Enhancing Information Management for Digital Learners

Stefanie Sieber
Media Informatics Group, University of Bamberg
Feldkirchenstrasse 21, D-96052 Bamberg, Germany
www.uni-bamberg.de/minf/
stefanie.sieber@uni-bamberg.de

Abstract

The ongoing digitalisation of the learning process not only brings many advantages for learners but also poses new challenges like the increasing amount of information that need to be faced. This paper conceptually sketches a system employing and combining different existing Information Retrieval (IR) techniques to support digital learners. The system establishes a personal learning space including all learning content by building an appropriate index of learning material. In order to store learning content and allow an additional classification, adjusted to the specific needs of a single user, two supplementary index levels are created and connected to the primary index level. Different maintenance and search facilities allow an intensive use of this individual structure. Summarising, a personalised learning service, meeting the challenges of modern learning, is delivered by extended searching capabilities.

Keywords: personal search, personal information management

1. INTRODUCTION & MOTIVATION

The nature of learning has changed over the last decades, especially because computers and the Internet are, by now, quite naturally integrated into a typical learning environment. This technical progress and the consequential integration cause a shift from traditional learning to learning concepts located in a continuum of blended learning as defined by Jones [1]. Furthermore, not only the preparation of content itself but also the way of distributing learning material and communication in a learning environment are in the process of being completely changed. This digitalisation of the learning process on the one hand, but also the nowadays common assumption of and awareness for lifelong learning as discussed in [2], bear new challenges, which have to be faced by learners as well as lecturers.

On the one hand, Learning Management Systems (LMS) (cf. [3]) are, in the majority of cases, employed to provide a foundation for modern learning. Whether it comes to distance learning or not, all required material, including additional information for further reading and tasks controlling the learner's success, is provided online, such that learners can easily access all desired information any time and any place. On the other hand, advanced technical possibilities also allow an integration of multimedia content into learning and teaching concepts, as for example suggested in [4]. Furthermore, regarding the concept of lifelong learning, learning is ubiquitous. Newspapers or journals are more often read on the Internet and various information is also checked on the Internet—on stationary as well as mobile devices.

Of course, this development is pleasant and softens some difficulties that usually need to be faced by learners. The general availability of information, for example, undoubtedly increased due to this technical progress. However, for this very reason, composing the big picture is more difficult, also because skimming through collected information is not satisfactorily supported by now. Moreover, the rising number of different kinds of content that needs to be considered is even more calling for an appropriate support of collecting and arranging different kinds of learning content. Besides, there is also the fact that most of the systems used so far have an institutional point of view rather

than the perspective of a single learner.

Put another way, personal and professional experiences with LMS (cf. [5]) and their functionalities, often far from being satisfactory, actually lead to proposing our system as described within this paper.

Briefly spoken, the vision of the presented approach lies in the provision of an environment, allowing a single user to administer his learning material and learning context in order to facilitate learning insights. Hence, the concept for our future system is trying to build a *personal learning space* including all objects somehow related to the learner and the qualification he is trying to attain. Existing IR techniques are employed and combined in order to smoothly integrate the system into the learning process of a single learner.

The rest of the paper is structured as follows: Section 2 provides a brief overview on related concepts that need to be considered when actually designing the system. Focus of attention is on the presentation of the system concept in section 3. Within this section, the three design parts, that is the main ideas behind the system components for the collection, preparation, and presentation of learning objects are described. Finally, a short outlook in section 4, providing insights into future work, concludes this paper.

2. RELATED WORK

Of course, the promotion of a personal learning space immediately suggests personal learning environments (PLEs). By general definition, a PLE is a system helping learners to control and manage their own learning: "A PLE is a single user's e-learning system that provides access to a variety of learning resources, and that may provide access to learners and teachers who use other PLEs and/or VLEs [Virtual Learning Environments]." [6]

In order to allow a classification, Sandra Schaffert and Wolf Hilzensauer [7] identified seven crucial aspects for personal learning environments: the role of the learner as active, self-directed creator of content, personalisation, content without limitations, social involvement, ownership of learner's data, educational & organisational culture, and technological aspects. Evaluating these criteria, most of the stated aims of the proposed system coincide, such that our system can also be classified as PLE. However, PLEs often focus on examining and actively controlling the learning process and progress, which is particularly not the main objective of our system.

Mostly, PLEs have their origins in the demand for learner-centred approaches. Forms and specific definitions of PLEs range from sets of tools, each supporting a single asset of learning, as for example examined by Dalsgaard [8], to sophisticated systems, such as those briefly described by van Harmelen [6].

Besides some current research projects like TENCompetence¹ or MATURE², there are, of course, several existing systems already dealing with a more general view of learning than a typical PLE. However, as far as known, there is no existing system trying to smooth the process of assembling all information related to the learning process as well as fostering new insights by providing, first, assistance for easily setting up a personal learning space, second, a sufficient search and possibilities to efficiently glance through learning material and, third, visualisations for the two previously named tasks, showing the big picture by pointing out connections of single information pieces.

Concerning the technical implementation different research areas need to be considered. First of all, the proposed system follows a user-oriented approach to IR. Therefore, developments and findings of this area, as for example described in [9], are the relevant for our system. Secondly, user interaction is an important principle of user-centred systems in general and, of course, learning environments in particular. Relying on theories of constructivism (cf. [10]) interaction is a basic principle for learning, therefore interactive IR, examined in detail by Xie [11], needs to

¹<http://www.tencompetence.org>

²<http://mature-ip.eu>

be particularly considered. Thirdly, the idea of working with single information pieces, called learning objects, is not new. While generally working with learning objects, ideas, thoughts, problems, and solutions as for example discussed in [12] need to be taken into account. Also, storing learning objects suggests digital libraries like examined in the DELOS project³.

Moreover, since different existing IR techniques are supposed to be combined, actual research of the specific area needs to be considered as well. Selected references are directly included in the description of section 3.

3. SYSTEM CONCEPT

The future system, trying to accomplish the vision described above, is now described in more detail. By now, the system has not been realised in full depth and detail, for this reason descriptions need to be understood as conceptual sketches. The system will be divided into three core components—a collection component, a preparation component, and a presentation component (cf. figure 1). Each component is depicted in a separate subsection.

The foundation and content of the system are *learning objects*. Basically, a learning object is any possible physical representation of information, such as locally saved files in different formats or an online web page, related to the learning process. These learning objects are assembled and traditionally stored in a database by the *collection component*. Subsequently, the *preparation component* processes the assembled objects and extracts or assigns additional information. The presentation component, in contrast, represents the interface enabling user interaction and passing results along to the user.

FIGURE 1: User Interaction and Interplay of System Components

Concerning the actual realization, in summary, the system is ought to adopt familiar Web2.0-paradigms such as a rich and user friendly interfaces and tagging of information assets in order to facilitate intensive usage. Additionally, the system is meant to smoothly operate in the background for most of the time by default; variations, adjustments and a more intensive usage are of course fostered for active users.

3.1. COLLECTION COMPONENT

The basic component of the components triad is the *collection component*. This component is employed to actually build the personal learning space by assembling local as well as web content by an automatic, though controlled approach in contrast to crawling the Internet in general. Technically spoken, this component is based on a combination and extension of existing persistence and search frameworks such as Hibernate⁴ and Lucene⁵. At the moment possible compositions are examined, since there are already different existing approaches and projects like Hibernate Search⁶—integrating Lucene into the Hibernate framework—or the Compass Project⁷—allowing a flexibel combination of Lucene and different Object-Relational Mapping frameworks as well as supporting Spring⁸ integration.

³<http://www.delos.info>

⁴<https://www.hibernate.org/>

⁵<http://lucene.apache.org/>

⁶<https://www.hibernate.org/410.html>

⁷<http://www.compass-project.org/>

⁸<http://www.springsource.org/>

FIGURE 2: Levels of Indexing connected with Learning Objects

As core feature this component is not just extracting and processing textual information explicitly stored in *learning objects*, but generating three different index levels, building upon the database of learning objects, each of them organised in a specific hierarchy. This concept is related to the area of Hypertext IR (cf. [13]). The three levels—*learning material*, *learning context*, and *reference*—are loosely connected and organised in two tiers, as shown in figure 2. All different index steps can be triggered automatically for real-time indexing or manually for time-shifted adding of learning objects.

Since the technical granularity of *learning objects* is more or less arbitrary and learning objects are—at best—only structured by a manually created file structure; there is no existing structure that can be build upon. Hence, the index level *learning material*, respectively the first tier, includes all learning content but abstracts from learning objects. In this manner, a hierarchical structure linking to learning objects, however allowing file-independent connections of single learning assets, is formed. Furthermore, the supplementary index levels *learning context* and *reference* in the second tier are referencing to this particular structure. Put in other words, this level is used to abstract from physical data formats and build a well-defined hierarchy of learning material to allow precise references.

Technically, this level is similar to known indexing components. Text-based content is processed according to traditional full-text indexing mechanisms and an appropriate index is built. To successfully implement this index level as system foundation, related areas, such as XML Retrieval described by Kamps et. al [14], need to be considered.

Subsequently, the two supplementary index levels need to be populated. Therefore the collection component is trying to extract and store the *learning context* for a particular learning asset. Depending on the learner's situation, the learning material is ought to be classified as relevant, for example for a particular course or subject. Determining the learning context completely automatically is not meant to be achieved. Especially in the beginning, the detection will require additional manual effort of the user. However, preferences are supposed to be employed to structure and minimise manual input. It seems, for instance, applicable to use actual course websites for each semester in LMS or similar systems to determine the context of learning assets invoked within a certain radius from these websites.

Independent from a certain, determined learning context, learning material can be further classified by using arbitrary tags. This possibility is represented by the third index level *reference*. As mentioned, adopting the Web2.0-philosophy, the user is allowed to freely assign tags to learning material using the reference level. These tags can, if desired, be organised in one or more hierarchies in order to depict implicit structures. Here, too, manual setup and maintenance are unpreventable for a successful realisation. The fact, that this level explicitly asks for the user's input, is meant to be attenuated by offering Web2.0-like manipulation options. Additionally, the

FIGURE 3: Reasoning across different Levels of Indexing

process is ought to be smoothened and supported by automatically extracting and assigning keywords like headings or captions from learning objects.

Using the three index levels organised in two tiers, as described, instead of a single index level or a database index, fosters personalization and learning insights by allowing an individual structure. As a result, the design of the collection component is closely related to faceted search and needs to consider issues of this area, as for instance examined by Henrich and Eckstein [15].

Finally, the collection component is completed by the design of an active learning mode. This mode basically activates automatic indexing. Enabling this active learning mode, all visited web pages passing a filter—validating criteria like file format or domain names—are added if a certain time threshold is exceeded. Furthermore, maintaining a white list of web pages, like course pages, that should automatically be revisited and checked for updates periodically is an absolute necessity in our opinion. These examples are just two possibilities for user defined filters, which should of course be covered by system functionalities. For local content the organisation is considerably simpler. Besides adding single files to the repository, the user should, of course, be allowed to add directories to be observed and periodically checked for updates. In this manner, it can be ensured that the learner is the active, self-directed creator of content. Regular learning actions—consuming as well as creating learning-relevant content—are employed to automatically collect the content of the personal learning space.

3.2. PREPARATION COMPONENT

Building upon the foundation created by the collection component, the *preparation component* processes and enhances the information collected and basically structured in the previous stage. The main focus is to use the three index levels for a reasoning mechanism, adopting spreading activation techniques, such as described by Crestani [16].

As depicted above, every information in the two supplementary levels of the second tier is connected to a particular node in the first tier. Hence, spreading activation models and theories can be employed to detect possible connections between single nodes of learning material that have not been obvious before. That way, learning objects not explicitly searched, however considered relevant through connections spread across the different index levels, can be recommended to the user. Figure 3 shows reasoning using connections spread across the three different indexing levels. Correlations can, for example, be used to add novel connections between learning material and tags of the third index level. The careful configuration of the reasoning component will be subject of further research and inquiries.

The preparation component, obviously, is also one possible starting point for collaboration scenarios incorporating social involvement. Especially because Web2.0 philosophy is employed and techniques like tagging are integrated, it is likely to avoid performing equal tasks by multiple users several times. The tagging of a particular learning material, for example, does not need to be computed several times but can be shared with other learners, using the same learning

objects. Of course, complete ownership of learner's data needs to be retained. Elaboration of feasible implementations will also be subject to further research.

3.3. PRESENTATION COMPONENT

Finally, the *presentation component* primarily forms the investigation interface offering the service as a whole to the user. The presentation component also enables user interaction for the two previously described components by allowing manual input, collection and learning-related preparation. Direct adding, accessibility, modification, and extension of learning objects or learning material of the first tier are also provided by the interface.

For this component, it is especially important to allow a lightweight and intuitive utilisation. Principles of exploratory search, as for example described by Marchionini [17], need to be applied in order to satisfy traditional search activities like fact retrieval but in particular the continuative search activities *learn* and *investigate*.

4. CONCLUSION AND OUTLOOK

In summary, our proposed system is supposed to support the learning process of a single learner, also allowing optional collaboration with other learners if desired by employing and combining IR techniques. One of the research challenges lies in carefully selecting and combining existing techniques in order to benefit from related research areas.

Basic searching is constituted by keyword search. Of course, an advanced search, allowing to use the three index levels as filters to narrow the search down to certain criteria, also needs to be part of the described system. Among others, it is very likely that parameters like the actual learning context—determined due to the web page just visited—can be employed to find related learning objects not directly included, but assigned to the search context as well.

Moreover, different visualisation components are supposed to be embedded into the search interface for additional presentation of various aspects. Search results can be visualised by tag clouds for textual information. The collection of learning objects can—based on a particular search or not—be visualised using associations of different levels, allowing browsing the repository on a visual level.

REFERENCES

- [1] Jones, N. (2006) *E-college wales, a case study of blended learning*. In Bonk, C.J. and Graham, C.R.: The handbook of blended learning: Global perspectives, local designs. 182–194. Pfeiffer.
- [2] Field, J. (2006) *Lifelong learning and the new educational order* (2nd edn). Trentham.
- [3] McGee, P., Carmean, C., Ali, J. (2005) *Course Management Systems for Learning: Beyond Accidental Pedagogy*. IGI Publishing.
- [4] Henrich, A., Sieber, S. (2009) *Blended learning and pure e-learning concepts for information retrieval: experiences and future directions*. Information Retrieval, 12(2), 117–147.
- [5] Henrich, A., Wolf, S.U., Sieber, S. (2008) *Evaluation, analysis, and future issues of a university-wide learning management system concluding a two-year initiation phase*. In Hambach, S., Martens, A., Urban, B.: e-Learning Baltics 2008: Proceedings of the 1st International eLBa Science Conference in Rostock, Germany, June 18-19, 2008. Fraunhofer IRB Verl., Stuttgart.
- [6] van Harmelen, M. (2006) *Personal learning environments*. Conference on Advanced Learning Technologies, 815–816.
- [7] Schaffert, S., Hilzensauer, W. (2008) *On the way towards personal learning environments: Seven crucial aspects*. eLearning Papers, (9).
- [8] Dalsgaard, C. (2006) *Social software: E-learning beyond learning management systems*. European Journal of Open, Distance and E-Learning, (2).

- [9] Xie, I. (2008) *User-Oriented IR Research Approaches*. In Xie, I.: *Interactive Information Retrieval in Digital Environments*. 1–28. IGI Global.
- [10] Twomey Fosnot, C. (2005) *Constructivism. Theory, Perspectives, and Practice* (2nd edn). Teachers College Pr.
- [11] Xie, I. (2008) *Interactive Information Retrieval in Digital Environments*. IGI Global.
- [12] Littlejohn, A. (2003) *Reusing Online Resources: A Sustainable Approach to E-Learning*. Routledge Falmer.
- [13] Agosti, M. and Smeaton, A.F. (1996) *Information Retrieval and Hypertext*. Kluwer Academic Publishers.
- [14] Kamps, J., Marx, M., Rijke, M., Sigurbjörnsson, B. (2003) *Xml retrieval: what to retrieve?* In SIGIR '03: Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval, 409–410, ACM.
- [15] Henrich, A. and Eckstein, R. (2008) *An integrated context model for the product development domain and its implications on design reuse*. In Marjanovic, D., Storga, M., Pavkovic, N. and Nojcetic, N.: 10th International design conference: DESIGN 2008, 761–768.
- [16] Crestani, F. (1997) *Application of spreading activation techniques in information retrieval*. Artificial Intelligence Review, 11(6), 453–482.
- [17] Marchionini, G. (2006) *Exploratory search: from finding to understanding*. Communications of the ACM, 49(4), 41–46.