High Level Semantic Networking – Using k-infinity to build a multiontological learning enviroment

Roland Böving[1], Udo Bleimann[1], Christoph Wentzel[1], and Paul Walsh[2]

[1]h_da – University of Applied Sciences Darmstadt, Germany
[2]Cork Instiute of Technoloy, Cork, Ireland
roland@boeving.net

Abstract: Semantic linking of learning community information brings many advantages – the integration of various data sources, for example documents, collaborative learning units, and the flow of information into tasks and project processes [BB10]. It is anticipated that semantic web technologies will influence the next generation of e-learning systems and applications. By considering the learner preferences, we will develop a system where we envisage that the needs of the students are evaluated more efficiently and the learning will be more effective and retention will last longer. This paper proposes a solution by using semantic technologies to find the most suitable solution for the learner.
The virtual learning communities of the University of Applied Science Darmstadt (h_da) will be connected to the h_da-Knowledge-World and benefit by the content of this complex web-based ontology with more than 200.000 topics and over 5 million related document objects.
This complex topic network has been created by matching all kinds of different ontologies from the university environment. The supporting technology used for this is k-infinit, which enables imports in RDF and OWL standards and makes further refinements and thorough evaluation of the data possible.
[App05, BLHL01, PDM09, SLWD04, Wes07]

1 Approach

Creating a semantic network that takes real-world experiences into account.

Students' social skills, family background, gender, native language and personal preferences must all be taken into account when preparing learning material and deciding how it is to be presented. The idea is to use IT architecture to help students learn in a way that reflects their needs [Ble04] [BR06] [Mar08]. A self-learning system which constantly adds empirical data is required, as this enables the system to promptly and uniquely adapt the content of the learning material to individual students as well as groups.

The requirements for such a complex system are very high. Fundamental to it is an automatic evaluation of all types of attributes as well as the relations of the different knowledge objects to one another. A semantic topic network, which meets the needs of all faculties, staff members and students of a university, is the basis of this architecture. This means that such a topic network must reflect all the topics of all faculties. By linking all types of documents, whether they are thesis papers, books or technical journals, or research projects, lectures and e-learning modules, each individual topic node will become a supporting knowledge object. Expert knowledge can be accessed via such a topic node (figure 1).

A university topic network of this kind can hold more than 100,000 topics, to meet the needs of all user groups. Defining such a huge quantity of topics and linking them meaningfully with one another is a mammoth task and requires appropriate resources. We are in the process of linking as many of the existing h_da ontologies to one another as possible, so that we can quickly set up a high-quality topic network. These ontologies include the library topic tree, faculty media wikis, and a bought-in ontology focussing on technology and chemistry. Amalgamating these topic networks, all with their very different approaches, into a resultant ontology understood by users as “a world”, is the first major step towards a new interdisciplinary learning infrastructure.

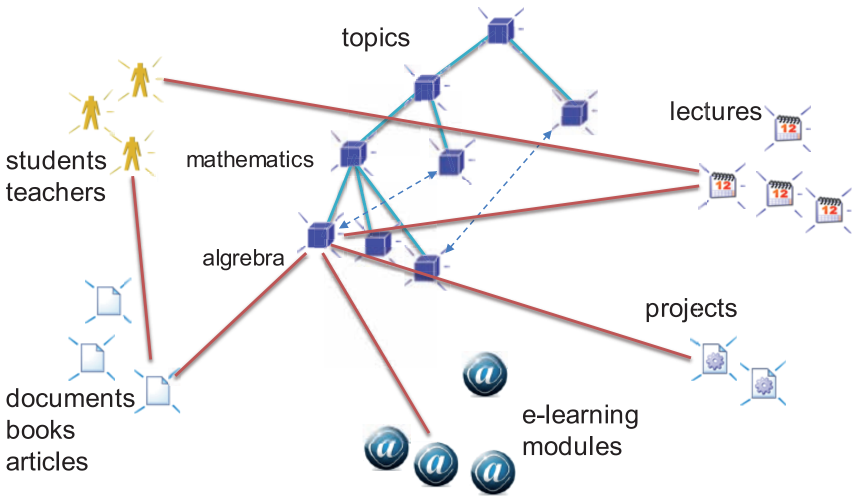

Figure 1: h_da-semantic web – showing relations between topics and content.

2 Semantic Network and User Interaction

The h_da-Knowledge World is a platform in which all the components for comprehensive support of knowledge processing will be synchronized with each other. From the simple definition of even the most complex knowledge models, through the use of semantic networks, knowledge can be widely distributed and used. A knowledge network of this type does not place words in a statistical context, but represent the context of the content. It represents the semantics of expressions because it stores knowledge objects that are linked by means of relations.

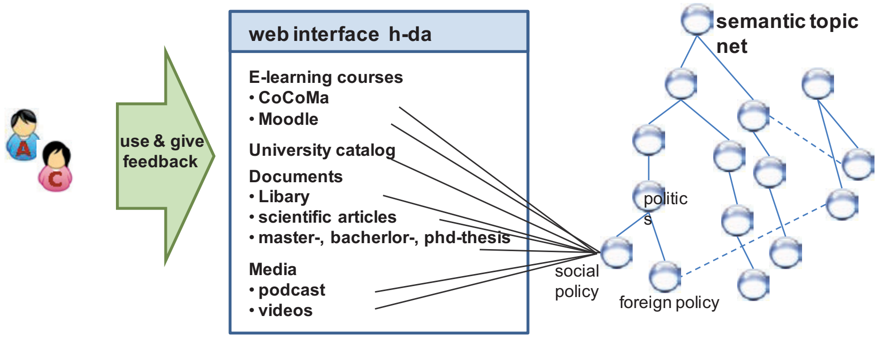

Figure 2: User interaction to the new semantic driven h_da-web interface.

In the future, all the learning options will be presented via the h_da web interface, and will be made directly available in some cases. On one hand, an overview of the seminars and e-learning offers will be given. Furthermore, the students have direct access to the appropriate course materials as well as further documents such as scientific articles, books or thesis papers. These materials are linked to each other through a semantic topic network and can be analysed even in very complex structures. Through continuous student feedback on the content (figure 3), the system can adapt to the learning requirements of the students and provides an individual, differentiated framework to every logged-in user.

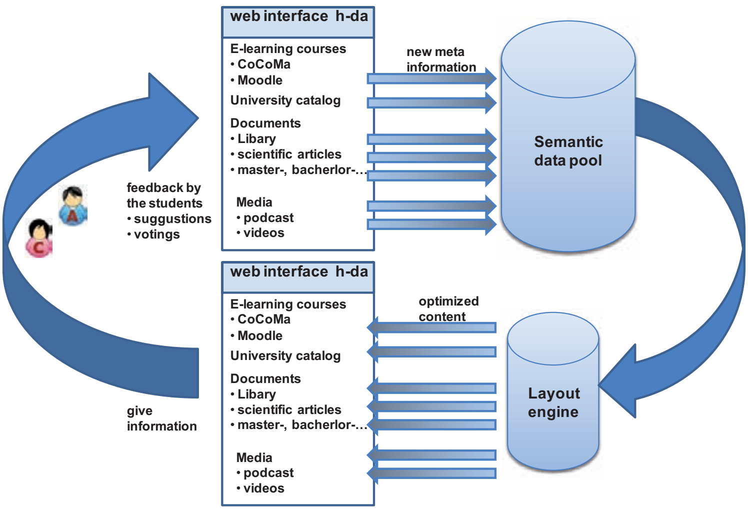

Figure 3: Approach: A self-learning system.

3 Technology

Use of semantic technology that interlinks all relevant content sources and therefore ensures comprehensive evaluation of inter-relationships and contents. This technology should offer the following features:

a) Fusion of complex ontologies which may be based on different data models

b) Attribute mapping and search functionalities for all object types

c) Links with external content sources

d) Unlimited synonym and language variants

e) Unlimited relationship variants

f) Attribute mapping based on relationship mapping

g) Graphic modeling

h) Adequate performance, even with ontologies involving more than 300,000 topics and 5 million instances

i) Evaluable and compatible with current web services

j) Ability to reason logically about properties and relations

As the main tool to create a own ontology, the software k-infinity will be used for this work. K-infinity combines its own data model for creating ontologies with graphical tools and opportunities for a direct web-based presentation of the results. The import of further ontologies based on RDF and OWL [Ras05] is possible. Imports of media wiki and xml-data records are also possible. One of the strengths of k-infinity is, to connect different complex data structures with each other and make knowledge evaluable.

4 System-supporting ontologies

What are the topics which students and staff members of a university are concerned with and how can structured data be accessed? The library is a main source of topics. The majority of h_da library books are tagged, with the tags, in turn, linked via different relations to one another. The quality of this data varies. On the one hand, it is a database which has grown historically and awareness of ontologies in 1960 was very different from today. Essentially, modern literature is tagged and linked in a different way. 38,729 topics, all directly linked to 577,300 books, were gained by importing h_da library master data. h_da library topics are not available in translation and do not have synonyms. Post-processing is necessary to combine language variants and similar topics.

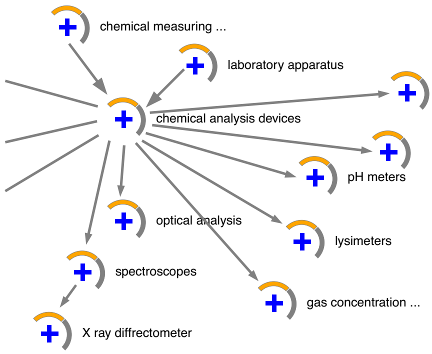

Figure 4: Detailed view on the h_da topic map. It contains more than 90.000 topics (40.000 translated, 130.000 synonyms)

The FIZ Technology thesaurus is a further source with very high-quality data, covering the technical sector in particular. Its import enabled 48,912 topics, directly connected to 3.5 million technical journals, to be integrated into the h_da knowledge world. The topics are also translated from German into English and have details of an additional 130,000 synonyms. Besides hypernym and hyponym relations the objects are to some extent linked to one another via the relation “is related to”. This relation appears 32,000 times. Identical object names between h_da library topics and FIZ Technology topics appear 6800 times.

Apart from these two main sources for topics we will also try to integrate as many h_da media wikis as possible, of which it is estimated that 60 are available. Media wikis have the advantage that in addition to the self-compiled topic description, they reflect the topics which the lectures and tutorials of the university are concerned with. In an initial trial, the IT-Management wiki was successfully imported from the Computer Science department with 350 objects. The next candidate is a media wiki from the Social Science department with approximately 3500 topic definitions.

5 Complex Ontology Matching – The Big Challenge

The import of these different data records into the h_da datapool means that a search can now locate more than 250,000 keywords in the form of topics, synonyms or translations.

Although the ontologies used have developed in very different ways, and strongly in content, to some extent they all contain the same topic areas and language affinities. Each source has its own metadata and specific content. However, all relations between the ontologies still require the critical judgment of users. The fact that certain topics can be found several times still remains a major problem. For example (figure 5), the term “management” appears in all of the three ontologies already imported and in each instance has its own content or is connected to other sources.

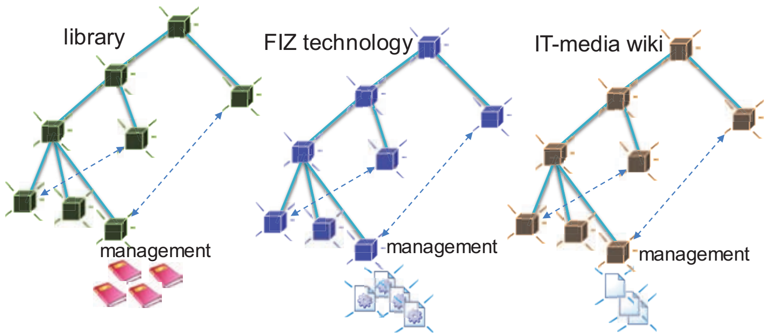

Figure 5: The term “management” appears in all of the three original ontologies.

It is vital that all the ontologies are combined to ensure that each search term can only be found once and that it is also only available once for future identification when books, professional articles and thesis papers are uploaded electronically. It must be noted that there will be regular updates from all data sources and it must be guaranteed that clear object identification remains. If only new topics are added this will never be a problem. What happens, however, when topics are renamed or disappear completely from ontologies? There is a big risk that h_da’s own topic content or links will be lost without any trace.

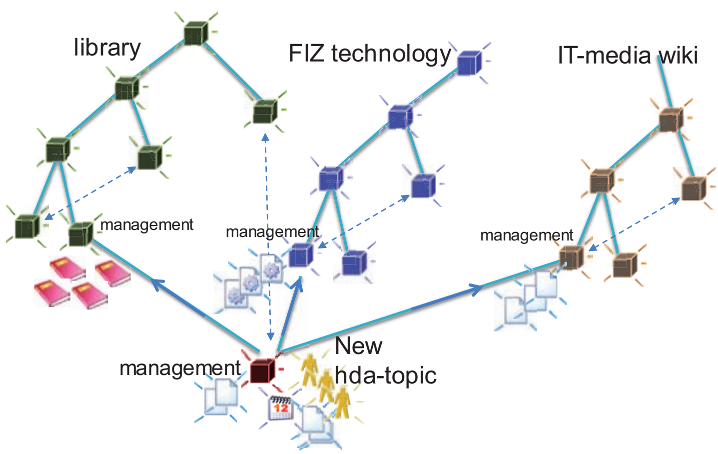

Figure 6: The solution – besides the independently imported ontologies, the ontologies already existing will generate a new h_da ontology, which will contain each knowledge object once. All additional content sources can be accessed via this h_da topic map, without impairing their structures.

The solution – besides the independently imported ontologies, the existing ontologies will generate a new h_da ontology, which will contain each knowledge object once. (figure 6).

This h_da topic network links all additional content sources such as projects, events, lectures, thesis papers, posters and professional articles of students and staff members. The content and structure of the imported ontologies such as FIZ Technology, the library and the media wikies are maintained within the structure of the new h_da ontology. They can be continually expanded by regular updates and remain consistent in themselves. They are summarized in the modified topic structure (figure 7) under foreign topics (imported), with the individual topics in the front end not available or traceable in either the semantic search or in the relation target searches of upload forms. The actual working plane is defined by the "topics h_da" network, which can be found under "own topics" in the structure. Any editorial changes in this network do not need to take account any other structures and the network can develop freely into a supporting ontology to meet the complex requirements of a university.

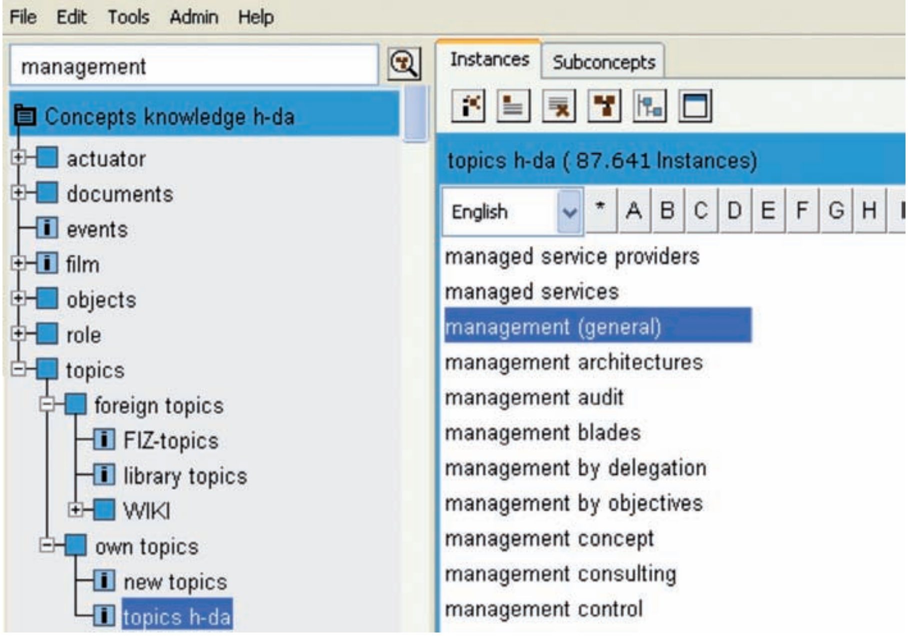

Figure 7: Concept structure in the k-infinity knowledge builder. Division into foreign topics and h_da's own topics is fundamental to this.

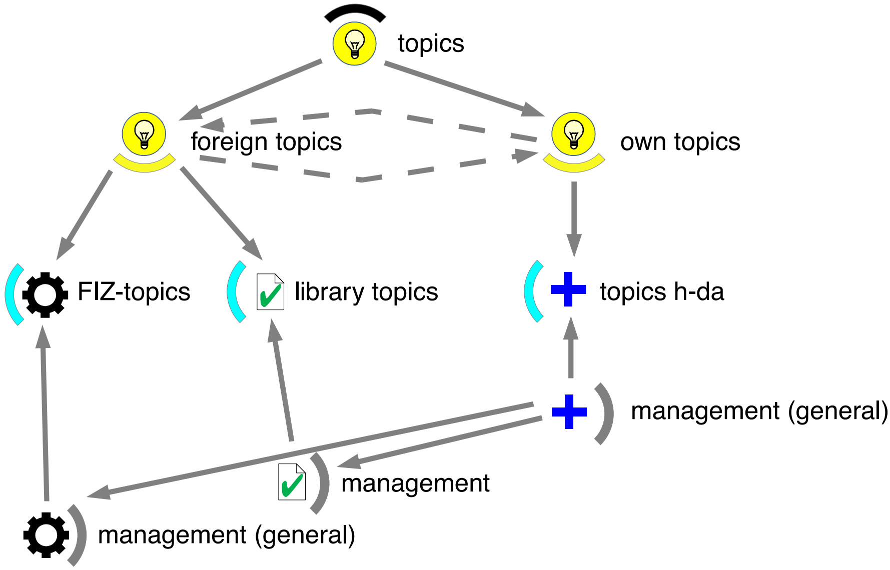

Figure 8: Diagram in the graph editor of the knowledge builder. The topics under h_da topics are linked via the "related foreign topic" relation, in case there are identical objects. The contents of these objects are evaluable and available via this link.

Staff members and students have the opportunity to assign their documents, such as professional articles, posters, bachelor and master thesis papers, to specific topics through upload functions. At the same time it can be assumed that in the initial period a large number of the required topics will still not be included in the "h_da topics" and are therefore available for tagging. In this case new topics can be suggested, which will be linked directly to the documents. These "new topics" can also be found in the topic hierarchy under "own topics" and must subsequently be embedded in "h_da topics". These new topic proposals mean that the new "h_da topics" will reflect the core knowledge of the university.

The content of foreign topics such as books or technical journal articles can be accessed via the relation, "related foreign topic", in case there are identical objects. The content of "foreign topics" is evaluable and available via this link (figure 8). "Foreign topics" objects are not shown in the layout through the definition of a shortcut relation, but through their linked content. The search in the corresponding web layout for "management" would only give one hit, this being the relevant topic from "h_da topics". This topic content provides all contents linked directly to this topic, and all other types of document, which are given as hits via shortcut relations (figure 9).

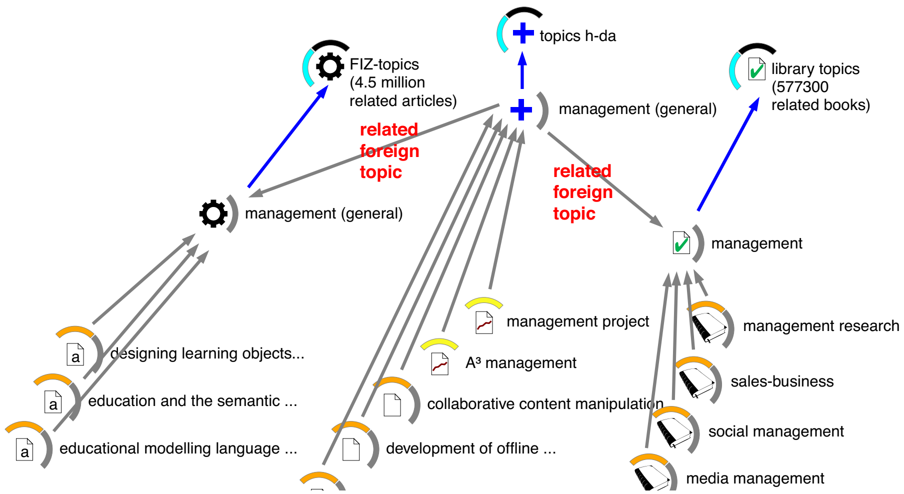

Figure 9: The topics under h_da topics are linked to foreign topics via the "related foreign topic" relation, in case identical objects exist.

6 Conclusions

The ideal way to quickly establish a topic network that meets the needs of both students and lecturers is to use existing ontologies, such as media wikis, library hierarchies and scientific ontologies, and to amalgamate them with one another for this purpose. These ontologies, which have arisen as a result of very different objectives, constitute the essential knowledge carriers of a university. However, the amalgamation process enables these original ontologies to remain consistent in themselves while being expandable. We call the resulting topic network the Next Generation Topic Map (NGTM), as it facilitates both independent data modelling and a continuous import with automatic linking to other ontologies. This allows the characteristics and features of the original ontologies to be maintained. The actual value of the new topic world is derived from linking the topics to lectures, projects, books, professional articles, thesis papers and much more. The key to this creation of value is a web-based semantic search, which can, for example, be offered directly on the university main page. The clustering of search results by document and object types ensures clarity. Such a system makes activities and expert knowledge transparent, to the great advantage of all university members.

References

[App05] W P Fok Apple. Ontology-driven Content Search for Personalized Education. Phd, City University of Hong Kong, Department of Computer Science, City University of Hong Kong, 2005.

[BB10] Roland Böving and Udo Bleimann. Semantic search scenarios to enhance University needs. In Steve Wheeler, editor, *Conference Proceedings of the 5th Plymouth e-Learning Conference – Learning without limits: Facing the Challenges*, Plymouth, UK, May 2010. University of Plymouth.

[Ble04] Udo Bleimann. Atlantis University – A New Pedagogical Approach Beyond E-Learning. *Campus-Wide Information Systems*, 21(5):191–195, 2004.

[BLHL01] Tim Berners-Lee, James Hendler, and Ora Lassila. The semantic web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, May 2001.

[BR06] Udo Bleimann and Franz Josef Röll. Extended Blended Learning – Innovative Learning Scenarios for Universities. In *INC 2006 Conference Proceedings*, pages 385–392, Plymouth, UK, July 2006.

[Mar08] Philippe Martin. Semantic Networks to Support Learning. In *International Conference on Conceptual Structures; Supplementary Proceedings of ICCS 2008*, Sophia-Antipolis, France, July 2008.

[PDM09] Niels Pinkwart, Darina Dicheva, and Riichiro Mizoguchi. Ontologies and Social Semantic Web for Intelligent Educational Systems. In *Conference Proceedings In conjunction with the 14th International Conference on Artificial Intelligence in Education*, Brighton, UK, July 2009.

[Ras05] Marcel Rasinger. Vergleich von RDF/RDF Schema und OWL. Website, 2005. Available online at http://alturl.com/58h9; visited on May 8th 2010.

[SLWD04] Demetrios Sampson, Miltiadis Lytras, Gerd Wagner, and Paloma Diaz. Ontologies and Semantic Web for E-Learning. *Educational Technology & Society*, 7(4):26–28, 2004.

[Wes07] Adam Westerski. *E-learning based on the Social Semantic Information Sources*, pages 172–186. Springer, 1 edition, 10 2007.