Picture this:
Mapping knowledge in higher education organizations

Willem-Olaf HUIJSEN
Telematica Instituut, P.O. Box 589, 7500 AN, Enschede, The Netherlands
Wolf.Huijsen@telin.nl

Harry VAN VLIET
Telematica Instituut, P.O. Box 589, 7500 AN, Enschede, The Netherlands
Harry.vanvliet@telin.nl

Henk PLESSIUS
Utrecht University of Professional Education, P.O. Box 182, 3500 AD, Utrecht, The Netherlands
h.plessius@fnt.hvu.nl

ABSTRACT

Universities of Professional Education have a vast amount of quality knowledge available in their staff, students, curricula, information systems and other sources. Although this is generally agreed upon, it is not easy to get an overview of this available knowledge from an individual perspective, say a lecturer or student. At the Utrecht University of Professional Education, a demonstrator has been developed of a knowledge-mapping tool to address this problem. Knowledge mapping is the field within knowledge management that aims to optimize the effective and efficient use of an organization's knowledge base by addressing the issues of how to support finding the knowledge that is available within an organization and how to build insight into the qualities of this knowledge. This paper describes the underlying principles of the knowledge mapping tool and the results of developing and introducing this tool within the Utrecht University of Professional Education.

Keywords: knowledge mapping, informatics, knowledge management, ontology's, information management, university of professional education

INTRODUCTION

Part of the strategy of Utrecht University of Professional Education, a University of Professional Education in the Netherlands, is to present itself through twelve key fields of knowledge. One of which is informatics. Currently staff, curricula, students and facilities regarding this key area are spread across three faculties (Economics & Management, Journalism & Communication and Science & Technology). The PICTURE project (http://www.picture.hvu.nl) was initiated in 2003 to stimulate closer collaboration between these faculties and to work towards one network organization that presents itself transparently to the outside world. The PICTURE project has set ambitions with regard to integrating curricula development, innovative research projects, changes in culture, and introducing supporting facilities for this key area of informatics. One supporting facility will be a portal, part of which will be a knowledge map representing the available knowledge and expertise on informatics across the three faculties in order to help lecturers and students to get the most out of the available knowledge. [1]

In many organizations, available knowledge and expertise is almost omnipresent: it is spread across people, documents and information systems (including Internet). So looking for certain knowledge typically involves approaching (close) colleagues one knows to have expertise on the subject. Still, there may be others with high-quality knowledge one simply misses because one does not know them (yet). Documents, as a second source, often come in large numbers, and with a poor structure to them, making a quick and effective search impossible. The third source, that of information systems, tend to be numerous too, and each system has a different interface and internal structure, so that finding knowledge and piecing together information from across a number of systems is a lot to ask.

This difficulty of getting a clear and complete overview of the knowledge that is present and available within the three faculties, or one organization for that matter, is tackled here by the concept of knowledge mapping. Knowledge mapping is about making the knowledge that is available within an organization transparent. In this paper we describe the process of knowledge mapping, its tools and specific issues that play a role in designing knowledge maps. Next we briefly present a prototype that was built: the 'Knowledge Cockpit'. Currently we are applying this knowledge in the context of the PICTURE project in order to have a demonstrator up and running in May 2004 in this specific educational context. We will conclude this paper with a short evaluation and discussion on the benefits of knowledge maps for Universities of Professional Education.

KNOWLEDGE MAPPING: PROCESS AND TOOLS

Knowledge mapping is the field within knowledge management that aims to optimize the effective and efficient use of an organization's knowledge base by addressing the issues of how to support finding the knowledge that is available within an organization and how to build insight into the qualities of this knowledge. An individual knowledge map is a view on one or more aspects of the knowledge that is available within an organization. [2]

Knowledge mapping tools provide computer support for addressing these issues. These tools may interface with any number of information systems to collect data on the knowledge that is available within the organization. For example, documents in the document-management system can provide clues on who knows about which subjects (making it possible to find experts), and project data may reveal the topics that are addressed by which projects, and who participates in those projects.

The information systems from which data may be gathered include the following:

- Administrative databases, such as calendar system, document-management system, hours database, library database (collection and usage), project database and yellow pages database.
- Communication systems, such as: community server, E-mail server, newsgroup server, telephone system and web server.
- Others, for instance presence server and elearning systems.

The functionality of the knowledge-mapping tool has two major components. On the one hand, it taps into as many sources of information as possible to extract and analyze all relevant data, so that users can be presented with overviews of the corporate knowledge base. On the other hand, the tool also aggregates and contextualizes the collected information so as to support the more abstract building of insights into the qualities of this knowledge.

The knowledge-mapping process can be said to consist of the following steps (see Figure 1). First, raw data is acquired from one or more sources through system-specific interfaces. This also typically involves some basic processing such as filtering or keyword extraction. The resulting first-order data is stored in the knowledge-mapping database. Although this data may be interesting in itself, and some knowledge maps can be constructed on the basis of this data alone, in order to obtain more meaningful information, it may be further analyzed, aggregated, and conceptualized, resulting in higher-order data. By visualizing the first-order and higher-order data in specific ways, and making into account user preferences, knowledge maps can be produced that provides the user with relevant insights into the knowledge that is available within the organization.

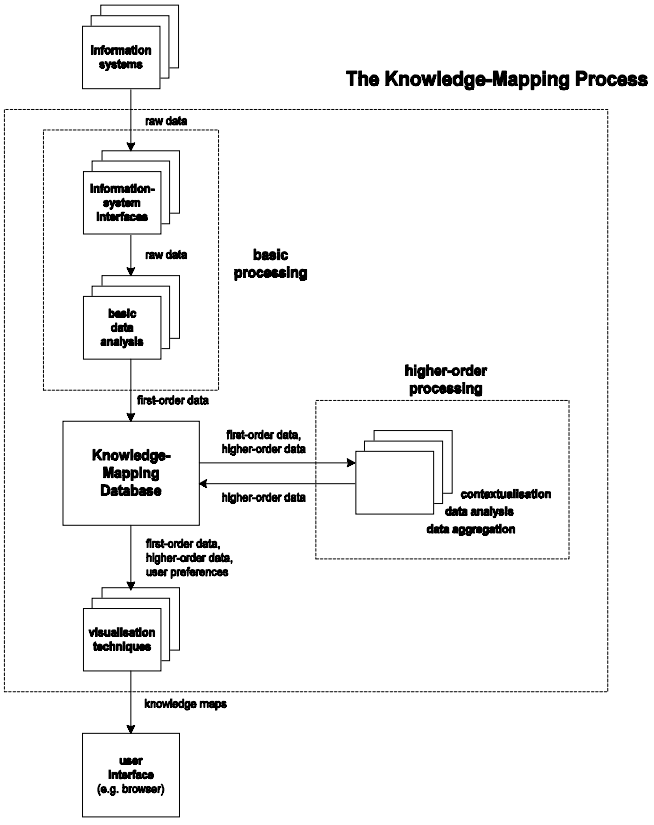

Figure 1: The knowledge mapping process

A knowledge-mapping tool also typically contains a conceptualization of the subject domain, for instance informatics. Such a conceptualization makes it possible to classify knowledge resources. This may be a simple list of (controlled) keywords, a more complex thesaurus with synonymy and other basic relations between concepts, sophisticated ontology's with lots of domain knowledge encoded into them, or even semantic networks. A knowledge-mapping tool uses the conceptualization in several situations: (1) in the indexing of knowledge sources, (2) in concept naming for user interaction, (3) in browsing the concept space, and (4) in relevance ranking.

Each of these situations has its own requirements on the conceptualization. [3] From the perspective of knowledge mapping the following requirements seem valid. First, the conceptualization must include, for every concept, (a) all terms that describe the concept: synonyms and spelling variants, (b) one of those terms must be declared as the preferred term, and (c) related concepts such as hyperonyms, hyponyms, and others. Second, the conceptualization must include a set of different relation types. At a minimum, subsumption and association are required. More relations are needed if the application requires so, or if the number of relation instances with the same name becomes impractical.

Given these requirements simpler structures such as keyword lists and taxonomies – although very useful for other purposes – are not considered likely candidates for a knowledge map application, as they fail to meet basic requirements such as concept browsing. Regarding the other possible means of conceptualization (thesauri, ontology's and semantic networks) the case is not so clear-cut which to prefer. From a technical point of view – does the conceptualization account for fully integrated conceptualization? - ontology's and semantic network satisfy these criteria, whereas thesauri do not. From a user perspective – how easily can a conceptualization be built? - the thesaurus is the simplest and most straightforward to build (and maintain). Ontology's and semantic networks are much more difficult to build (and maintain). The complexity of the semantic network is higher than for the ontology, due to the fact that no difference is made between concepts and instances, and the possible relations are endless. Thus, when we have to make a choice between these conceptualizations we advise, after the user defines his user needs, to choose between the simple approach by using the thesaurus for the knowledge mapping systems or the complex approach, by choosing for the semantic network. Although from a pragmatic view ontology's may be a way out if there are already ontology's available. [4] [5] [6]

ISSUES IN KNOWLEDGE MAPPING

Several issues are important in developing and implementing a knowledge-mapping tool. We will discuss five of them: intended user roles, manual entry versus automatic extraction, ranking, security and privacy, and portability. [7] [8]

Who to support? In constructing a knowledge-mapping tool, one must consider who will use it, for what purpose, how it will be used, and which user requirements are stated. These factors determine the exact functionality of the application, and which data must be collected as input to the system. In turn, this has implications for the sources from which this data is collected. It is our view knowledge maps are relevant to everyone within an organization since all members of an organization deal with organizational knowledge on a regular basis. However, different organizational roles and different personal preferences mean different user needs. A knowledge-mapping tool can address this by having different sets of knowledge maps for different organizational roles and by

having options to customize the knowledge maps to personal preferences.

Common roles and their typical needs with regard to knowledge maps, in an educational context, are:

- Lecturer: a lecturer needs to know others lecturers who have expertise on a specific technical subject, so that they can contact them for help. This requires very detailed data on the subjects that lecturers have expertise on. Typical questions are: Who has knowledge on subject X? Who has knowledge on subject area Y? Who can help me with subject Z?
- Student: in the same way as the lecturer the student wants to know interests and expertise of others students in order to work together on projects, but they may also like to know the expertise of lecturers for choosing classes or apprenticeships.
- Management: Management needs to set organizational strategy taking into account the available expertise, and, in alignment with this, needs to define the course for knowledge development in line with students demand, developments in the market, etcetera. This requires a broad overview of the knowledge that is available within the organization. Typical questions are: How many people have expert knowledge in subject area X? How many people have ambition for developing their knowledge in subject area Y? Is there a gap between our current knowledge and future market developments?

Manual entry versus automatic extraction: Information such as personal profiles can be entered manually or extracted automatically. Manually entered data can easily be tailored to specific needs, and can be very accurate. Also, it is easy to include data about people's ambitions. On the downside, filling out (detailed) personal profiles is very tedious and labor-intensive. Such profiles tend to get out of date quickly too. The more detailed the profiles are, the more labor-intensive it becomes, and the faster they will be out of date. Automatic extraction on the other hand is always up to date, does not burden the work force with filling out profiles, and can be very detailed. However, it is hard to tailor to a specific purpose, and it may be less accurate since the process cannot benefit from human judgment. For example, mining from text is notoriously vulnerable to the complexity of natural language. Newcomers that pose a lot of questions on a specific subject may be mistaken for experts, and negations ("I do not know anything about X") may not always be caught. Furthermore, it is often impossible to mine for data on people's ambitions. Both methods have their advantages and disadvantages, and each is applicable more readily to one type of data than to others. It is our belief that any nontrivial system should use both, applying each where it works best, thereby combining the strengths of both. For instance start with automatic extraction and then let people correct or add information.

Ranking: Whenever data from multiple sources is combined to yield a single conclusion, the ranking problem arises: what weight should be associated to each of the sources so that the resulting combination is satisfying? For example, consider a set of documents for which we want to know how relevant they are to a certain subject (e.g. ‘communities of practice’). One can think of several approaches. For example, one may simply count the number of occurrences of a term that describes the subject (say "community"). Thus, a document containing 42 occurrences would be ranked higher than a document containing only 18 occurrences. Of course, the latter document might still be more relevant to communities of practice than the former, for example because the former is a much larger document that only refers to communities without really discussing them, while the latter is shorter but has communities as a central theme. Thus, one may further refine the combining formula, for instance by taking into account the document length or document age.

Security and Privacy: In a corporate setting the information stored in an organization’s documents and information systems is to a large extent confidential, either to protect corporate interests (security) or personal interests (privacy). For example, overviews of corporate expertise in its key areas of interest must in general clearly not be made accessible outside of the organization. For the same reasons, the information in a knowledge-mapping tool must be protected. One can even argue that because such a tool makes accessing and interpreting corporate knowledge so much easier, it is even more vital that the information be protected. The situation in an education (public) organization is different exactly because of its public nature, although still some information can be classified as confidential (financial figures, assessment information of people, etcetera).

Another issue is that of how far a knowledge-mapping tool should be allowed to go in collecting and analyzing the behavior of individual employees. In the technical perspective, pushing the limits of what is feasible opens up many new possibilities. In the personal perspective, some of this can be seen as an invasion of privacy. For example, by analyzing an employee's e-mails and phone conversations it can be inferred which subjects people are discussing, and what their social networks are. Technically speaking this can provide valuable clues, amongst others, on available expertise, on who the recognized experts are, and on knowledge flows within the organization. Nevertheless, knowing that all communication is monitored by computer systems, and not being sure of which aspects are analyzed, may very well be disconcerting. Especially since many organizations allow employees to use e-mail and phones for (limited) personal use. But even if applied only to professional communication, monitoring all communication may still not feel right. One way out of this is giving employees control over which results from the analysis may be used in the knowledge-mapping tool. Consider as an example scanning e-mail for topics. After analyzing the e-mail communication of the past week, the tool can present to the employee the topics it Identified, and can ask the employee to mark which results it may and which it may not use. Only after the employee has done this will the results become available for further processing in the knowledge-mapping tool.

Although the privacy issue is a very delicate one, the value of the information that can be gained is so great that it pays to have a good look at what can and may be done without invading individual employees' privacy or even without giving the feeling that that might be the case.

Portability: Ideally, there would be a single knowledge-mapping tool that can be used by any organization. Unfortunately this is not as easy as it sounds. By their very nature, knowledge-mapping tools have a number of aspects that bind them to a specific organization. These aspects are:

1. The knowledge domain: Each organization has its own knowledge domain with different concepts, with different definitions of those concepts, and with different terms to describe those concepts. The concepts and terms may not even exist outside the organization.
2. The information systems: Each organization has his own set of information systems. There are not only differences in terms of the specific systems used, but also in terms of the type of data stored, the level of detail of the data, and the specific format used.
3. The user requirements: Different organizations are likely to have (slightly) different employee roles and most certainly have different individual employees. For a knowledge

mapping tool, this means different users requirements. Fortunately, the user requirements for a knowledge mapping tool for different roles and individuals tend to overlap a lot, so that it is very well possible to specify a set of user requirements that cover the majority of user requirements that one can think of.

In a corporate environment these aspects seems to hold up nicely, although the dominance of certain information systems (e.g. Windows software, SAP) bring about some homogeneity. In an educational context the knowledge domain, such as informatics, is shared and very often thesauri and ontology's are available, and also the user requirements may be more alike than across corporate organizations. But that does not mean everything is clear-cut: in the situation of the Utrecht University of Professional Education there are four geographically dispersed locations and there are different information systems at the different faculties.

A PROTOTYPE: 'THE KNOWLEDGE COCKPIT'

As part of the METIS project (http://metis.telin.nl), in which Basell, Océ Technologies, Telematica Instituut and several universities are collaborating, a proof-of-concept prototype was built at the Telematica Instituut. This prototype 'The Knowledge Cockpit', implements a knowledge-mapping tool that interfaces to existing information systems to scan the document-management system, project data, and web traffic, presenting tabular and graphic knowledge maps with various views on the collected data. A graphical concept map can be used for navigation. The Knowledge Cockpit is used as a testbed for trying out new ideas. The screen depicted in Figure 2 shows a knowledge map of people found with expertise in knowledge management.

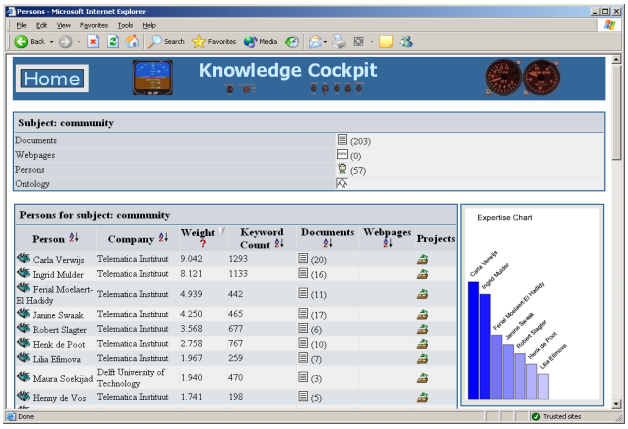

Figure 2: The METIS 'Knowledge Cockpit'

Furthermore a mock-up was constructed, 'The Knowledge Visualiser', that gives an impression of more sophisticated knowledge maps that can be constructed if more aggregation, analysis, and visualization is done (see figure 3). Promising new ideas from the Knowledge Visualiser can be implemented in the Knowledge Cockpit. Below is a screen from showing expertise of different people taking into account the amount of expertise (height), the recency of this expertise (color) and the free capacity in hours of the person (width).

Figure 3: The Knowledge Visualiser: a Mock-up

KNOWLEDGE MAPS IN THE PICTURE PROJECT

As outlined in the introduction, within the Picture project, our ambition is to make available all ICT knowledge in the Utrecht University of Professional Education to the lecturers and students, so as to stimulate the exchange of knowledge within the university. Different from the situation in the Metis project, the University of Professional Education of Utrecht does not have a central document-management system. However, the University does have a well-developed course information system in which for every course given, a number of vital characteristics are registered. This system makes it possible to match lecturers to course descriptions, and in this way to derive a first approximation of the knowledge of every lecturer. To improve on this information, lecturers will be asked to supplement this first derivation with documents from various given courses (sheets, hand-outs, etc), which can be used to refine the knowledge map for the lecturer. This way of working introduces quite a lot of maintenance, but we feel it may open up the introduction of a document-management system. Another system which proved very valuable is the employee database which lists a number of characteristics of each employee, e.g. faculty, room number, telephone number, e-mail address, etcetera. We decided to give the employees the possibility to augment this data with more personal characteristics (like personal websites, other e-mail addresses, comments, etcetera).

As a number of information systems will be renewed in the years ahead, we decided to build a data warehouse to guarantee data independence. For the structured data in this warehouse, we chose SQL Server, as quite a lot of experience with this DBMS is available in the University, and for the documents we set up a simple directory structure. Since it should possible to query the knowledge map with a browser, a HTML-based interface seemed a logical choice. We decided to keep the ontology basically simple. An element in the ontology may have one ore more parents and one or more children (subsumption relations). All other relations are mapped as associations. Furthermore, an element may have any number of variants, which include synonyms, different spellings, shorthand's and translations in other languages (in our case, Dutch and English). The ontology may thus be implemented in SQL Server by the schema shown in figure 4. An initial ontology has been developed by starting with the ACM-ontology [9] and completing it by scanning a lot of available documents, which are used in the ICT courses of the University. As every query involving the ontology will be saved, we expect that the ontology may be further completed on a short term.

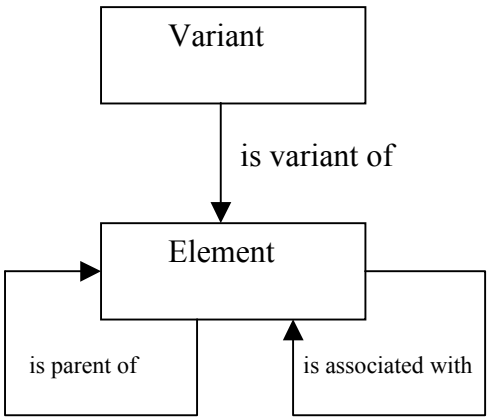

Figure 4: Schema for ontology implementation in SQL Server

Concerning the content in the user interface, five primary concepts should be visible including meaningful relationships between these concepts: employee, personal employee characteristics, knowledge, courses, and documents. As for the user interface itself, after a lot of discussion we decided on a 3-screen model. In the first screen a concept (like a knowledge element) can be chosen, which can then be related to a second concept (e.g. employee) in the second screen, thereby implementing the query ´all employees, which have the selected knowledge element´. In the third screen, a graphical representation of a relevant part of the ontology (if applicable) would be shown. Furthermore, it should be possible to select a concept in one screen and have the other screens updated directly.

As we wanted to build a demonstrator first, we decided to use the portal DotNetNukeXL for a ´quick and dirty´ implementation. We succeeded in a quick implementation, but due to restrictions of DotNetNukeXL not all functionality could be implemented. Especially the graphical screen could not be implemented in the way we wanted. Worse, the ´click anywhere´ function could not be implemented. For a final version we suspect we will have to start from scratch (probably using Java). Figure 5 shows the knowledge map for a given employee.

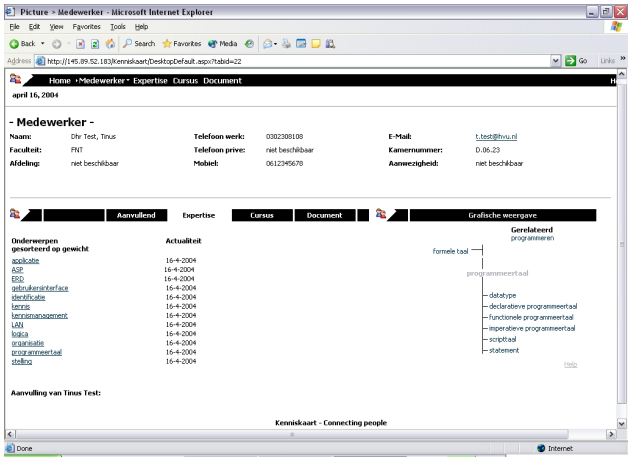

Figure 5: Knowledge map in the PICTURE project

The first screen (top) shows the characteristics of one employee. The second screen (bottom left), displays the most relevant concepts associated with the employee, while the third screen (bottom right) contains a ´graphical´ representation of a relevant part of the ontology is shown: in this example, a main element, *programmeertaal* (programming language) is shown in the middle together with its parent, *formele taal* (formal language) and its children. A related concept, *programmeren* (programming), is also shown in the upper right hand corner.

EVALUATING THE KNOWLEDGE MAPS APPLICATION IN THE PICTURE PROJECT

While building the knowledge map we sent a newsletter to inform interested employees of our progress. This proved very valuable since this made people feel involved in the development process, and they reacted accordingly. In this way we kept in touch with a number of employees, the characteristics of whom have been included in the demonstrator system. The first reactions are encouraging: employees like the automatic generation of the knowledge map. On the other hand, not everyone wants to enter personal characteristics in the knowledge map for privacy reasons (although most employees do not object). As the knowledge-mapping tool is a new type tool for most if not all employees, they find it hard to say how they expect to use the knowledge map in their daily work. We expect that usage will grow in time, especially when students have access to the knowledge map.

Based on the feedback we plan to build a new version of the knowledge map in the fall of this year to be rolled out among the lecturers on the Picture project. We also plan to introduce one or more communities of practice [10] that will be tightly coupled to the knowledge mapping tool, thereby creating a new platform from which we can derive elements for the knowledge map.

Based on the reflections above and our experience with developing 'The Knowledge Cockpit' and the demonstrator in the PICTURE project, several guidelines were formulated concerning the introduction of a knowledge management tool:

Strategic:

(1) Carefully consider security and privacy issues.
(2) Carefully communicate to the employees, staff councils, and managers regarding the privacy issues.
(3) As with any system development effort, it cannot be stressed enough that one must get the system requirements right. If this is not done well, the risk is high that a lot of time will be spent on developing the wrong things. Amongst others, clearly define the target user population, along with a realistic view on how the maps will be used.
(4) Follow an evolutionary path: first introduce a small-scale system that has a low technical risk and that is likely to be received well. Start small and build from there.
(5) Keep an eye on maintenance issues: don't build anything fancy that is hard to maintain.

Technical:

(6) Prepare to build an ontology for your organization, preferably re-using and building on existing ontology's.
(7) Standardize interfaces and exchange formats to information systems (APIs, SQL, XML).
(8) Build knowledge mapping tools with knowledge management databases that have a generic structure, so that typical data from the information systems can easily be mapped to the KMDB.
(9) Use a data warehouse to guarantee data independence.

User Interfacing:

(10) Analyze that questions can be posed and the 'way' to get to the answers.
(11) Design the 'central screen'.
(12) Visualize the relationship between main concept and related concepts in the interface.

DISCUSSION

Knowledge mapping is an exciting part of knowledge management research and implementation. Ongoing research deals with issues such as knowledge maps that use aggregation and visualization, the use of ontology's and Semantic Web technology, time dynamics of knowledge, different information-retrieval techniques (for example, techniques for ranking results), and portability of knowledge-mapping tools.

Developing and building a demo within the PICTURE project gives us further insight in how knowledge mapping tools 'behave' in different domains: co-operate versus public domain. As far as the ambitions go these are no less than any corporate organization would conceive: it is about quality and efficiency. The nowledge-mapping tool in the PICTURE project should contribute to the following results:

- Better education of young professionals by a more expanding cluster of experts;
- New courses through collaboration of different knowledge areas;
- Efficient use of resources (people and facilities);
- Better coordination in curricula;
- No unnecessary competition between ICT-curricula;
- Better distribution of knowledge.

Based on the development of the demonstrator and the reactions towards it within the Utrecht University of Professional Education, we think the results are promising enough to consider a further broadening of the knowledge mapping tool, this could be: 1) use more available information systems in building the knowledge maps, 2) include other key fields of knowledge, 3) connect more Universities of Professional Education. Anyone interested in this can monitor our progress on the PICTURE website (http://www.picture.hvu.nl).

REFERENCES

[1] H. Plessius, S. Koenders, F. Vodegel, **Business Plan Picture**. Utrecht University of Professional Education, Utrecht, 2003.

[2] W. Huijsen (ed.), **Knowledge Mapping**. Metis project deliverable D3.1.1. Telematica Instituut, Enschede, 2003.

[3] W. Huijsen, S.J. Driessen & J.W.M. Jacobs, "Explicit Conceptualizations For Knowledge Mapping", In: **Proceedings of ICEIS 2004, Sixth International Conference on Enterprise Information Systems**, Porto, Portugal, April 14-17, 2004, Volume 3, pp.231-236.

[4] W3C Semantic Web Activity (2002), http://www.w3.org/2001/sw/

[5] International Standard ISO 2788, Documentation. **Guidelines for the Establishment and Development of Monolingual Thesauri**, Second edition, 1986-11-15, International Organization for Standardization, 1986.

[6] "Guidelines for the Construction, Format, and Management of Monolingual Thesauri", NISO Press, Betheseda, MD. http://www.niso.org/standards/resources/Z39-19.html.

[7] J. L. Gordon, "Creating knowledge maps by exploiting dependent relationships," **Knowledge-Based Systems**, 13, 2000, pp. 71-79.

[8] A. Kolonić, "Océ Knowledge Map," A thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts in Information Management, Tilburg University, July 4, 2002.

[9] The ACM Computing Classification System [1998 Version], Association for Computing Machinery, Inc., http://www.acm.org/class/1998/.

[10] R. McDermott, "Knowing in Community: 10 Critical Success Factors in Building Communities of Practice", **IHRIM Journal**, 2000.