Establishing a Cooperative Digital Library for Teaching Materials - A Case Study

Florian Matthes, Ulrike Steffens
Software Systems Institute,
Technical University Hamburg-Harburg, Hamburg, Germany
{f.matthes,ul.steffens}@tu-harburg.de
www.sts.tu-harburg.de

Abstract

This paper reports on the experience gained in establishing a cooperative digital library for teaching materials in the area of computer science. Based on a detailed description of the content management and collaboration services relevant for such a digital library, an integrated conceptual model is developed. Furthermore, the paper gives an outline of the system's architecture, which makes use of widespread technologies and exploits mature off-the-shelves system components where possible. The paper ends with a report on the practical experience gained using the system at our academic institution and a discussion of future project perspectives expanding the use of the library across institutions and enterprises.

1 Motivation

The rapid advancement of information technology challenges computer science education [3]. For avoiding the isolation of the ivory tower and preparing graduates to meet the industry's growing demands regarding their skills and knowledge, an ever increasing amount of content has to be taught. The need for a high-quality collection of teaching materials benefitting from the possibilities of modern information systems comes along with this development. Computer science lecturers find themselves faced with the following situation:

- Teaching materials are subject to a reduced half-life. Content significant today can be obsolete tomorrow.
- The enlargement of the subject area enforces new foci in planning curricula and designing courses.

- The given possibilities of weaving multimedia effects into digital documents stimulate the demand for a high quality of teaching materials and increase the production costs.
- The increasing globalization of markets requires an increasing internationalization of education resulting in the need for multilingual teaching material.
- Teaching materials have to be unitized in an adequate manner in order to be linked together across the boundaries of different teaching scenarios and also of different subject areas [11].

The lion's share of computer science teaching materials is already stored in digital form. It comprises documents in a variety of formats representing textual or multimedia content, as slides, exercises, sample programs, or animations. Of course, archiving these documents within file systems or administering them within lists of web links does improve access, but this cannot meet all requirements described above. The solution lies in the development of advanced information systems which help to archive, to reuse, and to understand computer science materials and which finally enable an easy integration of new educational services as soon as they arise.

In this paper we present a digital library system for teaching materials. The system uses concept maps as organizing structures for the administered content. Concept maps can be represented in a straightforward datamodel (see section 2) which guarantees the system's extensibility. Section 3 describes a number of library services not only including options for retrieval and authoring but also comprising facilities for a personalized and cooperative use of the system. The system's web-based architecture, which makes use of widespread technologies and exploits mature off-the-shelves system components where possible, is discussed in section 4, before the paper closes with a number of conclusions drawn from our own experiences with employing the digital library system (see section 5).

2 Structuring Teaching Materials in Concept Maps

Structuring knowledge within *concept maps* in order to present or recall it is a prominent approach in education science [8]. We follow this point by employing a concept map as organizing structure for the digital library of teaching materials. In this way it is possible to split up teaching content into reusable modules of reasonable size and to later retrieve these modules in a semantically intuitive manner. This section briefly describes the data model behind the idea of a concept map depicted in figure 1.

A concept map consists of a number of *concepts* which form a taxonomy over the teaching content. There are two different kinds of concepts. A

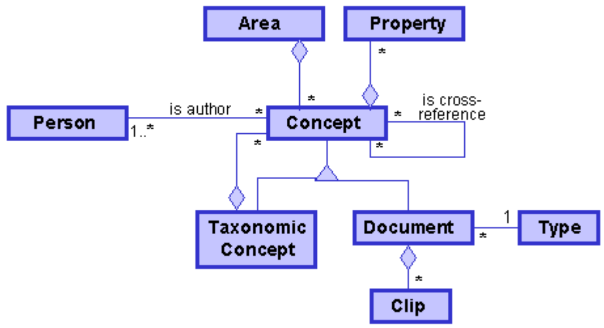


Figure 1: The class diagramm for concepts

concept representing and briefly explaining a notion within the subject area is called a *taxonomic concept*. A node named *"Functional Programming"* providing a short definition, for example, is considered a taxonomic concept. *Documents* storing supplementary relevant materials are also specific concepts. A PowerPoint-file containing a number of slides on functional programming, for example, is considered a document. Documents always have a well-defined document *type* and can be devided into a number of *clips* representing the smallest reasonably usable multimedia units within a document.

Concepts are linked together in a network. They can be related in two different ways. On the one hand, a concept can be crossreferenced by another. On the other hand, taxonomic concepts can have a number of subconcepts. In contrast, documents are regarded as leaves of the concept map with no further child concepts.

Each concept has a number of *properties* which contain metadata important for retrieval, cooperation, or personalization purposes. Examples for properties are the date of concept creation, a set of attached keywords or a set of user rankings. Moreover, the concepts in the map can be divided into subsets, which are called *areas*. Areas represent a further guidance system which can be accessed for navigation and retrieval. There is always at least one *person* who signs responsible as the author of a concept. This assignment is important with respect to feedback and correction processes within the library system (see section 3.4).

Not only in education science but also in computer science concept maps gain more an more popularity. We based the data model introduced here on the model described by the *Topic Map* standard [9], but omitted features like link types as we believe that they increase the price of archiving and retrieving without substantially improving the usability of a digital library.

3 Cooperative Use of a Concept Map of Teaching Materials

The members of a teaching institution approach the digital library system in two different roles. On the one hand, they act as *content providers* adding new or updating existing concepts. Their contributions usually originate in the various daily working tasks within the institution. Some of those tasks, such as preparing lectures or writing articles or books, on the other hand, do not only produce new content but also require a lot of input which might again be drawn from the digital library. A person approaching the system in order to find already existing material has the role of an *internal content user*. Internal content users collect useful parts and pieces from the library and other sources and join them within revised teaching materials which they present in their specific courses to a specific group of learners, the *external content users*.

The concept map structures all materials within the library and should therefore be agreed upon by all content providers and internal content users. A consent of this kind certainly cannot be reached in one single step, but it has to develop over time based upon discussions between the members of the teaching institution. As the concept map is a vital part of the cooperative working processes, it can be useful to install an *editorial board* consisting of a small number of persons having the exclusive responsibility for the map's manipulation.

Against the background of the aforementioned usage scenario, the digital library should not only provide services to store and retrieve concepts, but should also offer ways for the teaching staff to both communicate on the existing content as well as to create personal working contexts for collecting, reusing and extending concepts. An overall view on the library system's functionality is depicted in figure 2.

Each of the library's services falls into one of four different categories, which are determined by the user's intention when addressing the system:

- services for *knowledge research*, enabling internal content users to search or browse through the digital content in order to gain a personal understanding of certain, possibly linked concepts and an overview over material already existing within digital documents,

- services for *content acquisition*, enabling content providers to easily change or extend the concepts within the library on site without requiring a deeper system knowledge,

- services for *personalized knowledge exploitation*, enabling content users to reuse and adapt the library's content for their own working purposes,

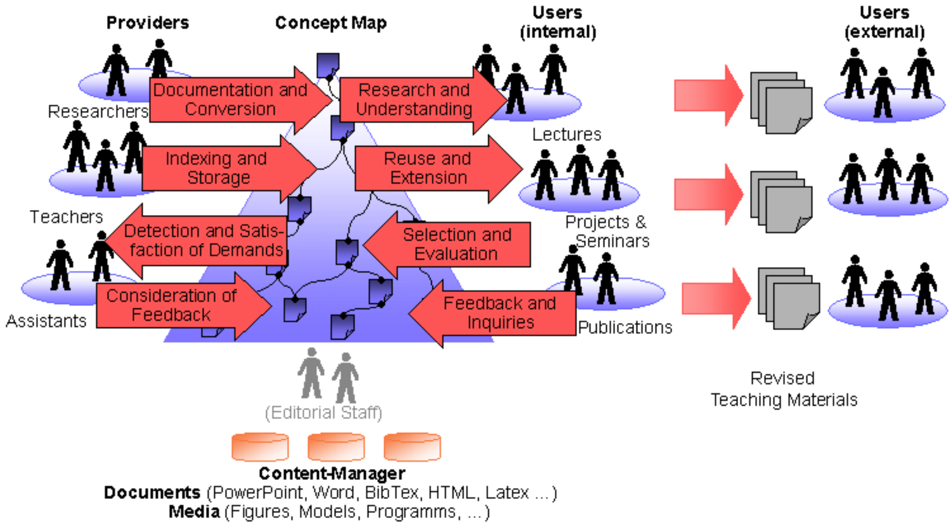


Figure 2: Cooperation in the Digital Library of Teaching Materials

- services for *cooperative content evolution* enabling feedback and discussions about the concept map.

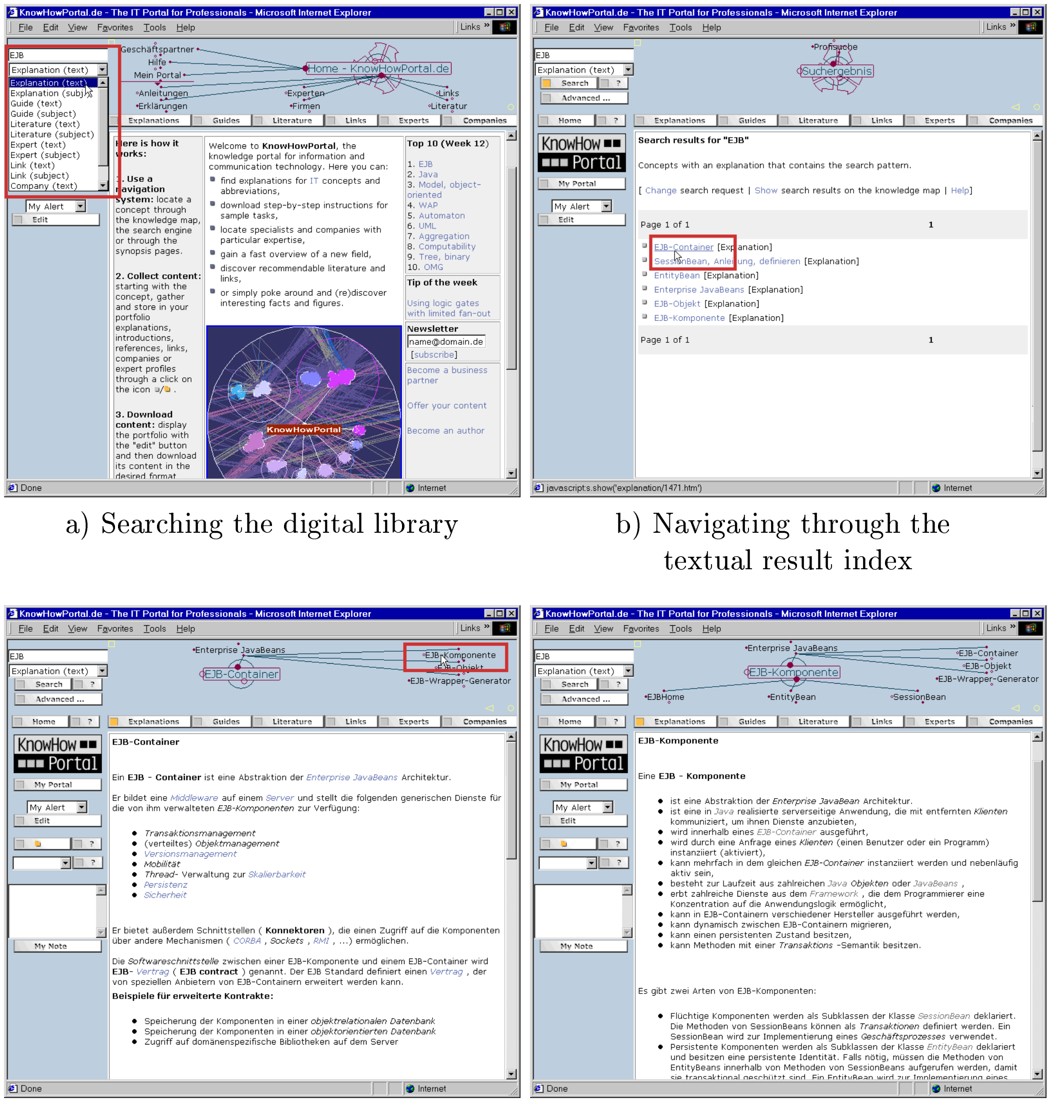
The choice of individual services for a digital library of teaching material depends on the respective working context. Therefore, the system introduced relies on a component based architecture, enabling the easy integration of new services.

3.1 Research and Understanding

The digital library of teaching materials offers services for both, searching particular concepts, as well as browsing through the concept map step by step, exploring different concepts and gaining an overview of the content involved. Figure 3 gives an impression of the user interface, which allows the use of searching and browsing facilities from anywhere in the concept map.

Apart from a service for fulltext retrieval, there is also a possibility for explicitly searching for concepts within the concept map (see Fig. 3 a)). Furthermore, the user can employ an advanced search service on concept metadata, as for example "author", "date", or "format". The set of stored metadata elements is generic and can be extended or adapted in behalf of the specific application context [5]. All kinds of search can be executed either within the whole library or within a specific area.

Navigation is supported by textual as well as graphical services. A textual navigation service is available through a number of indexes over the included concepts, ordered for example chronologically, alphabetically or by



a) Searching the digital library

b) Navigating through the textual result index

c) Using the concept navigator

d) Viewing a digital document

Figure 3: Searching and browsing services

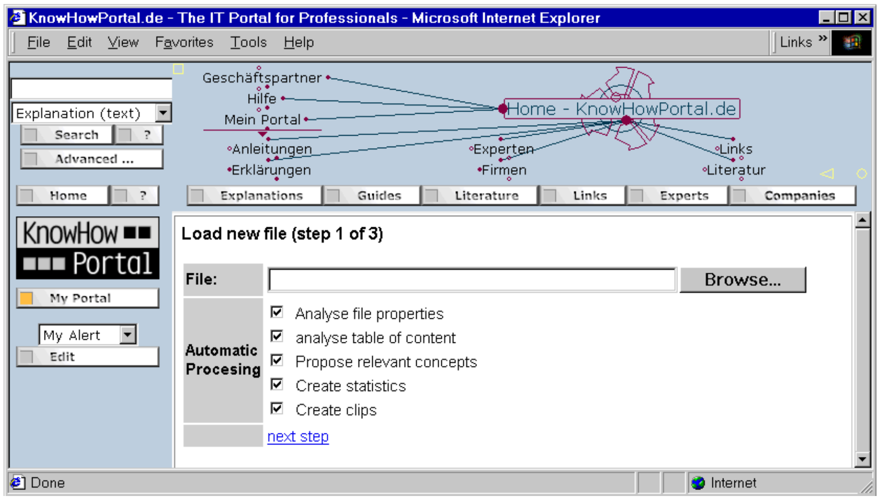


Figure 4: Importing a new document

area. Evidently, all concepts are reachable directly from the indexes via hyperlinks (see Fig. 3 b)).

In addition to these text indexes, the broker offers two ways of graphical navigation through the library. The *map overview service* provides the ability to zoom in and out of the concept map, providing an overview over all teaching materials, over selected areas, or over subareas. Using this service helps internal content users to obtain a general understanding of what kind and what number of concepts can be retrieved from the library.

In contrast to the bird's-eye view of the overview service, a graphical *concept navigator* offers a hiker's view for a specific concept currently visited (see Fig. 3 c)). This concept is always placed in the navigator's center surrounded by its direct neighbours within the map. Following the graphical representation of the links between the concepts, the user can explore the concept's neighbourhood and trace the semantic connections represented.

3.2 Content Acquisition

Archiving new teaching materials within the digital library means extra work beyond the simple creation of the respective digital files. Nevertheless, discipline and insistence on the part of the content provider determine the library's quality. Hence, inserting new concepts should be possible with least effort.

In order to import new digital teaching materials, a content provider does not have to do anything more than to enter the respective directory path into the web input interface of the library system (see Fig. 4). Thereby he is offered a number of automatic facilities for the semantic processing of the documents to be included. The system is able to automatically update

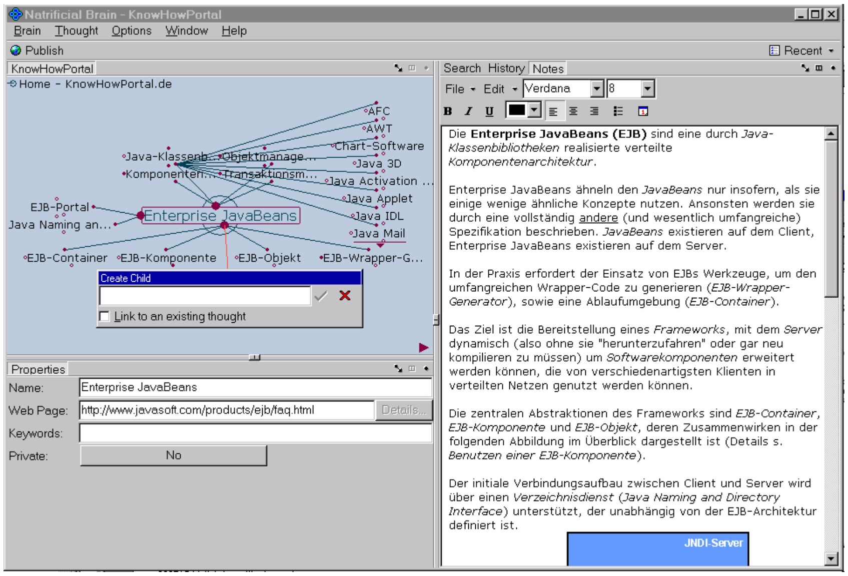


Figure 5: The graphical concept map editor

the retrieval indices that underlie the search services and to augment the library's content by creating possible hypertext links between new and existing concepts. Since we believe that a content provider usually has a deeper understanding of how his own materials suit into the subject area than an information system could have, the classification of a new concept is done semi-automatically. The system generates some suggestions on the area into which the new concept could fit and on what kind of concept links could be added, but leaves the actual classification with the content provider.

For reasons of convenience, the web input interface is supplemented by a mail import service and an FTP-monitor via which new or updated concepts may also be loaded. For obtaining a first visual feedback on his submission, a content provider can turn to a preview service which gives him an idea of how his newly-generated concept will look like in the web.

Especially in large teaching institutions, the quality of a digital library might profit from the fact that not all members of the institution are able to update arbitrary parts of the content. Therefore, the library system allows to equip different individuals or groups of persons with diverse sets of access rights. The group of people which has write access to the whole library together with its underlying concept map can then be regarded as the library's editorial board. In order to support the editorial board the library system is backed by a content management system, which strictly differentiates between the update and the publication of a concept. In this way, a number of content providers can make tentative updates, whereas an editor is responsible for the publication of those updates within the library.

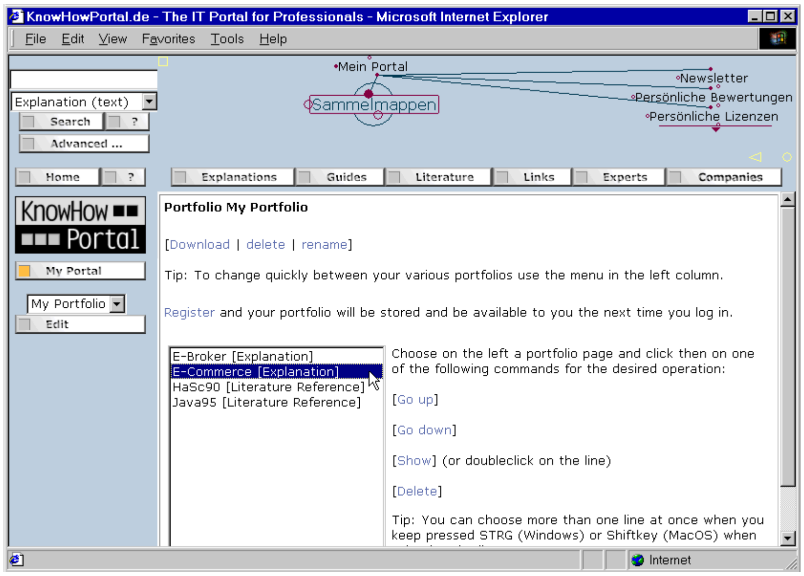


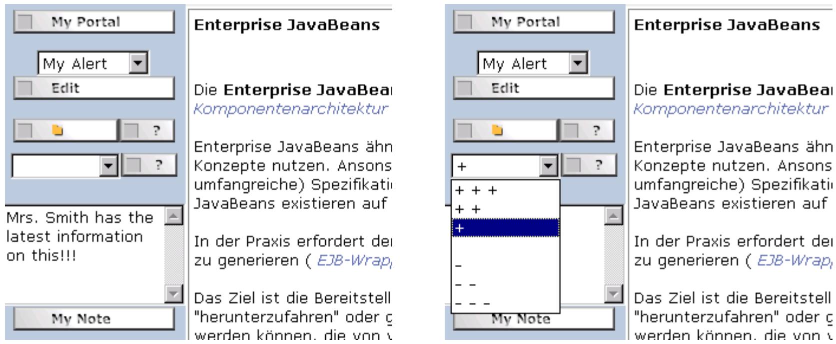
Figure 6: A personal portfolio

In most cases, the editorial board is also responsible for the management of the concept map underlying the digital library. For this task a graphical concept map editor enabling the easy manipulation of links and nodes within the concept map is at disposal (see Fig. 5). The interface of this service is very similar to the concept navigator's interface, but also allows to create or delete concepts within the map or to manipulate the concept links.

3.3 Personalized Knowledge Exploitation

Internal content users approach the digital library to explore the existing teaching materials for fragments which they can reuse within their own teaching context. The members of the teaching institution have the possibility to create an arbitrary number of personal portfolios, in which they can collect individually relevant concepts. For this reason, all concepts, whether they are directly visited or they appear in index lists, search results, etc., are equiped with a uniform icon to be clicked, if the user wants to include the respective piece of information into his portfolio. The user can view these portfolios whenever he likes to and then has the opportunity to delete or re-arrange the contained items, but also to directly revisit them via hyperlinks (see Fig. 6). Of course, the library system stores portfolios over multiple sessions.

Once the content user has included some concepts within a portfolio, he has the possibility to download all or some of them, automatically converting them into his preferred document format. The library system offers document generators for several popular formats, among them e.g. HTML,



a) Making an annotation b) Making a personal judgement

Figure 7: Making personal remarks

MS Word, MS PowerPoint, or PDF.

While members of the teaching staff inspect the library's content, new aspects, comments, questions etc. might come to their minds. It is straightforward to connect these personal ideas to the respective concepts directly within the library system which in this case functions as a personal working environment for its users. Therefore, each user has the opportunity to add both annotations as well as personal judgements to arbitrary concepts (see Fig. 7). These individual remarks are not visible to the public, but to the respective user whenever he visits one of the items in question. He may change or delete them whenever he thinks it is appropriate. Downloading concepts from a portfolio also includes downloading personal metadata, as they serve as an important basis for a potential reuse and extension.

Apart from adding their personal remarks to a concept, content users can also pass on their comments and judgements to the rest of the institution (see section 3.4). The judgements are then collected by a recommendation engine [1] generating personalized recommendations for users who subscribe to this service.

To keep up with the ongoing changes on the concepts they are concerned with, content users can also rely on the library system which offers them a number of subscription services. A general newsletter reports on changes in either the whole library or in some of its areas and is distributed to its subscribers via e-mail at regular intervals. To achieve an even finer granularity, the user can put single concepts on which he wants to be kept informed into a special alert portfolio. If one of the included items changes, the user receives a notification (see e.g. [10]).

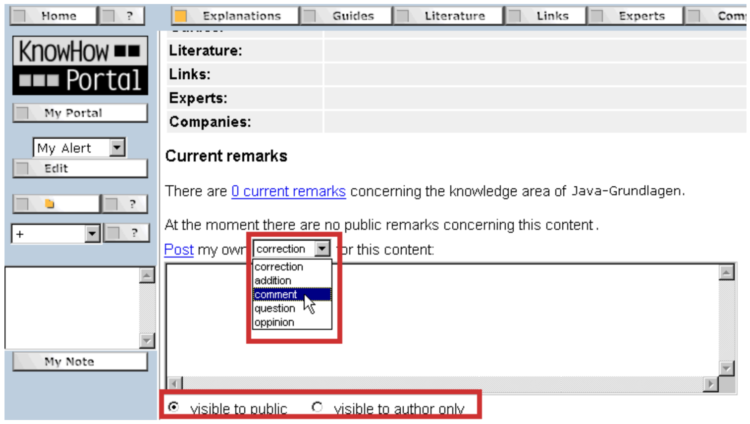


Figure 8: Shared Remarks

3.4 Cooperative Content Evolution

The library system provides different feedback services which can also be regarded as communication channels between readers and authors of concepts or among particular interest groups within the teaching institution.

Users can share their personal remarks and judgements on concepts by making them public for the other members of the teaching institution (see Fig. 8). For the sake of readability and further computations, remarks are classified as *"opinions"*, *"questions"*, *"corrections"* etc. Judgements furthermore serve as input for an underlying recommendation engine in order to generate user-related recommendations. Both, remarks and judgements, can be analyzed statistically and therefore serve as basis for searches like *"Concepts with a lot of questions on them"* or *'Concepts with no corrections on them"*.

The library system administers author information for each concept. This fact enables users to confidentially pass their remarks and judgements directly to a content provider obscuring them from the public. The content provider then has the possibility to answer by e-mail and to adapt his concept according to a received criticism or question.

4 System Architecture

Figure 9 gives an overview on the library system's architecture. The system is composed of two servers. On the one side, the internal content management server is concerned with the administration of the concept map and the attached digital documents. On the other side, an external content brokerage server deals with the visualization of the existing data as well as with the interaction with content users.

On the side of the content management server a production engine pro-

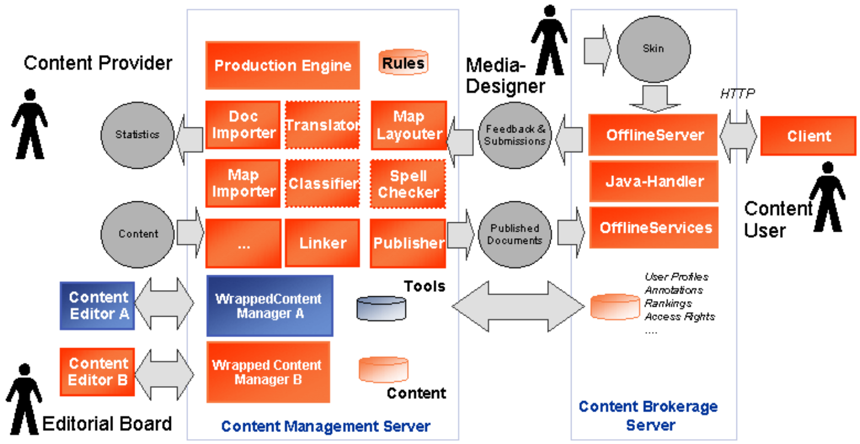


Figure 9: The architecture of the digital library for teaching materials

cesses the incoming content. It makes use of numerous components in order to produce a concept map and digital documents which later can be published over a web interface. These tools realize among other things the visualization of different views on the concept map, the automatic classification of concepts, and the automatic creation of hyperlinks. Where possible, mature off-the-shelves system components are exploited instead of reimplementing services from scratch. The external representation of the concept map, for example, is provided by *TheBrain's* Java-based tools for information management [2].

The same strategy is applied for content management, where a commercial system is in use. This system differentiates between the editing and publication of concepts and thus enables the editorial board to check the quality of concepts before they are finally published over the external content brokerage server. Furthermore, the content management system is able to administer access rights for different subsets of the content and therefore fosters a task sharing between the editors.

The content brokerage server equips published concepts with the library's specific look and feel making use of a skin of templates. The concepts finally presented to the content users over a web client are HTML-documents enriched with JavaScript function calls for the library operations offered by the server. The content brokerage server is responsible for the whole interaction between the library system and the content users. It therefore stores all user related information like profiles, annotations, or access rights and passes user feedback and online submissions of new concepts back to the content management server. Apart from approaching the library system over the Internet it is also possible to run it offline, for example from CD-ROM, by the help of an integrated offline server.

5 Conclusions

Employing the presented digital library system for the adminsitration of our own teaching materials, we can draw a number of conclusions, which on the one hand show the advantages currently drawn from using the system and which on the other hand point at the possible directions for future work.

Presenting teaching materials within a concept map seems to be intuitive to the great majority of library users. The modularity of the content helps to retrieve and reuse required pieces of information quickly, but beyond it supports the understanding of the single modules. Making use of the system directly within our lectures by presenting chosen paths through the concept map, we learnt that students attentively follow the lecturer's train of thoughts. That is why we decided to not only open the library to the staff of our institution but also to our students as an offline version on CD-ROM which they now employ to recapitulate teaching content in order to prepare for the respective exams.

This example shows already, that the actual use of the library system soon changed its nature from the simple task of archiving and retrieving teaching materials to working with pieces of knowledge in a number of very different scenarios. A number of underlying business models seem to feasible in this respect. In a next project step, we plan to open the system for other universities in order to establish a common pool of teaching materials which profits from the different experiences of all the persons involved and eases the creation of new documents. Furthermore, computer science content nowadays also represents an important asset outside the academic life. Currently, we are developing a business model which enables us to circulate parts of our digital library to companies against payment or other kinds of reward.

All these extensions are supported by the simple data model underlying the digital library system (see section 2). The notion of linked documents and concepts can be found in most of the potential applications and services. Various fields of research connected with digital libraries [6] or teaching methods [4] produce metadata, which are assembled within the concepts' extensible set of properties turning the library into a flexible system adaptable to the different business models described above. Moreover, administering persons as objects within the system opens the library for personalized and cooperative services. At present, we are for example extending the system by a component for online self-assessment and by a cooperative service making use of aging links [7].

Acknowledgements: The research described in this paper was supported by the DFG Project KOLIBRI (DFG Schm 450/7-1 and DFG MA 2005/1- 2).

References

[1] V. Bouthors and O. Dedieu. Pharos, a Collaborative Infrastructure for Web Knowledge Sharing. In S. Abiteboul and A. Vercoustre, editors, *Proceedings of the Third European Conference on Research and Advanced Technology for Digital Libraries, Paris, ECDL'99*, September 1999.

[2] TheBrain.com. http://www.TheBrain.com, 2000.

[3] M. Broy and J.W. Schmidt. Informatik: Grundlagenwissenschaft oder Ingenieurdisziplin? *Informatik-Spektrum*, 22(3):206–209, June 1999.

[4] IMS. The IMS Metadata Specification. http://www.imsproject.org/metadata, 2000.

[5] F. Matthes and U. Steffens. PIA - A Generic Model and System for Interactive Product and Service Catalogs. In S. Abiteboul and A. Vercoustre, editors, *Proceedings of the Third European Conference on Research and Advanced Technology for Digital Libraries, Paris, ECDL'99*, September 1999.

[6] J. Milstead and S. Feldman. Metadata: Projects and Standards. *Online*, pages 33–40, January/February 1999.

[7] C. Niederée, U. Steffens, J.W. Schmidt, and F. Matthes. Aging Links. In *Proceedings of the Fourth European Conference on Research and Advanced Technology for Digital Libraries, Lisbon, ECDL'00*, September 2000. To appear.

[8] J.D. Novak and D.B. Gowin. *Learning how to learn.* Cambridge University Press, Cambridge, 1984.

[9] S. Pepper. Navigating Haystacks and Discovering Needles: Introducing the New Topic Map Standard. *Markup Languages: Theory and Practice*, 1(4):47–74, 1999.

[10] U. Rauschenbach. Supporting Awareness in Shared Workspaces Using Relevance-dependent Event Notifications. In *Proceedings of the CVE'96 Workshop, Nottingham*, September 1996.

[11] J. Schlichter and G. Teege. Web-based Information Management for University Education. In J. Multisilta and J.-P. Niemi, editors, *Proceedings of The LeTTeT 98 and MaTILDA98 Joint Conference*, pages 99–106, Pori, Finland, May 1998.