6th International Conference on Applied Informatics
Eger, Hungary, January 27–31, 2004.

Towards Knowledge Management using Galois Lattices*

László Szathmáry[a,b], Amedeo Napoli[a]

[a]LORIA/Inria-Lorraine, 615 rue du Jardin Botanique, BP 101
F-54600, Vandœuvre-lès-Nancy, France
{szathmar, napoli}@loria.fr

[b]University of Debrecen, Institute of Mathematics and Informatics
Debrecen H-4010, P.O. Box 12, Hungary
szathml@delfin.unideb.hu

Abstract

In this paper we investigate the application of Galois (or concept) lattices on different information or data sources (e.g. web documents or bibliographical items) in order to organise knowledge that can be extracted from the data. This knowledge organisation can then be used for a number of purposes, e.g. knowledge management in an organisation, document retrieval on the Web, etc. Galois lattices can be considered as classification tools for knowledge units in concept hierarchies that can be used within a knowledge-based system. Moreover, Galois lattices can be used in parallel with domain ontologies for building more precise and more concise concept ontologies, and for guiding the knowledge discovery process.

Keywords: knowledge management, Galois lattices, concept lattices, ontology, knowledge discovery

1. Introduction

Today knowledge management has become an important task in an enterprise. "Knowledge is power", and realising this more and more companies start their own knowledge management (KM) projects. We have made experiments on KM within our research team, because it can also be considered a small enterprise [1]. Experience shows that researchers very often do not know exactly what other people

*This publication was made in the frame of Hungarian - French scientific and technological cooperation supported by Hungarian OM and French EGIDE (Balaton F-23/03)

[1]We will use the term "enterprise" to denote a set of people working together and having a common goal, e.g. machine design, sales, research, ...

are working on, not even within the same team. With our work we want to analyse the global work of the team to find interconnections between the members, to know which are the main/marginal works in the team, i.e. to carry out a diagnosis on the research work.

In a research team, the interests of a person can be described the best way by his publications. This is why we have chosen to analyse the bibliographical items. We have worked with the BibTeX descriptions that provide us metadata about a paper, e.g. title, authors, keywords, abstract, year of publication, etc. A BibTeX entry is similar to Dublin Core. The Dublin Core metadata standard is a simple yet effective element set for describing a wide range of networked resources, especially HTML pages [Dub01]. Metadata is often defined as "data about data". The Dublin Core is a set of fifteen metadata elements that were set down after extensive discussions. BibTeX descriptions are a standard for scientific papers, and having a "controlled vocabulary", which is a limited set of consistently used and carefully defined terms, they can be interpreted in terms of the Dublin Core.

To analyse the publications we have used classification as data mining technique in the knowledge discovery process. For classification we have used Galois lattices connected with domain ontologies. In general, ontologies provide a shared and common understanding of a domain. In our case, we have built some ontologies to explain our knowledge about the members and publications of the team. With ontologies a more intelligent way of knowledge management and document search can be performed, based on the content (semantics) of the manipulated document.

In this paper we propose a method to join Galois lattices and ontologies for guiding the knowledge discovery process between people, documents, and topics. According to our knowledge our approach of using domain ontologies for data cleaning and for data mining connected with concept lattices is unique.

The rest of the paper is organized as follows. In Section 2 we give an overview of ontologies and we describe our domain ontology. Section 3 defines knowledge discovery, presents the role of ontologies in the KDD process, and gives an introduction to the basics of concept lattices. In Section 4 we discuss our new approach of data mining with Galois lattices and ontologies. In this section we give three examples with our bibliographical items. Section 5 gives a synthesis, and Section 6 provides some perspectives for the future research.

2. Ontologies

Ontologies were developed in artificial intelligence to facilitate knowledge sharing and reuse [McG02]. An ontology is a shared and common understanding of some domain that can be communicated among people and application systems. An ontology is a formal, explicit "specification of a shared conceptualization" [Gru93]. *Ontologies represent knowledge about domains.* They 1) identify the key concepts in a domain, 2) identify key relations between these concepts, and 3) identify a vocabulary for the concepts and relations. They specify the meaning of the vocabulary terms precisely enough so that they can be shared between a) humans and humans, b) humans and machines, and c) machines and machines.

To explain and access knowledge about our team we need some ontologies. There are several simple ontologies freely available on the Web, like the DAML Ontology Library[2] or the DMOZ [3] (Directory Mozilla) effort to generate large simple ontologies. Unfortunately we have not found an appropriate ontology about scientific keywords covering our research area, thus we have decided to build one from scratch. Our initial idea was to reuse some existing ontologies by merging them with an ontology mapping method [KS03].

An extract of our ontology is illustrated in Figure 1. A node is a concept, and a directed edge represents the class–subClassOf relation. All concepts are subclasses of a root concept, called "Top". We store the keywords of our publications in the concepts. For more details see Section 4.2.

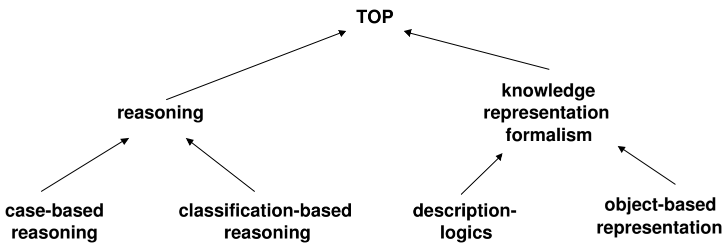

Figure 1: Ontology of keywords

3. Knowledge discovery in databases

In this section we define the terms "knowledge discovery" and "data mining". Then, we list what an ontology can be used for in the KDD process. Thereafter, we introduce the basics of the theory of Galois lattices.

3.1. Definition of KDD

Knowledge discovery and data mining are techniques to discover strategic information hidden in (very large) databases. The terms *knowledge discovery in databases* and *data mining* are often used as synonyms. Actually KDD is a process for finding valid, useful and understandable patterns in data. Data mining is just a part of this process used for the extraction of patterns from data [GG99]. Additional steps include data preparation, data selection, data cleaning, and interpretation/evaluation of the results to derive knowledge at the end [FPSS96].

[2]www.daml.org/ontologies/
[3]www.dmoz.org

3.2. Ontologies in the KDD process

Using ontologies we can do the following within the KDD process:

- Data cleaning. It allows mapping data to a single naming convention, and handling noise and errors when possible (see Section 4.2).
- Knowledge organisation (with Galois lattices). Formal concept analysis is a classification method in data mining. After having built a Galois lattice it is possible to automatically extract rules between the attributes of objects (see Section 4.3.1 and 4.3.2).
- Information retrieval. We can do query answering on our bibliographical items, for instance ranking them by their relevance on a keyword (see Section 4.3.3).
- Creating a family of lattices based on the depth of properties in the ontology. We can construct different lattices by descending in the ontology from the Top concept and taking into account more and more concepts. Comparing these lattices seems an interesting experiment, and this is at the first place of our research plans.

In this paper we examine the first three items in this list.

3.3. Galois lattices

Galois (or concept) lattices provide a natural and formal setting to discover and represent concept hierarchies. "Formal concept analysis" is mainly used for data analysis, i.e. for investigating and processing explicitly given information [GW98, Gan99, GW99]. As input data, we consider a binary relation between a set of individuals and a set of properties, e.g. a set of documents (D) and a set of keywords (W). A *context* is a triple (D, W, R), where $R \subseteq D \times W$. $R(d, w)$ means: the document d has the keyword w. We can think of the set of keywords associated with a document as a bit vector where each bit is on or off depending on whether a document has the keyword or not.

From such a binary correspondance one may derive for each document its keyword pattern (that is, the set of all keywords that apply to it). Similarly, one may derive for each keyword its document pattern (that is, the set of all documents to which it applies) [GVM93].

A *concept* (C) is determined by its extension and intension: $(Extension(C), Intension(C))$, where the extension consists of all objects that share the attributes in the intension, and vice versa, the intension consists of all attributes that are common to the objects in the extension. These concepts are *formal*, which means that they are mathematical entities and must not be identified with concepts of the mind.

The *subsumption relation*, or also called *partial ordering* ($\leq$) is defined between concepts: $C_1 \leq C_2$ (C_1 is subsumed by C_2, or C_2 subsumes C_1), iff:

$Extension(C_1) \subseteq Extension(C_2)$, and dually $Intension(C_2) \subseteq Intension(C_1)$. A *Galois lattice* is a set of concepts defined by the context (D, W, R), and organised by the subsumption relation $(\leq)$.

In the context (D, W, R) let X and $Y \subset W$, where $X \cap Y = \emptyset$. An association rule is an implication of the form $X \Rightarrow Y$. The rule $X \Rightarrow Y$ is true in the context (D, W, R) iff every document in D that contains the keywords X also contains the keywords Y. For further details on association rules see [AMS$^+$96, KMR$^+$94, PBTL99].

Example: let us consider a context (D, W, R), where $D = \{d_1, d_2, d_3\}$ and $W = \{w_1, w_2, w_3, w_4\}$. In the following table *1* means that a document contains the given keyword (see Table 1).

	w_1	w_2	w_3	w_4
d_1	1		1	
d_2	1	1		
d_3			1	1

Table 1: A simple example

The documents d_1 and d_3 are the only documents that have the keyword w_3, and vice versa, the keyword w_3 is possessed only by documents d_1 and d_3. Thus $\{d_1, d_3\} \times \{w_3\}$ is a concept of the Galois lattice. Figure 2 shows the so-called Hasse diagram of the Galois lattice. In this diagram each node is a concept and an edge represents the subsumption relation $(\leq)$. Two concepts are called *neighbors* if there is a direct edge between them in the Hasse diagram. We can also determine two association rules in this lattice: $w_2 \Rightarrow w_1$, and $w_4 \Rightarrow w_3$.

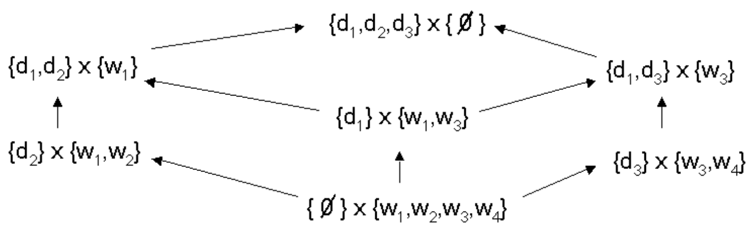

Figure 2: Hasse diagram of the lattice

We can see that a Galois lattice has two special concepts. The *Top* concept $(\{d_1, d_2, d_3\} \times \{\emptyset\})$ has an extension that contains all the objects, and the *Bottom* concept $(\{\emptyset\} \times \{w_1, w_2, w_3, w_4\})$ has an intension that contains all the attributes. To construct the lattice we have used the incremental algorithm described in [GMA95]. It builds a lattice by adding a new object to an already existing Galois lattice, without reconstructing the lattice from scratch.

Galois lattices are not used just in data mining; as a classification method they can play an important role in object oriented systems and in knowledge representation [Bog03].

4. Galois lattices connected with ontologies

4.1. General architecture of the system

In the following we describe our experiment based on the application of Galois lattices combined with an ontology for organising knowledge in a given domain. The general architecture of our system is depicted in Figure 3.

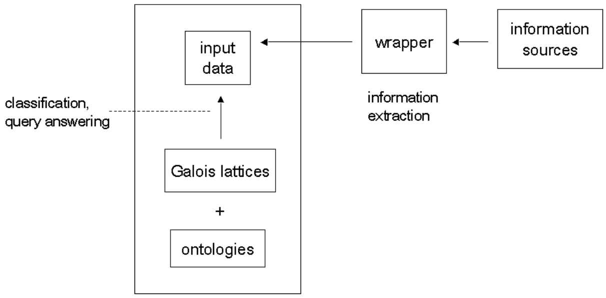

Figure 3: General architecture of the system

We have chosen to study the correlations existing between people, documents and topics within our research team. A person has written an article, possibly with other person, on a given topic defined by a list of keywords. For this experiment, we have used the bibliography of the team that is stored in a database recording the whole bibliography of the laboratory. The global bibliography archive can be accessed through a Web interface. We have designed a wrapper for extracting the BibTeX descriptions of bibliographical items of the team from the HTML pages. Then, according to our interest, we build a Galois lattice based on the relation people $\times$ articles, or articles $\times$ keywords, or people $\times$ keywords. Moreover, Galois lattices allow us to define and organize concepts from data, but also for querying the data as we explain below.

4.2. Data cleaning with ontologies

In this section, we detail why and how an ontology can be used for enhancing a keyword-based information retrieval, just as a thesaurus does. Using just a simple keyword-based search over the publications may lead to several problems:

- Based on an exact string matching, the documents that are returned must be indexed by the exact keyword used for the search. Moreover, if a syntactically incorrect keyword is attached to a document (typo) then even if a correct search term is used, the appropriate document cannot be found.

- Synonyms: keywords may have various synonyms, and, as the association of a keyword to a document is not based on definite rules or a grammar, more than one keyword may be attached to a document for the same topic.

- Languages: as we are working with a bibliography where there are at least two languages, French and English, keywords may be used in both languages. This fact must also be taken into account within the search.

Using an ontology we can solve these problems by grouping the same/similar keywords in a concept (Figure 4). Now if we perform a search on "DL" for instance, it will give a much more precise result containing documents with all these keywords.

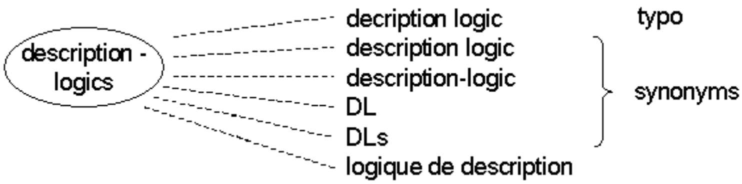

Figure 4: Grouping keywords

We used the ontology for data cleaning in the first step of the knowledge discovery process: we filtered all the keywords that were present in the publications through the ontology and as a result we got a reduced keyword set (Figure 5). For the construction of Galois lattices we took this keyword set. Grouping keywords as shown in Figure 4 can solve the three problems mentioned above. However, extending this kind of search to the Web is not really satisfactory yet because the list of keywords is not necessarily complete and documents may be missed. In this case an approximate or inexact matching may be mandatory.

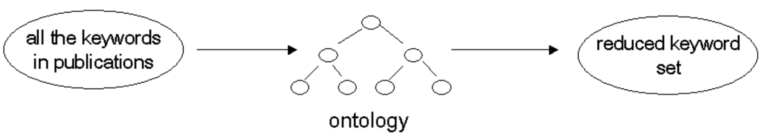

Figure 5: Filtering keywords

Until now, we have analysed 147 publications. The publications have had altogether 335 different keywords that we managed to insert in an ontology having

89 concepts, that means the reduced keyword set consists of 89 different keywords. For instance, taking the following original keywords ('DL', 'DLs', 'case-based problem solving', 'CBR', 'galois connection') a filtering through the ontology produces the following result: 'description logics', 'case-based reasoning', 'Galois lattices'.

4.3. Document organisation based on Galois lattices

In this section we show three examples of document organisation based on concept lattices. In all three cases the input data was the publication list of the team. In all cases we needed the keywords of publications, and for this we used the reduced keyword set mentioned in the previous section. As examples we pose three questions and we are looking for the answers using concept lattices.

4.3.1. First question: "Which documents are written about a common topic?"

In our team first we were interested which documents are about a common topic, as it would be a serious help for our members to know which documents they should consult in their research area within the team. It is also useful for finding similar works that one is not aware of.

In general, we are looking for an interaction between x and y, for example "which documents (x) are on a common topic (y)?". To answer the question we have studied the relations between x and y (see Table 2), and for this we have constructed a Galois lattice.

	ont.	sw	cbr	assoc. rules	bioinfo.	adapt.
cadot03b				x		
cherfi03c				x	x	
daquin02a						x
daquin03a	x	x	x			
lieber02a			x			x

Table 2: Articles × keywords input table

(In this example we have taken some of our publications, but to keep the example in moderated size we have only taken some keywords into account). In the corresponding Hasse diagram (Figure 6) the extension set of concepts contains documents, the intension set of concepts consists of keywords describing these documents. From the diagram we can read the answer to our question by examining the concepts: we can see that *{cadot03b, cherfi03c}* are on *association rules*, *{daquin02a, lieber02a}* are on *adaptation*, and *{daquin03a, lieber02a}* are both written about *case-based reasoning*. In our experiment we constructed a lattice with all the publications (147) and keywords (89), having 253 concepts in the lattice.

It is possible to extract association rules from the lattice, for example: "bioinfo." ⇒ "association rules", which means: every paper that has the keyword "bioinfo." also contains "association rules".

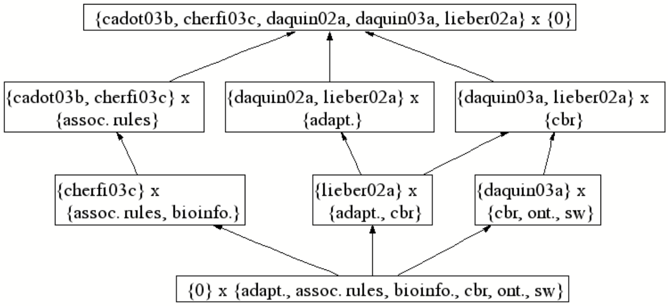

Figure 6: Documents on a common topic

4.3.2. Second question: "Who are the persons working on a common topic?"

In the next step we were interested to know who is working on a common topic within the team. For the members in a team it is useful to know who else work on the same topic, whom they can contact to consult on a problem. In this case we need to look for interactions between people and topics (see Table 3). We have constructed another Galois lattice, which differs from the previous example in the sense that the extension set of concepts contains authors. The intension set consists of keywords that are common in publications of all the people present in the extension set.

	repr.	bioinfo.	classification	cbr
Lieber, J.	x	x	x	x
Brachais, S.	x			
Le Ber, F.			x	x
d'Aquin, M.	x	x		x

Table 3: Authors $\times$ keywords input table

Again, the example contains just a few keywords to keep the lattice in moderated size. From the Hasse diagram (Figure 7) we can read the answers we are interested in, i.e. *{Brachais, S., Lieber, J., d'Aquin, M.}* work on *representation*; *{Le Ber, F., Lieber, J., d'Aquin, M.}* work on *case-based reasoning*; *{Lieber, J., d'Aquin, M.}* work on *{bioinformatics, case-based reasoning, representation}*; and *{Le Ber, F., Lieber, J.}* work on *{case-based reasoning, classification}*. In our experiment we took all the authors and co-authors of our publications (96) and the reduced keyword set (89 keywords). The resulting Hasse diagram has 287 concepts.

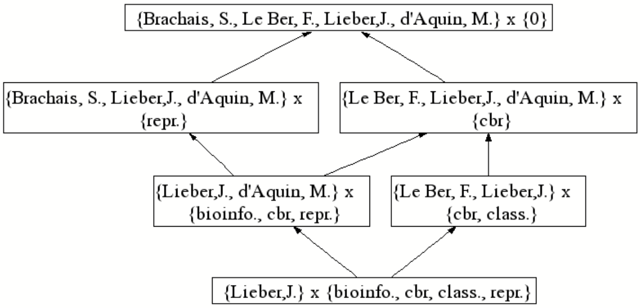

Figure 7: People working on a common topic

4.3.3. Third question: "How to rank documents by a keyword?"

In the final example we will show how we rank our publications by their relevance of a given keyword. For this we have adopted and implemented the algorithm described in [CR00]. The input matrix is similar to the one shown in Section 4.3.1, but the ranking algorithm requires to change the place of documents and keywords, e.g. we need to store the keywords in the extension set of concepts, while we keep the documents in the intension. To avoid the problems mentioned in Section 4.2, we have used again the cleaned keyword set.

Let us consider the following example with six documents and seven attributes describing them (see Table 4). For the sake of simplicity we will use "d_1", "d_2", etc. instead of document IDs. Suppose we want to order the documents by the keyword "ontology".

	d_1	d_2	d_3	d_4	d_5	d_6
1 graphs	x					x
2 description logics	x					x
3 case-based reasoning	x		x			x
4 data mining		x		x	x	
5 agronomy		x				x
6 ontology	x		x	x	x	
7 semantic web			x		x	

Table 4: Keywords × documents table for the ranking method

In the first step the Hasse diagram of the lattice has to be constructed. After that the query (in our example it is the keyword "ontology") is mapped on the concept whose extension contains only this keyword, it will be the 'query node'. Then each node has to be labelled by its distance from the query node. This third step can be accomplished by a breadth-first search.

If the intension of the Top concept is empty, it can be deleted. We can do the same thing with the Bottom concept if its extension is also empty. These two steps are optional and they only make sense in visualisation. The resulting graph after all these steps is depicted in Figure 8.

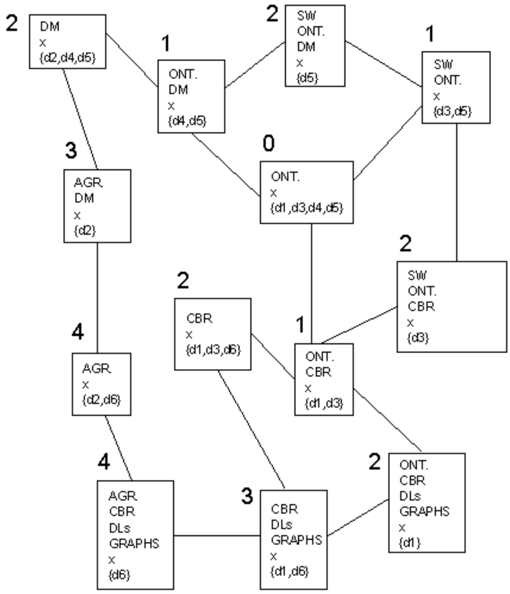

Figure 8: The resulting graph

To order the documents, we need to group the document sets at the different distances (Figure 8):

Level 1: $\{d_1, d_3\}$, $\{d_3, d_5\}$, $\{d_4, d_5\}$
Level 2: $\{d_1, d_3, d_6\}$, $\{d_2, d_4, d_5\}$, $\{d_1\}$, $\{d_3\}$, $\{d_5\}$
Level 3: $\{d_2\}$, $\{d_1, d_6\}$
Level 4: $\{d_2, d_6\}$, $\{d_6\}$

The final step of the ranking process: the rank of a document equals the level of the smallest set that contains the document. By this the rank of a one-sized set is unambiguous (d_1, d_3, d_5, d_2, d_6). The document d_4 is present in two sets ($\{d_4, d_5\}$, $\{d_2, d_4, d_5\}$) of which $\{d_4, d_5\}$ has less elements. The level number of this set is 1, thus the rank of d_4 is 1. The final result (rank indicated in parenthesis): the most relevant document is $d_4(1)$, then $d_1(2)$, $d_3(2)$ and $d_5(2)$, finally $d_2(3)$ and $d_6(4)$.

The advantage of this method is that it can order all the documents, even if a document does not contain the specified keyword. Naturally, the most relevant documents are at the beginning of the ranked list, thus setting up a limit can be a good idea, such as displaying only the first 20 documents found.

4.4. State of the system

For the implementation we have used Java Servlet/JSP technology. Under Java there are several ontology APIs available for RDF(S), OWL, etc. thus manipulating ontologies in this language is fairly easy. We have used the Jena2 ontology API [Jen]. To create ontologies we used the Protégé-2000 editor, and as an ontology language we have used RDF(S). RDF is a datamodel for relations between "things", and RDF Schema adds vocabulary for RDF. The newest ontology language is OWL, which is an extension of RDFS. It is likely that we will change to OWL soon.

To visualise lattices we have used the freely available Graphviz[4] package that permits to draw the whole graph, but the navigation in a large lattice with more hundred concepts can become difficult. We plan to try other visualisation methods like the fisheye view technique (from the focus node the other nodes are displayed in decreasing levels of detail and at increasing graphical distance), and hypertrees.

5. Synthesis

In this paper we have investigated the application of Galois lattices in a small enterprise, namely within our research team, for guiding the knowledge discovery process. We have proposed to use domain ontologies in several steps of the KDD process. We have shown how to use ontologies for data cleaning to avoid some problems that can arise by an exact keyword-based search. Furthermore, we have used the same ontology to connect it with Galois lattices. We were interested to find relations between people, documents, and topics in our team, and we have confirmed the results with several examples.

For our work we consider three projects as reference works. Galicia is a tool for lattice construction and visualisation. Galicia implements several algorithms in both batch and incremental mode, including iceberg lattices [Gal]. The Weka tool is a collection of machine learning algorithms for data mining. Weka supports several tasks of the KDD process, including classification and extracting association rules [WF99]. Both of these tools are implemented in Java and freely available. SEAL (SEmantic PortAL) is a framework to build community web sites. It uses ontologies as key elements for managing community web sites and web portals. SEAL serves as the core methodology underlying the OntoWeb[5] portal. We have chosen these projects for our reference works because our final goal is to create a semantic portal for the team with integrated data-mining capabilities. Building an intranet portal is today a standard first step in knowledge management, and with a portal it is possible to draw together on the desktop all the important information from both inside and outside an enterprise [MSS+02].

Our approach can also be used on different data sources to reveal relations between object sets and attribute sets. In the first step a domain ontology has to

[4]http://www.graphviz.org
[5]www.ontoweb.org

be built (from scratch or rather reusing already existing ones) that can function as a filter for cleaning the attribute set. In the second step the reduced attribute set can be used in Galois lattices.

6. Perspectives

As future work, first we would like to investigate the automatic rule extraction from Galois lattices such as "In 68% of cases, author A has published with author B." Next we plan to change to the OWL ontology language after RDF(S) as OWL is a more expressive language.

Another interesting area is the application of iceberg concept lattices. Iceberg concept lattices show only the top-most part of a concept lattice, and they can be used in KDD as conceptual clustering tool, as a visualisation method, as a base of association rules, and as a visualisation tool for association rules [STB+02]. We want to make experiments with iceberg concept lattices connected with a domain ontology. By descending in the ontology and increasing the granularity we can build a family of lattices that would reveal more and more information.

While our keyword ontology is appropriate for our bibliographical items, it is not adequate enough for searching on the Web since the keyword list is not complete for this task. On the one hand a more complete list has to be built, for example by reusing similar ontologies with an ontology mapping method [KS03]. On the other hand, some kind of approximate matching may be needed.

We started to build a portal for our team that can only handle internal sources (bibliographical items) at the moment. We want to extend it with the capability to use external sources too to be able to answer questions like "What are the conferences in the next half year in which I may be interested?". The realization of such a system raises several questions but the main idea here is that machines would be searching the Web to find important information. Our view relies on the Semantic Web principles: the idea of having data defined and linked in a way that it can be used by machines not just for display purposes, but for automation, integration and reuse of data across various applications, and reasoning on documents [BLHL01].

References

[AMS+96] R. Agrawal, H. Mannila, R. Srikant, H. Toivonen, and A. I. Verkamo. Fast discovery of association rules. In *Advances in knowledge discovery and data mining*, pages 307–328. American Association for Artificial Intelligence, 1996.

[BLHL01] T. Berners-Lee, J. Hendler, and O. Lassila. The Semantic Web. *Scientific American*, 284(5):34–43, May 2001.

[Bog03] K. Bognár. Aspects théoriques de la classification à base de treillis. *Teaching Mathematics and Computer Science*, 1(1):125–135, 2003.

[CR00] C. Carpineto and G. Romano. Order-Theoretical Ranking. *Journal of the American Society for Information Science*, 51(7):587–601, May 2000.

[Dub01] Dublin Core Metadata Initiative, 2001. `http://dublincore.org/`.

[FPSS96] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. From data mining to knowledge discovery in databases. *AI Magazine*, 17:37–54, 1996.

[Gal] GaLicia – Galois lattice interactive constructor. `http://www.iro.umontreal.ca/~galicia/`.

[Gan99] B. Ganter. Attribute exploration with background knowledge. *Theoretical Computer Science*, 217(2):215–233, 1999.

[GG99] M. Goebel and L. Gruenwald. A survey of data mining and knowledge discovery software tools. *SIGKDD Explorations Newsletter*, 1(1):20–33, 1999.

[GMA95] R. Godin, R. Missaoui, and H. Alaoui. Incremental Concept Formation Algorithms Based on Galois (Concept) Lattices. *Computational Intelligence*, 11(2):246–267, 1995.

[Gru93] T. R. Gruber. A Translation Approach to Portable Ontologies. *Knowledge Acquisition*, 5(2):199–220, 1993.

[GVM93] A. Guénoche and I. Van Mechelen. Galois Approach to the Induction of Concepts. In I. Van Mechelen, J. Hampton, R. S. Michalski, and P. Theuns, editors, *Categories and Concepts: Theoretical Views and Inductive Data Analysis*. Academic Press Ltd, 1993.

[GW98] B. Ganter and R. Wille. Applied lattice theory: formal concept analysis. In G. Grätzer, editor, *General lattice theory*, pages 591–605. Springer Verlag, 2 edition, 1998.

[GW99] B. Ganter and R. Wille. *Formal concept analysis: mathematical foundations*. Springer, Berlin/Heidelberg, 1999.

[Jen] Jena – A Semantic Web Framework for Java. `http://jena.sourceforge.net/`.

[KMR+94] M. Klemettinen, H. Mannila, P. Ronkainen, H. Toivonen, and A. I. Verkamo. Finding interesting rules from large sets of discovered association rules. In *Proceedings of the third international conference on Information and knowledge management*, pages 401–407. ACM Press, 1994.

[KS03] Y. Kalfoglou and M. Schorlemmer. Ontology mapping: the state of the art. *The Knowledge Engineering Review*, 18(1):1–31, 2003.

[McG02] D. L. McGuinness. Ontologies Come of Age. In D. Fensel, J. Hendler, H. Lieberman, and W. Wahlster, editors, *Spinning the Semantic Web*. MIT Press, 2002.

[MSS+02] A. Maedche, S. Staab, N. Stojanovic, R. Studer, and Y. Sure. SEmantic PortAL - The SEAL Approach. In D. Fensel, J. Hendler, H. Lieberman, and W. Wahlster, editors, *Spinning the Semantic Web*. MIT Press, Cambridge, MA, 2002.

[PBTL99] N. Pasquier, Y. Bastide, R. Taouil, and L. Lakhal. Efficient mining of association rules using closed itemset lattices. *Inf. Syst.*, 24(1):25–46, 1999.

[STB+02] G. Stumme, R. Taouil, Y. Bastide, N. Pasquier, and L. Lakhal. Computing Iceberg Concept Lattices with TITANIC. *Data and Knowledge Engineering*, 42(2):189–222, 2002.

[WF99] I. H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kaufmann Publishers, 1999.