Positioning of Learners in Learning Networks with Content, Metadata and Ontologies

Open Universiteit
www.ou.nl

Positioning of Learners in Learning Networks with Content, Metadata and Ontologies

M. Kalz[*], J. van Bruggen, E. Rusman, B. Giesbers & R. Koper[1]

[1]*Open University of the Netherlands, The Netherlands*

Positioning in learning networks is a process that assists learners in finding a starting point and an efficient route through the network that will foster competence building. In the past we explored computational approaches to positioning that are based on the contents of the learning network and the behavior of those participating in it, more or less ignoring different efforts to stimulate positioning and competence development from a top-down-perspective. In this paper we introduce a research agenda for positioning in learning networks, discuss several cases and give an outlook on the development of a positioning service for learning networks.

Introduction

Technology enhanced lifelong learning promises learners the possibility to learn and build competencies in every context and every phase of their life. To meet this promise the individual should stand in the centre of every effort in lifelong learning instead of institutions and organizations. The concept of learning networks offers a framework to bridge the different distributed parts of current technology enhanced lifelong learning (Koper, Rusman & Sloep 2005). A learning network connects actors, humans as well as agents, institutions and learning resources which are organized in competence development programs. Information and communication technologies are used in such a way that the network self-organizes. The actors in the learning network share one common goal: furthering the development of competence by learners. A common approach to overcome the limitations of institutional dependencies in is the concept of Accreditation or Recognition of Prior Learning (APL/RPL) (Merryfield, McIntyre & Osaigbovo 2000). APL offers methods and techniques to identify prior learning experiences from formal and informal education. This procedure is especially important if a person crosses the boundaries between work and learning or between academic disciplines. Most of the methods for APL rely on experts who study the learners' profiles and decide which parts of educational programs could be exempted and which ones are best suited as starting point for the students. However, this way of analyzing prior learning experiences is a very time-consuming and expensive method (Skinner 2005). We propose therefore as an alternative the usage of computational approaches to address this problem for lifelong learning in learning networks. Previous work at the Open University of the Netherlands focused on content-based approaches to address this problem in learning networks (van Bruggen, Rusman, Giesbers & Koper 2006). This article widens the focus to metadata and ontologies and provides a research agenda for an ongoing project that aims at the research and development of a web service for learner positioning in learning networks .

---

*Email: marco.kalz@ou.nl

Positioning in learning networks

Positioning in learning networks has to take into account various forms of learning, including non-formal learning. No matter if the competence development programs are formal or informal, learners engage in series of learning activities that may take a long time to complete. In learning networks for lifelong learning, prolonged interruptions of such series of learning activities are likely to occur. Moreover, learners may engage intermittently in different types of learning. Whenever such a learner enters or returns to a learning network we are faced with what we call the 'positioning problem': Taking into account the goals and the history of the learner, what route or routes of learning activities through the learning network can we advise and what is the best place for the learner to start (van Bruggen et al. 2004)? Positioning means to compare the already acquired (levels of) competencies of a learner to the (levels of) competencies that result from a particular competence development program in the current learning network. Assume that this learning network contains pre-arranged routes towards particular goals and that every route is a competence development program. Then, the positioning problem is one of determining which learning activities in the routes need to be completed and which ones can be skipped, because they do not add to the competencies, skills and knowledge that the learner has acquired in the past. How exactly the competencies of the learner and his history can be mapped onto the learning outcomes of activities in the learning network, depends to a great extent on the given data. Learner data may result from formal, accredited learning as well as from experience gained in informal learning situations. Description of competencies may range from completely absent to being based on an ontology or at least a controlled vocabulary. To address the positioning problem in learning networks we assume that learners will enter a learning network with a variety of different data stored in learner profiles or electronic portfolios.

On the other hand the learning network itself can consist of a loosely coupled collection of material or a very well structured collection of learning activities where we have a connection between them and competencies or competence levels. We surmise that alternative approaches to positioning need to be based on the type of competence descriptions (of learners as well as programs) that are available. We seek computational approaches to positioning that ultimately fulfil the criteria of reliability (the same situation leads to the same recommendation) as well as validity (the recommendation matches that of experts). A reliable positioning service has to bring always the same result from the same given data, while the validity can only be compared to human performance for the positioning problem. The different situations and data for positioning are shown in the Positioning Situations-Matrix in figure 1:
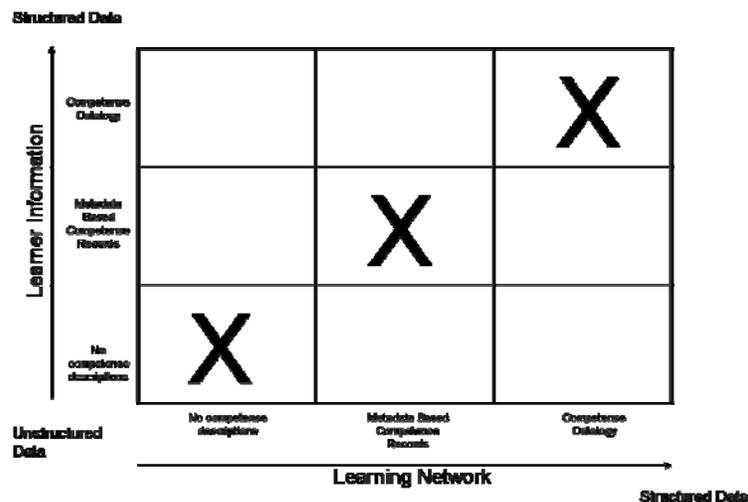

Figure 1: The Positioning Situations Matrix

The three cases discussed here represent the "symmetrical positioning" where similar data are compared. The more complicated positioning would be the "asymmetrical positioning" where for example a competence ontology in the learner profile should be mapped to the content of a competence development program. To cover all these different situations and to ensure the best achievable position inside a learning network we compare different situations and approaches. In this paper we limit the discussion to three cases of symmetrical positioning:

Case 1: Informal descriptions
The learner enters an educational environment without any explicit competence descriptions. The competence development program is highly informal without information about the resulting competence of learning activities. An example: A learner wants to update his competences in accounting. He enters a learning network that deals with the finance and accounting domain. His competence development goal will be reached through an informal collection of learning activities. His electronic portfolio contains only some documents he produced in his former education. Here, a content based approach, as discussed in the next section is best suited for positioning.

Case 2: Metadata based positioning
If a learner enters with a standards-compliant ePortfolio the situation would be different for a positioning service. To take the same example as in case one the learner enters with a standards-based description of his competencies and the activities in his chosen learning network have detailed information about requirements or competence result for a learning activity. In section four we review and discuss standards and the way they can support the positioning process.

Case 3: Ontology-based positioning
If there are competence ontologies inside the learner profiles and the competence development program the positioning problem can be based on mappings between the ontologies. The same learner as in case one and two enters a learning network with a very detailed competence-ontology or competence map that shows his already acquired competencies. The learning network contains an agreed upon domain-ontology where all aspects of the domain are modelled. Additionally a competence-ontology has relations to the domain ontology. This highly structured description in the learning networks allows a direct comparison between the competence-ontology in the learner profile and the competence-ontology in the current learning network.

A Content-Based Approach to the Positioning Problem

The rationale and the research agenda for a content-based approach to positioning was described in (van Bruggen et al. 2004). The approach rests on the following assumptions. Because it would require extensive assessment it does not aim to directly demonstrate that the learner has already acquired knowledge, skills and competences that are equivalent to the outcomes of learning activities within the routes considered. The core assumption is that equivalence of outcomes will be reflected in, or can be approximated by, the similarity of the contents of (learning) materials studied or produced by the student (source material) and the material contained in the learning activities in the learning network (target). If a positioning service determines that the content of source and target materials overlap substantially, the target activity is exempted. In our content-based positioning service document similarity is computed using latent semantic analysis (LSA) (Deerwester at al. 1990). LSA is based on word (co)-occurrences in documents, thus all order (syntax) of words or semantics in the original documents is ignored. All analyses are performed on a Term-by-Document matrix with word frequencies in the cells. The dimensions of this matrix are computed and the largest dimensions found (the semantic factors) are retained to reproduce the original matrix (Landauer & Dumais 1997). In the

reproduced matrix each document is represented as a vector. The smaller the angle between two document vectors the higher they are correlated, that is, they are expected to contain materials that have substantial overlap. Learners are represented by one or more documents that they have produced or studied. If one or more of these learner document vectors demonstrate a high correlation with learning material vectors, then the learning material may be considered redundant. Although the content-based approach has modest requirements on the way data are expressed, there are several limitations and assumptions that we need to consider like the amount and quality of available material in the learner profiles for content-based positioning.

Metadata-Approaches for the Positioning Problem

Metadata are used to describe learning resources as well as learner profiles. Several efforts from standardization bodies and working groups aim at unifying competence descriptions and competence levels. The IMS Reusable Definition of Competency or Educational Objective (RDCEO) specification aims at a standard description of competencies and educational objectives for online and distributed learning. RDCEO is expected to promote common understanding of competencies that can be used in competency development (learning and career development) or in specifying learning pre-requisites or learning outcomes (IMS 2002). The RDCEO offers a unique identifier to assign an unstructured competency description to an object for example in a Unit-of-Learning (UoL). Based on the RDCEO a draft standard for Reusable Competency Definitions (RCD) is being defined in the IEEE. Although RCD does not intent to offer a solution to the aggregation of competencies from sub-competencies the data-model allows the integration of relational information or competence ontologies through embedding additional metadata (IEEE LTSC 2006). For portfolios two specifications are of interest. The IMS Learner Information Package Specification (LIP) is designed to package learner information for the exchange of data (IMS 2001). The IMS ePortfolio specification builds on the LIP specification to ensure portability and exchange of ePortfolio records for learners (IMS 2005). The specification is addressing different usage possibilities (assessment, planning of learning) and it can store produced artifacts form the learner and formal achievement records like references. A slightly different approach comes from the HR-XML Consortium. The consortium develops a standard suite of XML-specifications to allow the exchange of Human-Resource-related data, such as a competency schema for a variety of business contexts that is applicable in recruitment processes (HR XML 2004). The model allows the evaluation, rating and ranking of competences which are an important issue in recruiting processes.

While these metadata were all related to the learner profiles different standards in the learning network are also important for the positioning problem. The IMS Learning Object Metadata (LOM) is used to assign metadata to learning objects. For the positioning service it is important that there is no element in the LOM standard to store competence related information at the moment (Ng, Hatala & Gasevic 2006). They could be stored in the educational segment of the metadata as proposed in (Sanchez-Aloso & Sicilia 2005) but this does not seem to be a widely adopted solution to the problem. On the authoring level, IMS Learning Design (IMS LD) can also be used to take into account prior knowledge (IMS 2003). One of the use-cases states that IMS LD can be used to reduce the content in a learning path to reduce the time required to reach learning objectives. Through conditions a learning activity can be skipped if a learner already knows enough about the specific subject of the learning activity.

For the positioning service not only the specifications and standards are important but also the way they can be compared to each other. Since the underlying data models differ in the above presented standards the interoperability of competence related metadata is a problem. Chang and Zeng (2006a) present different options for comparing and mapping metadata. One solution for this problem is the development of a

crosswalk for competence-related metadata. A crosswalk is a specification for mapping one metadata standard to another (St. Pierre & LaPlant 1998). While crosswalking works well when the number of involved schemas is small it can become a very complicated matter when a high number of schemas should be compared. Especially the exponential increase of relationships for the different schemas leads to a complexity problem. Therefore Chang and Zeng (2006) introduce the method of metadata switching. Instead of using a many-to-many relationship the switching method uses one schema as the central relation for all other schemas. Other options for interoperability are a metadata framework or a metadata registry. As a metadata framework tries to integrate all solutions in a common architecture a metadata registry collects information about different schemas in an environment and allows crosslinking and mapping.

One can imagine that it could still make sense to combine these approaches with one that is based on the content of learner profiles and the learning network. The specifications discussed here allow the integration of external competence models. They make (meta-)data available for a positioning service, and may serve the purpose of opening more data for content-based positioning. The standardization activities alone, however, have a limited usefulness for competence mapping and the formalized description of complex competence relationships. The interoperability standards discussed above serve the purpose of sharing data. They themselves do not ensure the semantics of the data, i.e. there are still different ways to describe the same learning outcomes, such as competencies.

Positioning with Competence Ontologies

The missing link between the standards and the competency mapping may emerge from the use of competence ontologies and semantic web technology (Koper 2004). Ontologies are metadata schemas providing a controlled vocabulary of concepts and they can be useful to share common understanding in a domain in a machine-readable way. For competence development ontologies or taxonomies can be used to define competences related to learning activities. Competence ontologies could be either added to the learner profiles (Dolog & Schaefer 2005), learning objects (Ng, Hatala & Gasevic 2006) or the competence development programs (Woelk 2002). But the design and implementation of competence ontologies is still a very complex and time consuming task. Su (2002) presents three different situations for ontology mismatch: The single ontology approach where all information sources are related to a unified global ontology, a multiple ontology approach where every information source has its own ontology without a shared vocabulary and a hybrid approach where all information sources have their own ontology but they use a unified shared vocabulary. In an ideal situation every learning network could share a common understanding of the competences needed for successful running through a competence development program based on ontologies. In this case positioning inside a learning network can be based on the relations between a domain ontology and the competence ontology (Pozea & Harzallah 2004). The process of adding competencies in the learner profile could be derived from successfully finished learning activities in the learning network. Parts of the competence ontology in the learning network could be added to the learner profiles step-by-step after they have successful passed the related assignments. For the multiple ontology-approach, ontology similarity is the key factor for successful positioning (Maedche & Staab 2002). In the next part of the paper we will discuss the presented approaches and try to give an outlook for our research on positioning in the future.

Discussion

Positioning a learner in a learning network for lifelong learning is a complex task by itself and this is exacerbated by conditions that prevent any simple mapping of learner profiles and competency descriptions onto the educational resources. The two most extreme situations that we considered are the clearest: (1) no competency descriptions inside the learner profiles and the learning network and (2) competence ontologies in the learner profile and the learning network. In the first case a content-based approach is the one to take. The content-based approach to the positioning problem has the advantage that it can be used for positioning right now, where most learners do not have a detailed profile with explicit competence descriptions. The drawback of the approach is that it is only related to the produced content of the learner and not to his earned competences. So the success is dependent on the amount of text the learner can provide in relation to his educational history. If he can for example only add content from several parts of his educational background, the positioning recommendation will be biased. Additionally, the concentration on content may effectively limit the approach to domains with a strong verbal character. For the same reason, domains with psycho-motor content, for example practical skills, may not be adequately represented. In the second case a mapping of ontologies could be a feasible technique to reach the ideal position for the learner. For the positioning problem all the data models can be useful because having machine-readable information about the competences of the learner simplifies the positioning task if we have also competence descriptions inside the chosen learning network. But, there are drawbacks to this approach and those related to it: all the presented metadata-based initiatives offer a way to ensure a standardized description of competence related data. The models differ from openness (in terms of the possibility to embed ontologies or taxonomies) and intention (packaging focus or description focus). But the biggest drawback with metadata- and ontology-based approaches is the economical side of the medal. A huge amount of work has to be invested to enrich learning resources and learner profiles with metadata and competence ontologies. Another problem is that metadata and ontologies are always arbitrary models to a knowledge domain and that objective ontologies don't exist (Shirky 2003). Besides it is very expensive to let domain experts guarantee the quality of the metadata used. Several experiences from repositories have shown that it is not an advisable idea to pass the burden of metadata-enrichment to the users. The quality of user created metadata cannot be compared to the quality of experts (Barton, Currier & Hey 2003).

Outlook

Our research will focus in the future on computational approaches to address the three presented cases for positioning of learners in learning networks. While there are already several individual experiences in all of the presented cases a combination of them is a new and previously untried approach to address the positioning problem. Because of feasibility reasons we will concentrate for the development of a positioning service on a situation from formal education where we will have predefined activities in the learning network. All experiments will be done in an introductory psychology learning network of the Open University of the Netherlands. The research project is divided into three phases defined by the above described cases. The result of every phase should add a model and techniques to the development of a prototypical positioning service. For the first experiment we will concentrate on the use of content-analysis, for the second experiment we will additionally make us of metadata and for the third experiment we analyze a combination of content, metadata and competence ontologies. For every development cycle the validity and reliability of the results should be controlled through the comparison with the prior learning assessment results from domain experts who analyse the same material as we use for the positioning service. While the focus of the project is on the comparison of similar data it is still an open question how an asymmetrical positioning could be addressed.

References

Barton, J., Currier, S. & Hey, J. M. N. (2003). Building quality assurance into metadata creation: an analysis based on the learning objects and e-prints communities of practice. In Sutton, S. and Greenberg, J. and Tennis, J., Eds. Proceedings 2003 Dublin Core Conference: Supporting Communities of Discourse and Practice - Metadata Research and Applications, 2003 Seattle, Washington (USA).

Cheetham, G. & Chivers, G. (2005). Professions, Competence and Informal Learning, Cheltenham:Edward Elgar Publishing.

Cheng, L. M. & Zeng, M. L. (2006). Metadata Interoperability and Standardization – A Study of Methodology Part 1. D-Lib Magazine, Vol 12. No. 6, ISSN 1082-9873. Retrieved 1st of August from

Deerwester, S., Dumais, S.T., Furnas, G. W., Landauer, T. and Harshman, R. (1990).Indexing by latent semantic analysis. Journal of the American Society for Information Science, vol. 41, 391-407.

Dolog, P., Schaefer, M. (2005). A Framework for Browsing and Manipulating and Maintaining Interoperable Learner Profiles, Proc. of UM2005 - 10th International Conference on User Modeling, July, 2005, Edinburgh, UK. Springer Verlag.

HR XML Consortium (2004). Competencies (Measurable Characteristics). Retrieved 1st of July, 2006 from http://ns.hr-xml.org/2_3/HR-XML-2_3/CPO/Competencies.html

IEEE Learning Technology Standards Committee (2006). Reusable Competency Definitions (P1484.20.1, Draft 3). Retrieved 1st of July, 2006 from http://www.ieeeltsc.org/wg20Comp/wg20rcdfolder/IEEE_1484.20.1.D3.pdf

IMS Global Learning Consortium (2001), IMS Learner Information Packaging Specification. Retrieved 1st of July, 2006 from http://www.imsglobal.org/profiles/

IMS Global Learning Consortium (2002). IMS Reusable Definition of Competency or Educational Objective Specification. Retrieved 1st of July, 2006 from http://www.imsglobal.org/competencies/

IMS Global Learning Consortium (2003). IMS Learning Design Best Practice and Implementation Guide.

IMS Global Learning Consortium (2005). IMS ePortfolios Specification. Retrieved 1st of July, 2006 from http://www.imsglobal.org/ep/index.html

Koper, R. (2004). Use of the Semantic Web to Solve Some Basic Problems in Education: Increase Flexible, Distributed Lifelong Learning, Decrease Teacher's Workload. Journal of Interactive Media in Education, 2004 (6). Special Issue on the Educational Semantic Web.

Koper, R., Rusman, E. & Sloep, P. (2005). Learning Network connecting people, organisations, software agents and learning resources to establish the emergence of effective lifelong learning, *LLine: Lifelong Learning in Europe*, vol. 9, no. 1, 18-27.

Landauer, T. K. & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychological Review, vol. 104, no. 2, 211-240.

Maedche, A. & Staab, S. (2002). Measuring Similarity between Ontologies. Proceedings of the European Conference on Knowledge Acquisition and Management (EKAW). Springer.

Merrifield, J., McIntyre, D & Osaigbovo, R. (2000) Mapping APEL: Accreditation of Prior Experiential Learning in English Higher Education. London.

Ng, A., Hatala, M. & Dragen Gasevic (2006). Ontology-Based Approach to Learning Object Formalization, In: Sicilia, M.-A. (Ed.), Competencies in Organizational E-Learning: Concepts and Tools, Idea Publishing 2006.

Van Bruggen, J., Sloep, P., van Rosmalen, P., Brouns, F., Vogten, H., Koper, R. & Tattersall, C. (2004). Latent semantic analysis as a tool for learner positioning in learning networks for lifelong learning, British Journal of Educational Technology, vol. 35, no. 6, 729-738.

Van Bruggen, J., Rusmann, E., Giesbers, B. & Koper, R. (2006). Content Based Positioning in Learning Networks. Proceedings of the

Posea, V. & Harzallah, M. (2004). Building an Ontology of Competencies. Proceeding of Workshop on Ontology and Enterprise Modelling: Ingredients for Interoperability. In Conjunction with 5th International Conference on Practical Aspects of Knowledge Management. December 2, 2004. Vienna, Austria.

Sanchez-Aloso, S . & Sicilia, M.-A. (2005). Normative Specifications of Learning Objects and Learning Processes: Towards Higher Levels of Automation in Standardized e-Learning, International Journal of Instructional Technology and Distance Learning. Vol 2. No. 3., ISSN 1550-6908.

Shirky, C. (2003). The Semantic Web, Syllogism and Worldview, 2003, Available from http://www.shirky.com/writings/semantic_syllogism.html

Skinner, K. (2005). Continuing Professional Development for the Social Services Workforce in Scotland. Dundee:Scottish Institute for Excellence in Social Work Education.

St. Pierre, M. & LaPlant, W.P. (1998). Issues in Crosswalking Content Metadata Standards, National Information Standards Organization (NISO), Available from http://www.niso.org/press/whitepapers/crsswalk.html

Su, X. (2002). A text categorization perspective for ontology mapping. Technical report, Norwegian University of Science and Technology, Norway.

Woelk, D. (2002). e-Learning, Semantic Web Services and Competency Ontologies, Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2002.