodology and a system for semi-automatically enriching an ontology from a collection of texts. In P. Eklund et al. (Eds.). *ICCS LNAI* (Vol. 5113, pp. 203–216). Springer.

Berrey, L. (Ed.). (1962). *Roget's international thesaurus* (3rd ed. New York: Crowell.

Bertaux, A., Le Ber, F., Braud, A., & Tremolieres, M. (2009). Identifying ecological traits: A concrete FCA-based approach. In S. Ferré & S. Rudolph (Eds.). *ICFCA LNAI* (Vol. 5548, pp. 224–236). Springer.

Besson, J., Pensa, R. G., Robardet, C., & Boulicaut, J. -F. (2005). Constraint-based mining of fault-tolerant patterns from boolean data. In *KDID 2005* (pp. 55–71).

Besson, J., Robardet, C., Boulicaut, J.-F., & Rome, S. (2005b). Constraint-based concept mining and its application to microarray data analysis. *Intelligent Data Analysis, 9*(1), 59–82.

Beydoun, G. (2008). Using formal concept analysis towards cooperative e-learning. In D. Richards et al. (Eds.). *PKAW LNAI* (Vol. 5465, pp. 109–117). Springer.

Beydoun, G., Kultchitsky, R., & Manasseh, G. (2007). Evolving semantic web with social navigation. *Expert Systems with Applications, 32*, 265–276.

Birkhoff, G. (1973). *Lattice theory* (3rd ed.). American Mathematical Society Coll Publ. 25 Providence RI.

Blinova, V. G., Dobrynin, D. A., Finn, V. K., Kuznetsov, S. O., & Pankratova, E. S. (2003). Toxicology analysis by means of the JSM-method. *Bioinformatics, 19*(10), 1201–1207.

Boutari, A. M., Carpineto, C., & Nicolussi, R. (2010). Evaluating term concept association mesaures for short text expansion: two case studies of classification and clustering. In *CLA 2010* (pp. 163–174).

Breu, S., Zimmermann, T., & Lindig, C., (2006). Mining eclipse for cross-cutting concerns. In *MSR*. Shanghai, China. May 22–23.

Brüggemann, R., & Bartel, H. G. (1999). A theoretical concept to rank environmentally significant chemicals. *Journal of Chemical Information and Computer Sciences, 39*, 211–217.

Brüggemann, R., Voig, K., & Steinberg, C. E. W. (1997). Application of formal concept analysis to evaluate environmental databases. *Chemosphere, 35*(3), 479–486.

Bruno, M., Canfora, G., Penta, M.D., & Scognamiglio, R. (2005). An Approach to support Web Service Classification and Annotation. In *Proc. IEEE Int. Conf. on e-Technology, e-Commerce and e-Service* (pp. 138–143).

Busch, P., & Richards, D. (2004). Modeling tacit knowledge via questionnaire data. *ICFCA LNAI* (2961, pp. 321–328). Springer.

Carpineto, C., & Romano, G. (2004a). *Concept data analysis: Theory and applications*. John Wiley & Sons.

Carpineto, C., & Romano, G. (2004b). Exploiting the potential of concept lattices for information retrieval with CREDO. *Journal of Universal Computing, 10*(8), 985–1013.

Carpineto, C., & Romano, G. (2005). Using concept lattices for text retrieval and mining. In B. Ganter et al. (Eds.). *Formal concept analysis. LNAI* (Vol. 3626, pp. 161–171). Springer.

Carpineto, C., Michini, C., & Nicolussi, R. (2009). A concept lattice-based kernel for SVM text classification. In *ICFCA 2009* (pp. 237–250).

Chang, Y. H. (2007). Automatically constructing a domain ontology for document classification. In *Proc of the sixth IEEE int. conf on machine learning and cybernetics*, (pp. 1942–1947). August 19–22.

Chang, Y. H., & Huang, H. V. (2008). An automatic document classifier system based on naïve bayes classifier and ontology. In *Proc. seventh IEEE int. conf. on machine learning and cybernetics* (pp. 3144–3149). July 12–15, Kunming.

Chi, Y. L., Hsu, T. Y., & Yang, W. P. (2005). Building ontological knowledge bases for sharing knowledge in digital archive. In *Proc. Fourth IEEE Int. Conf. on Machine Learning and Cybernetics* (Vol. 4, pp. 2261–2266). August 18–21, Guangzhou.

Cho, W. C., & Richards, D. (2004). Improvement of precision and recall for information retrieval in a narrow domain: reuse of concepts by formal concept analysis. In *Proceedings of the 2004 IEEE/WIC/ACM international conference on web intelligence (WI '04)* (pp. 370–376). Washington, DC, USA: IEEE Computer Society.

Choi, V., Huang, Y., Lam, V., Potter, D., Laubenbacher, R., & Duca, K. (2008). Using formal concept analysis for microarray data comparison. *Journal of Bioinformatics and Computational Biology, 6*(1), 65–75.

Choi, V., Laubenbacher, R., Duca, K., Lam, V., & Huang, Y. (2006). Using formal concept analysis for microarray data comparison. Presented at DIMACS Workshop on Clustering Problems in Biological Networks, Piscataway, NJ, May 9–11.

Chou, C. Y., & Mei, H. (2008). Analyzing tag-based mashups with fuzzy FCA. In *Proceedings of the 2008 IEEE international symposium on service-oriented system engineering (SOSE '08)* (pp. 257–262). Washington, DC, USA: IEEE Computer Society.

Christopher, A. (1965). A city is not a tree. Architectural form *122(1)*, April, pp. 58–62 and *122(2)*, May, pp. 58–62.

Cimiano, P. (2006). *Ontology learning and population from text – algorithms, evaluation and applications*. Springer.

Cimiano, P., Hotho, A., Stumme, G., & Tane, J. (2004). Conceptual knowledge processing with formal concept analysis and ontologies. In P. Eklund (Ed.). *ICFCA, LNAI* (2961, pp. 189–207). Springer.

Cimiano, P., Hotho, A., & Staab, S. (2005). Learning concept hierarchies from text corpora using formal concept analysis. *Journal of Artificial Intelligence (JAIR), 24*, 305–339.

Cole, R., & Becker, P. (2005). Navigation spaces for the conceptual analysis of software structure. In B. Ganter et al. (Eds.). *ICFCA LNAI* (3403, pp. 113–128). Springer.

Cole, R., & Eklund, P. (1996) Application of formal concept analysis to information retrieval using a hierarchically structured thesaurus. In *International conference on conceptual graphs, ICCS '96* (pp. 1–12). University of New South Wales, Sydney.

Cole, R., & Eklund, P. (2001). Browsing semi-structured web texts using formal concept analysis. In H. Delugach et al. (Eds.). *Conceptual structures: Broadening the base LNAI* (2120, pp. 319–332). Berlin: Springer.

Cole, R., Eklund, P., & Stumme, G. (2003). Document retrieval for e-mail search and discovery using formal concept analysis. *Applied Artificial Intelligence, 17*, 257–280.

Cole, R., Tilley, T., & Ducrou, J., (2005). Conceptual exploration of software structure. A collection of examples. In Belohlavek, R. et al. (Eds.), *CLA* (pp. 135–148).

Cole, R., & Stumme, G. (2000). CEM – A conceptual email manager. In *ICCS 2000* (pp. 438-452).

Cole, R. J. (2000). The management and visualization of document collections using formal concept analysis. Ph. D. Thesis, Griffith University.

Colton, S., & Wagner, D. (2007). Using formal concept analysis in mathematical discovery. In M. Kauers et al. (Eds.). *MKM/calculemus. LNAI* (4573, pp. 205–220). Springer.

Curé, O., & Jeansoulin, R. (2008). An FCA-based solution for ontology mediation. *Proceedings of the 2nd international workshop on Ontologies and information systems for the semantic web (ONISW '08)* (pp. 39–46). New York, NY, USA: ACM.

Dau, F., Ducrou, J., & Eklund, P. (2008). Concept similarity and related categories in searchsleuth. In P. Eklund et al. (Eds.). *ICCS. LNAI* (5113, pp. 255–268). Springer.

Dau, F., & Knechtel, M. (2009). Access policy design supported by FCA methods. In F. Dau et al. (Eds.). *ICCS LNAI* (5662, pp. 141–154). Springer.

De Souza, K. X. S., & Davis, J. (2004a). Aligning ontologies and evaluating concept similarities. In R. Meersman et al. (Eds.). *CoopIS/DOA/ODBASE LNCS* (3291, pp. 1012–1029). Springer.

De Souza, K. X. S., & Davis, J. (2004b). Using an aligned ontology to process user queries. In C. Bussler et al. (Eds.). *AIMSA. LNAI* (3192, pp. 44–53). Springer.

De Souza, K. X. S., Davis, J., & Evangelista, S. R. M. (2006). Aligning ontologies evaluating concept similarities and visualizing results. *J. on data semantics V, LNCS* (3870, pp. 211–236). Springer.

Del Grosso, C., Penta, M. D., & Guzman, I. G. R. (2007). An approach for mining services in database oriented applications. In *Proceedings of the 11th European conference on software maintenance and reengineering (CSMR '07)* (pp. 287–296). Washington, DC, USA: IEEE Computer Society.

Ducrou, J., & Eklund, P. W. (2007). SearchSleuth: The conceptual neighbourhood of an web query. In J. Diatta, P. Eklund, & M. Liquière (Eds.), *Proc. CLA 2007*, LIRMM & University of Montpellier II.

Dufour-Lussier, V., Lieber, J., Nauer, E., & Toussaint, Y. (2010). Text adaptation using formal concept analysis. *Case-based reasoning research and development (ICCBR 2010), lecture notes in artificial intelligence* (Vol. 6176, pp. 96–110). Alessandria, Italy: Springer. July.

Ebner, M., Mühlburger, H., Schaffert, S., Schiefner, M., Reinhardt, W., & Wheeler, S. (2010). Getting granular on twitter: Tweets from a conference and their limited usefulness for non-participants. *Advances in Information and Communication Technology, 324*, 102–113.

Egho, E., Jay, N., Raissi, C., & Napoli, A. (2011). A FCA-based analysis of sequential care trajectories. In Amedeo Napoli, Vilem Vychodil (Eds.), *CLA 2011* (pp. 363–376). INRIA Nancy – Grand Est and LORIA.

Eisenbarth, T., Koschke, R., & Simon, D. (2003). Locating features in source code. *IEEE Transactions on Software Engineering, 29*(3).

Eklund, P., Ducrou, J., & Brawn, P. (2004). Concept lattices for information visualization: Can novices read line-diagrams? In P. Eklund (Ed.). *ICFCA, LNAI* (2961, pp. 57–73). Springer.

Elzinga, P., Poelmans, J., Viaene, S., Dedene, G., & Morsing, S. (2010). Terrorist threat assessment with formal concept analysis. In: *Proc. IEEE International Conference on Intelligence and Security Informatics* (pp. 77–82).

Elzinga, P., Wolff, K. E., Poelmans, J., Viaene, S., & Dedene, G. (2012). Analyzing chat conversations of pedosexuals with temporal relational semantic systems. Accepted for Contributions to ICFCA 2012.

Endres, D., Adam, R., Giese, M., & Noppeney, U. (2012). Understanding the semantic structure of human fMRI brain recordings with formal concept analysis. Accepted for LNAI proceedings ICFCA 2012.

Falk, I., & Gardent, C. (2011). Combining formal concept analysis and translation to assign frames and thematic grids to french verbs. In Amedeo Napoli, Vilem Vychodil (Eds.), *CLA 2011* (pp. 223–238). INRIA Nancy Grand Est and LORIA.

Falk, I., Gardent, C., & Lorenzo, A. (2010). Using formal concept analysis to acquire knowledge about verbs. In *CLA 2010* (pp. 151–162).

Fan, L., & Xiao, T. (2007). An automatic method for ontology mapping. In B. Apolloni et al. (Eds.). *KES/WIRN, Part III, LNAI* (4694, pp. 661–669). Springer.

Fang, K., Chang, C., & Chi, Y. (2007). Using formal concept analysis to leverage ontology-based acu-point knowledge system. In D. Zhang (Ed.). *ICMB, LNCS* (4901, pp. 115–121). Springer.

Fellbaum, Ch. (1990). English verbs as a semantic net. *International Journal of Lexicography, 3*, 278–301.

Fenza, G., Loia, V., & Senatore, S. (2008). Concept mining of semantic web services by means of extended fuzzy formal concept analysis (FFCA). In *IEEE int. conf. on systems, man and cybernetics* (pp. 240–245).

Fenza, G., & Senatore, S. (2009). Friendly web services selection exploiting fuzzy formal concept analysis. *Soft Computing, 14*(8), 811–819.

Formica, A. (2006). Ontology-based concept similarity in formal concept analysis. *Information Sciences, 176*, 2624–2641.

Formica, A. (2008). Concept similarity in formal concept analysis: an information content approach. *Knowledge-Based Systems, 21*, 80–87.

Gamallo, P., Lopes, G. P., & Agustini, A. (2007). Inducing classes of terms from text. In V. Matousek et al. (Eds.). *TSD, LNAI* (4629, pp. 31–38). Springer.

Ganter, B., & Wille, R. (1989). Conceptual scaling. In F. Roberts (Ed.), *Applications of combinatorics and graph theory to the biological and social sciences* (pp. 139–167). Berlin: Springer.

Ganter, B., Grigoriev, P. A., Kuznetsov, S. O., & Samokhin, M. V. (2004). Concept-based data mining with scaled labeled graphs. In K. E. Wolff et al. (Eds.). *ICCS, LNAI* (3127, pp. 94–108). Springer.

Ganter, B., & Kuznetsov, S. O. (2001). Pattern structures and their projections. In G. Stumme, & H. Delugach (Eds.), *Proc. 9th int. conf. on conceptual structures, ICCS'01, lecture notes in artificial intelligence* (vol. 2120, pp.129–142).

Ganter, B., & Wille, R. (1999). *Formal concept analysis, mathematical foundations*. Springer.

Gebert, J., Motameny, S., Faigle, U., Forst, C. V., & Schrader, R. (2008). Identifying genes of gene regulatory networks using formal concept analysis. *Journal of Computational Biology, 15*(2), 185–194.

Geng, L., Korba, L., Wang, Y., Wang, X., & You, Y. (2008). Finding topics in email using formal concept analysis and fuzzy membership functions. In S. Bergler (Ed.). *Canadian AI, LNAI* (5032, pp. 108–113). Springer.

Girault, T. (2008). Concept lattice mining for unsupervised named entity annotation. In R. Belohlavek & S. O. Kuznetsov (Eds.), *Proc. CLA 2008* (pp. 35–46). Olomouc: Palacký University.

Gruber, T. (2009). Ontology. In L. Liu et al. (Eds.), *Encyclopedia of database systems*. Springer.

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human Computer Studies, 43*(5–6), 907–928.

Hauff, B. M., & Deogun, J. S. (2007). Parameter tuning for disjoint clusters based on concept lattices with application to location learning. In A. An et al. (Eds.). *RSFDGrC, LNAI* (Vol. 4482, pp. 232–239). Springer.

Hesse, W., & Tilley, T. (2005). Formal concept analysis used for software analysis and modelling. In *Formal concept analysis 2005* (pp. 288–303). Sprimger.

Hotho, A., Maedche, A., & Staab, S. (2002) Ontology-based text document clustering. KI 16(4), pp. 48–54.

Hsieh, T. C., Tsai, K. H., Chen, C. L., Lee, M. C., Chiu, T. K., & Wang, T. (2007). Query-based ontology approach for semantic search. In *Proc. of the sixth int. conf. on machine learning and cybernetics*, Hong Kong, 19–22 August.

Huang, D., & Li, H. (2008b). *The research of web page recommendation model based on FCA and enterprise ontology*. *Proceedings of the 2008 IEEE Pacific-Asia Workshop on computational intelligence and industrial application – volume 01 (PACIIA '08)* (Vol. 1. Washington, DC, USA: IEEE Computer Society (pp. 232–236). Washington, DC, USA: IEEE Computer Society.

Huang, D., & Zhu, H. (2008a). *Rough formal concept analysis for semi-automatic marine ontology*. *Proceedings of the 2008 fourth international conference on natural computation – volume 05 (ICNC '08)* (Vol. 5. Washington, DC, USA: IEEE Computer Society (pp. 114–118). Washington, DC, USA: IEEE Computer Society.

Hwang, S. H., Kim, H. G., & Yang, H. S. (2005). FCA-based ontology construction for the design of class hierarchy. In O. Gervasi et al. (Eds.). *ICCSA, LNCS* (3482, pp. 827–835). Springer.

Ignatov, D. I., & Kuznetsov, S. O. (2008). Concept-based recommendations for internet advertisement. In Radim Belohlavek, & Sergei O. Kuznetsov (Eds.), *Proceedings of the sixth international conference concept lattices and their applications (CLA'08), CLA 2008* (pp. 157–166). Palacky University, Olomouc.

Ignatov, D. I., & Kuznetsov, S. O. (2009). Frequent itemset mining for clustering near duplicate web documents. In S. Rudolph, F. Dau, & S. O. Kuznetsov (Eds.). *ICCS 2009, LNCS* (Vol. 5662, pp. 185–200). Springer-Verlag Berlin Heidelberg.

Ignatov, D. I., Mamedova, S., Romashkin, N., & Shamshurin, I. (2011). What can closed sets of students and their marks say? In *EDM 2011* (pp. 223–228).

Jay, N., Kohler, F., & Napoli, A. (2008a). Using formal concept analysis for mining and interpreting patient flows within a healthcare network. In S. B. Yahia et al. (Eds.). *CLA 2006, LNAI* (Vol. 4923, pp. 263–268). Springer.

Jay, N., Kohler, F., & Napoli, A. (2008b). Analysis of social communities with iceberg and stability-based concept lattices. In R. Medina et al. (Eds.). *ICFCA, LNAI* (Vol. 4933, pp. 258–272). Springer.

Jiang, G., Ogasawara, K., Endoh, A., & Sakurai, T. (2003). Context-based ontology building support in clinical domains using formal concept analysis. *International Journal of Medical Informatics Journal, 71*, 71–81.

Jiang, G., Pathak, J., & Chute, C. G. (2009). Formalizing ICD coding rules using formal concept analysis. *Journal of Biomedical Informatics, 42*(3), 504–517.

Jiang, G., & Chute, C. G. (2009b). Auditing the semantic completeness of SNOMED CT using formal concept analysis. *Journal of the American Medical Informatics Association, 16*(1), 89–102.

Kaytoue, M., Duplessis, S., Kuznetsov, S. O., & Napoli, A. (2009). Two FCA-based methods for mining gene expression data. In S. Ferré et al. (Eds.). *ICFCA, LNAI* (Vol. 5548, pp. 251–266). Springer.

Kaytoue, M., Kuznetsov, S. O., Napoli, A., & Duplessis, S. (2011). Mining gene expression data with pattern structures in formal concept analysis. *Information Sciences, 181*(10), 1989–2001.

Keller, B. J., Eichinger, F., & Kretzler, M. (2012). Formal concept analysis of disease similarity. *AMIA Summits on Translational Science Proceedings, 2012*, 42.

Kim, D. S., Hwang, S. H., & Kim, H. G. (2007a). *Concept analysis of OWL ontology based on the context family model*. *Proceedings of the 2007 international conference on convergence information technology (ICCIT '07)*. Washington, DC, USA: IEE Computer Society.

Kim, H. L., Hwang, S. H., & Kim, H. G. (2007b). FCA-based approach for mining contextualized folksonomy. In *SAC*. March 11–15, Seoul, Korea.

Kim, M., & Compton, P. (2004). Evolutionary document management and retrieval for specialised domains on the web. *International Journal of Human Computer Studies, 60*(2), 201–241.

Kim, M., & Compton, P. (2006). A hybrid browsing mechanism using conceptual scales. In A. Hoffmann et al. (Eds.). *PKAW, LNAI* (Vol. 4303, pp. 132–143). Springer.

Kirchberg, M., Leonardi, E., Tan, Y. S., Link, S., Ko, R., & Lee, B. S. (2012). Formal concept analysis for semantic web data. Accepted for *LNAI proceedings ICFCA 2012*.

Kiu, C. C., & Lee, C. S. (2008). Ontological knowledge management through hybrid unsupervised clustering techniques. In Y. Zhang et al. (Eds.). *APWeb, LNCS* (Vol. 4976, pp. 499–510). Springer.

Koester, B. (2005). Conceptual knowledge processing with google. In Bauer, Kröner, Brandherm (Ed.), *LWA 2005 – lernen, wissensentdeckung und adaptivität*. German Research Center for Artificial Intelligence (DFKI), Saarbrücken.

Koester, B. (2006). Conceptual knowledge retrieval with fooCA: Improving web search engine results with contexts and concept hierarchies. In P. Perner (Ed.). *ICDM, LNAI* (Vol. 4065, pp. 176–190.). Springer.

Krohn, U., Davies, N. J., & Weeks, R. (1999). Concept lattices for knowledge management. *BT Technology Journal, 17*(4), 108–116.

Krone, M., & Snelting, G. (1994). On the inference of configuration structures from source code. In *Proceedings of the 16th international conference on software engineering (ICSE '94)* (pp. 49–57). Los Alamitos, CA, USA: IEEE Computer Society Press.

Krotzsch, M., Hitzler, P., & Zhang, G. Q. (2005). Morphisms in context. In F. Dau et al. (Eds.). *ICCS, LNAI* (Vol. 3596, pp. 223–237). Springer.

Kuznetsov, S. O., & Samokhin, M. V. (2005). Learning closed sets of labeled graphs for chemical applications. In *Inductive Logic Programming* (pp. 190–208). Berlin, Heidelberg:Springer.

Kuznetsov, S., Obiedkov, S., & Roth, C. (2007a). Reducing the representation complexity of lattice-based taxonomies. In U. Priss et al. (Eds.). *ICCS. LNCS (LNAI)* (Vol. 4604, pp. 241–254). Springer.

Kuznetsov, S. O. (2007). On stability of a formal concept. *Annals of Mathematics and Artificial Intelligence, 49*, 101–115.

Lim, W. C., & Lee, C. S. (2005). Knowledge discovery through composited visualization, navigation and retrieval. In A. Hoffmann et al. (Eds.). *DS, LNAI* (Vol. 3735, pp. 377–379). Springer.

Lounkine, E., Auer, J., & Bajorath, J. (2008). Formal concept analysis for the identification of molecular fragment combinations specific for active and highly potent compounds. *Journal of Medicinal Chemistry, 51*, 5342–5348.

Lu, Y., & Zhang, G. (2008). A DL-based approach for eliminating concept restriction conflicts in ontology merging. In *Proceedings of the 2008 IFIP international conference on network and parallel computing (NPC '08)* (pp. 446–451). IEEE Computer Society.

Mackensen, K., & Wille, U. (1999). Qualitative text analysis supported by conceptual data systems. *Quality and Quantity: International Journal of Methodology, 2*(33), 135–156.

Maille, N., Statler, I., & Chaudron, L. (2005). An application of FCA to the analysis of aeronautical incidents. In B. Ganter et al. (Eds.). *ICFCA, LNAI* (Vol. 3403, pp. 145–161). Springer.

Maio, C. D., Fenza, G., Gaeta, M., Loia, V., Orciuoli, F., & Senatore, S. (2012). RSS-based e-learning recommendations exploiting fuzzy FCA for knowledge modeling. *Applied Soft Computing, 12*(1), 1 13–124.

Maio, C. D., Fenza, G., Loia, V., & Senatore, S. (2009). Ontology-based knowledge structuring: an application on RSS feeds. In *Proceedings of the 2nd conference on human system interactions (HSI'09)* (pp. 464–470). Piscataway, NJ, USA: IEEE Press.

Mens, K., & Tourwé, T. (2005). Delving source code with formal concept analysis. *Computer Languages, Systems & Stuctures, 31*, 183–197.

Messai, N., Bouaud, J., Aufaure, M. A., Zelek, L., & Séroussi, B. (2011). Using formal concept analysis to discover patterns of non-compliance with clinical practice guidelines: a case study in the management of breast cancer. In *Artificial intelligence in medicine* (pp. 119–128). Heidelberg: Springer Berlin.

Missaoui, R., & Kwuida, L. (2009). What can formal concept analysis do for data warehouses? In S. Ferré et al. (Eds.). *ICFCA. LNAI* (Vol. 5548, pp. 58–65). Springer.

Molloy, I., Chen, H., Li, T., Wang, Q., Li, N., Bertino, E., Calo, S., & Lobo, J. (2008). Mining roles with semantic meanings. June 11–13, Estes Park, Colorado, USA.

Motameny, S., Versmold, B., & Schmutzler, R. (2008). Formal concept analysis for the identification of combinatorial biomarkers in breast cancer. In: *ICFCA 2008* (pp. 229–240).

Myat, N. N., & Hla, K. H. S. (2005). A combined approach of formal concept analysis and text mining for concept based document clustering. In: *Proc. IEEE/WIC/ACM int. conf. on web intelligence*.

Nazri, M. Z. A., Shamsudin, S. M., & Bakar, A. A. (2008). An exploratory study on malay processing tool for acquisition of taxonomy using FCA. *Proceedings of the 2008 eighth international conference on intelligent systems design and applications – Volume 01 (ISDA '08)* (Vol. 1, pp. 375–380). Washington, DC, USA: IEEE Computer Society.

Okubo, Y., & Haraguchi, M. (2006). Finding conceptual document clusters with improved top-N formal concept search. In *Proceedings of the 2006 IEEE/WIC/ACM international conference on web intelligence (WI '06)* (pp. 347–351). Washington, DC, USA: IEEE Computer Society.

Old, L. J. (2006). Homograph disambiguation using formal concept analysis. In *ICFCA 2006* (pp. 221–232).

Pan, T., & Fang, K. (2009). Ontology-Based Formal Concept Differences Analysis in Radiology Report Impact by the Adoption of PACS. In S. Ferré et al. (Eds.). *ICFCA, LNAI* (Vol. 5548, pp. 267–283). Springer.

Pensa, R. G., Besson,, J., & Boulicaut, J.-F. (2004a). A methodology for biologically relevant pattern discovery from gene expression data. In *Discovery science* (pp. 230–241). Springer.

Pensa, R. G., Besson, J., Robardet, C., & Boulicaut, J.-F. (2004b). Contribution to gene expression data analysis by means of set pattern mining. *Constraint-based mining and inductive databases 2004*, 328–347.

Poelmans, J., Dedene, G., Verheyden, G., Van der Mussele, H., Viaene, S., & Peters, E. (2010). Combining business process and data discovery techniques for analyzing and improving integrated care pathways. *Advances in data mining. applications and theoretical aspects, 10th industrial conference (ICDM), lecture* *notes in computer science, Leipzig, Germany* (6171, pp. 505–517). Springer. July 12–14.

Poelmans, J., Elzinga, P., Viaene, S., & Dedene, G. (2008). An exploration into the power of formal concept analysis for domestic violence analysis. *Advances in data mining. applications and theoretical aspects, 8th industrial conference (ICDM), Lecture notes in computer science, Leipzig, Germany* (Vol. 5077, pp. 404–416). Springer. July 16–18.

Poelmans, J., Elzinga, P., Viaene, S., & Dedene, G. (2009). A case of using formal concept analysis in combination with emergent self organizing maps for detecting domestic violence. *9th industrial conf. on data mining, LNCS* (Vol. 5633, pp. 247–260). Leipzig, Germany: Springer. July 20–22.

Poelmans, J., Elzinga, P., Viaene, S., & Dedene, G. (2010a). 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. http://dx.doi.org/10.1002/isaf.319.

Poelmans, J., Elzinga, P., Viaene, S., & Dedene, G. (2010b). Formal concept analysis in knowledge discovery: A survey. *18th International conference on conceptual structures (ICCS): From information to intelligence, lecture notes in computer science* (6208, pp. 139–153). Kuching, Sarawak, Malaysia: Springer. July 26–30.

Poelmans, J., Elzinga, P., Viaene, S., & Dedene, G. (2010c). A method based on temporal concept analysis for detecting and profiling human trafficking suspects. In *Proc. IASTED international conference on artificial intelligence* (pp. 15–17).

Poelmans, J., Elzinga, P., Viaene, S., & Dedene, G. (2011a). Formally analyzing the concepts of domestic violence. *Expert Systems with Applications* (4), 3116–3130. http://dx.doi.org/10.1016/j.eswa. 2010.08.103.

Poelmans, J., Elzinga, P., Viaene, S., Dedene, G., & Kuznetsov, S. (2011b). A concept discovery approach for fighting human trafficking and forced prostitution. *19th International conference on conceptual structures, lecture notes in computer science* (Vol. 6828, pp. 201–214). Derby, England: Springer. July 25–29.

Poelmans, J., Ignatov, D. I., Viaene, S., Dedene, G., & Kuznetsov, S. O. (2012a). Text mining scientific papers: a survey on FCA-based information retrieval research. In: *Advances in data mining. Applications and theoretical aspects* (pp. 273–287). Heidelberg: Springer Berlin.

Poelmans, J., Elzinga, P., Ignatov, D. I., & Kuznetsov, S. O. (2012b). Semi-automated knowledge discovery: identifying and profiling human trafficking. *International Journal of General Systems, 41*(8), 774–804.

Poelmans, J., Kuznetsov, S. O., Ignatov, D. I., & Dedene, G. (2013). Formal Concept Analysis in Knowledge Processing: a Survey on Models and Techniques. *Expert Systems with Applications*.

Potter, D. P. (2005). A combinatorial approach to scientific exploration of gene expression data: An integrative method using formal concept analysis for the comparative analysis of microarray data. Thesis dissertation, Department of Mathematics, Virginia Tech.

Priss, U. (1998). Relational concept analysis: semantic structures in dictionaries and lexical databases. PhD Thesis, Verlag Shaker, Aachen.

Priss, U. (2004). Modeling lexical databases with formal concept analysis. *Journal of Universal Computer Science, 10*(8), 967–984.

Priss, U., & Old, J. L. (2006b). *An application of relation algebra to lexical databases*. Springer. ICCS 2006, pp. 388–400.

Priss, U., & Old, L. J. (2005). Conceptual exploration of semantic mirrors. In B. Ganter et al. (Eds.). *ICFCA, LNAI* (Vol. 3403, pp. 21–32). Springer.

Priss, U., & Old, L. J. (2010). Concept neighbourhoods in lexical databases. In *Formal concept analysis, lecture notes in computer science* (Vol. 5986, pp 283–295).

Priss, U., Riegler, P., & Jensen, N. (2012). Using FCA for modelling conceptual difficulties in learning processes. Accepted for *Contributions to ICFCA 2012*.

Qu, L., & Liu, D. (2007). Extending dynamic aspect mining using formal concept analysis. *Proceedings of the fourth international conference on fuzzy systems and knowledge discovery – Volume 03 (FSKD '07)* (Vol. 3, pp. 564–567). Washington, DC, USA: IEEE Computer Society.

Quan, T. T., Hui, S. C., & Fong, A. C. M. (2006a). Automatic fuzzy ontology generation for semantic help-desk support. *IEEE Transactions on Industrial Informatics, 2*(3), 155–164.

Quan, T. T., Hui, S. C., Fong, A. C. M., & Cao, T. H. (2004b). Automatic generation of ontology for scholarly semantic web. In S. A. McIlraith et al. (Eds.). *ISWC, LNCS* (Vol. 3298, pp. 726–740). Springer.

Quan, T. T., Hui, S. C., Fong, A. C. M., & Cao, T. H. (2006b). Automatic fuzzy ontology generation for semantic web. *IEEE Transactions on Knowledge and Data Engineering, 18*(6), 842–856.

Resnik, P. (1995). Using information content to evaluate semantic similarity in a taxonomy. In *Proc. int. joint conf. on artificial intelligence* (pp. 448–453). Montreal, Quebec, Canada, August 20–25, Morgan Kaufmann.

Richards, D. (2004). Addressing the ontology acquisition bottleneck through reverse ontological engineering. *Knowledge and Information Systems, 6*, 402–427.

Richards, D. (2006). Ad-hoc and personal ontologies: A prototyping approach to ontology engineering. In A. Hoffmann et al. (Eds.). *PKAW, LNAI* (Vol. 4303, pp. 13–24). Springer.

Richards, D., & Compton, P. (1997). *Combining formal concept analysis and ripple down rules to support reuse, software engineering knowledge engineering SEKE'97*. Springer Verlag.

Richards, D., & Malik, U. (2003a). Multi level knowledge discovery from rule bases. *Applied Artificial Intelligence, 17*, 181–205.

Richards, D., & Malik, U. (2003b). Mining propositional knowledge bases to discover multi-level rules. In O. R. Zaïane et al. (Eds.). *Mining multimedia and complex data, LNAI* (Vol. 2797, pp. 199–216). Springer.

Romashkin, N., Ignatov, D. I., & Kolotova, E. (2011). How university entrants are choosing their department? mining of university admission process with FCA taxonomies. In *EDM 2011* (pp. 229–234).

Roth, C., Obiedkov, S., & Kourie, D. (2008a). Towards concise representation for taxonomies of epistemic communities. In S. B. Yahia et al. (Eds.). *CLA 2006, LNAI* (Vol. 4923, pp. 240–255). Springer.

Roth, C., Obiedkov, S., & Kourie, D. G. (2008b). On succint representation of knowledge community taxonomies with formal concept analysis. *International Journal of Foundations of Computer Science, 19*(2), 383–404.

Rouane-Hacene, M., Toussaint, Y., & Valtchev, P. (2009). Mining safety signals in spontaneous reports database using concept analysis. In *Artificial intelligence in medicine* (pp. 285–294). Berlin Heidelberg: Springer.

Rudolph, S. (2004). Exploring relational structures via FLE. In K. E. Wolff et al. (Eds.). *ICCS, LNAI* (Vol. 3127, pp. 196–216). Springer.

Rudolph, S. (2008). Acquiring generalized domain-range restrictions. In R. Medina et al. (Eds.). *ICFCA, LNAI* (Vol. 4933, pp. 32–45). Springer.

Rudolph, S., Völker, J., & Hitzler, P. (2007). Supporting lexical ontology learning by relational exploration. In U. Priss et al. (Eds.). *ICCS, LNAI* (Vol. 4604, pp. 488–491). Springer.

Sato, K., Okubo, Y., Haraguchi, M., & Kunifuji, S. (2007). Data mining of time-series medical data by formal concept analysis. In Apolloni et al. (Eds.). *KES/WIRN, Part II, LNAI* (Vol. 4693, pp. 1214–1221). Springer.

Schnabel, M. (2002). Representing and processing medical knowledge using formal concept analysis. *Methods of Information in Medicine, 41*, 33–48.

Sertkaya, B. (2009). OntoComP: A protégé plugin for completing OWL ontologies. In L. Aroyo et al. (Eds.). *ESWC, LNCS* (Vol. 5554, pp. 898–902). Springer.

Siff, M., & Reps, T. W. (1997). Identifying modules via concept analysis. In *ICSM 1997* (pp. 170–179).

Sigmund, E., Zacpal, J., Sklenar, V., & Fromel, K. (2005). New approach to evaluation of IPAQ questionnaire physical aktivity data: formal concept analysis. In *Tudory, 4th international conference movement and health* (p. 173). Palacky University.

Sklenar, V., Zacpal, J., & Sigmund, E. (2005). Evaluation of IPAQ questionnaire by FCA. In R. Bělohlávek, & V. Snášel (Eds.), *Proc. CLA 2005*. Palacký University in Olomouc, VŠB – Technical University of Ostrava.

Snelting, G. (2005). Concept lattices in software analysis. In *Formal concept analysis 2005* (pp. 272–287). Springer.

Solesvik, M. Z., & Encheva, S. (2009). Partner selection for interfirm collaboration: The context of ship design. In Y. Luo (Ed.). *CDVE, LNCS* (Vol. 5738, pp. 141–144). Springer.

Soon, K., & Kuhn, W. (2004). Formalizing user actions for ontologies. In M. J. Egenhofer et al. (Eds.). *GIScience, LNCS* (Vol. 3234, pp. 299–312). Springer.

Stumme, G., & Maedche, A. (2001). FCA-MERGE: Bottom-up merging of ontologies. In *IJCAI 2001* (pp. 225–234).

Stumpfe, D., Lounkine, E., & Bajorath, J. (2011). Molecular test systems for computational selectivity studies and systematic analysis of compound selectivity profiles. In *Chemoinformatics and computational chemical biology* (pp. 503–515). Humana Press.

Su, Y., Zhou, X. W., & Zhang, M. Q. (2008). Approach on aspect-oriented software reverse engineering at requirements level. *Proceedings of the 2008 international conference on computer science and software engineering – Volume 02 (CSSE '08)* (Vol. 2, pp. 321–324). Washington, DC, USA: IEEE Computer Society,.

Tilley, T., & Eklund, P. (2007). Citation analysis using formal concept analysis: A case study in software engineering. In *18th Int. conf. on database and, expert systems applications (DEXA)*.

Tonella, P., & Ceccato, M. (2004). Aspect mining through the formal concept analysis of execution traces. In *Proceedings of the 11th working conference on reverse engineering (WCRE '04)* (pp. 112–121). Washington, DC, USA: IEEE Computer Society.

Villerd, J., Toussaint, Y., & Lillo-Le Louët, A. (2010). Adverse drug reaction mining in pharmacovigilance data using formal concept analysis. In *Machine learning and knowledge discovery in databases* (pp. 386–401). Berlin Heidelberg: Springer.

Völker, J., & Rudolph, S. (2008b). Lexico-logical acquisition of OWL DL axioms: An integrated approach to ontology refinement. In R. Medina et al. (Eds.). *ICFCA, LNAI* (Vol. 4933, pp. 62–77). Springer.

Wang, Y., Du, Y., & Chen, S. (2009). The understanding between two agent crawlers based on domain ontology. *Proceedings of the 2009 international conference on computational intelligence and natural computing – Volume 01 (CINC '09)* (Vol. 1, pp. 47–50). Washington, DC, USA: IEEE Computer Society.

Wermelinger, M., Yu, Y., & Strohmaier, M. (2009). Using formal concept analysis to construct and visualize hierarchies of socio-technical relations. In *Proc. ICSE*. May 16–24, Vancouver, Canada.

Wollbold, J., Guthke, R., & Ganter, B. (2008). Constructing a knowledge base for gene regulatory dynamics by formal concept analysis methods. In K. Horimoto et al. (Eds.). *AB, LNCS* (Vol. 5147, pp. 230–244). Springer.

Wollbold, J., Huber, R., Pohlers, D., Koczan, D., Guthke, R., Kinne, R., et al. (2009). Adapted Boolean network models for extracellular matrix formation. *BMC systems biology, 3*(1), 77.

Wollbold, J., Huber, R., Kinne, R., & Wolff, K. E. (2011). Conceptual representation of gene expression processes. In *Knowledge Processing and Data Analysis* (pp. 79–100). Springer:Berlin, Heidelberg.

Xu, H., & Xiao, D. (2009). Building information specification ontology for computer network management based on formal concept analysis. In *Proc. IEEE int. conf. on information and automation* (pp. 312–317). June 22–25, Zhuhai/Macau, China.

Xu, W., Li, W., Wu, M., Li, W., & Yuan, C. (2006). Deriving event relevance from the ontology constructed with formal concept analysis. In A. Gelbukh (Ed.). *CICLing, LNCS* (3878, pp. 480–489). Springer.

Yang, Y., Du, Y., Sun, J., & Hai, Y. (2008). A topic-specific web crawler with concept similarity context graph based on FCA. In Huang et al. (Eds.). *ICIC. LNAI* (5227, pp. 840–847). Springer.

Zhang, L., Pei, Z., & Chen, H. (2007b). Extracting fuzzy linguistic summaries based on including degree theory and FCA. In P. Melin et al. (Eds.). *IFSA, LNAI* (4529, pp. 273–283). Springer.

Zhao, Y., Halang, W. A., & Wang, X. (2007). Rough ontology mapping in e-business integration. *Studies in Computational Intelligence, 37*, 75–93.

Zhou, B., Hui, S. C., & Chang, K. (2005). A formal concept analysis approach for web usage mining. *Intelligent Information Processing II, 163*, 437–441.

Zhou, B., Hui, S. C., & Fong, A. C. M. (2006). An effective approach for periodic web personalization. In *Proceedings of the 2006 IEEE/WIC/ACM international conference on web intelligence (WI '06)* (pp. 284–292). Washington, DC, USA: IEEE Computer Society.

Zhou, W., Liu, Z. T., & Zhao, Y. (2007a). Ontology learning by clustering based on fuzzy formal concept analysis. *Proceedings of the 31st annual international computer software and applications conference – Volume 01 (COMPSAC '07)* (Vol. 1, pp. 204–210). Washington, DC, USA: IEEE Computer Society.