E-learning based on
the Social Semantic Information Sources

Adam Westerski[1,2], Sebastian Ryszard Kruk[1], Krystian Samp[1,2], Tomasz Woroniecki[1,2], Filip Czaja[1,2], and Caoimhin O Nuallain[1]

[1] Digital Enterprise Research Institute, NUI Galway, Ireland
[2] WETI, Gdansk University of Technology, Poland

<firstname.lastname>@deri.org

Abstract:

E-Learning is no longer just about the materials burned on CD-ROMs and sent across the country. With the potential offered by the Internet, there are new on-line courses emerging every day. However successful they are, they still do not utilize the potential of online communities or semantic descriptions of learning objects. As a result, learning material is usually static and not adjustable for the specific needs of the learner. What is more, there is no simple way to reuse existing e-learning material.

In this article we present how the concept of social semantic information sources can be adapted to the e-learning systems. We define a 2-layer learning objects enrichment architecture that allows interoperability between various e-learning systems. We present how dynamic courses can be delivered by utilizing user profile information.

Keywords:

E-learning, Semantic Web, personalisation, course composition, metadata vocabularies.

1. Introduction.

In the early 1990's, e-Learning was pushed as the "killer application" of our time, as it was to allow the delivery of education to everyone everywhere. It was also promised to allow a high level of flexibility, as the users could log on at any time and continue their learning when it suited them best. This led to a high degree of popularity for the concept of e-Learning and the potential that could be achieved through it in terms of learning. The development of e-Learning had associated with it the big advantage of cost efficiency over instructor-led training. However, much of the promise was just a gloss as the material available had a number of problems associated with it as "one size does not fit all" [7]. The material was seen to lack instructional design, educational potential, engagement and any feedback to the user. Despite the use of the Internet, the courses are very static and flat. Further enhancing of acceptability of courses is limited by the lack of bandwidth and limited access of people to computers and the internet.

1.1. Semantic Web 2.0.

When Tim Berners-Lee invented the World Wide Web it was a mere collection of HTML documents. Soon and rapidly it grew to the stage we have at the moment. The Internet has become more than just a source of information. It has become a source of entertainment, communication and last but not least – business opportunities. However, with the search

engines as robust as Google, everyone has the feeling that we cannot grow the Internet the way we did so far. Even now, many people just cannot efficiently search for information.

The second-generation Internet is currently the hot topic both in industry and academia. It is perceived as a remedy for all problems we know from the current Internet. However, academia and industry defines the future Web in two different ways.

Research centres mainly focus on the work on the Semantic Web. In this vision, the future Internet will be more than just human-understandable text. The idea is to add machine-processable meaning to the current and future information. Future search engines on the Semantic Web will be able to understand both the information they index and the users' queries they process. Business to business (B2B) systems will be able to cross the boundaries of heterogeneity, and find better deals with partners they cannot communicate with at the moment.

Web 2.0 is the Holly Grail of the contemporary Internet companies. Instead of making the information machine-understandable, Web 2.0 brings the whole communities of users to interact with the information and each other. Wikis allow groups of people to edit the information in truly collaborative fashion. Endeavours like wikipedia[1] proved the immense potential of community impact. Web 2.0 is also about the tagging. In services like deli.cio.us or Flickr, community users annotate bookmarks or photos they share with simple set of keywords. As opposed to the old Web everyone can annotate each resource and in contrast to the Semantic Web there is no meaning applied to each keyword.

Semantic Web also aims to grasp the potential of online communities by initiatives like FOAF (friend of a friend), that describes online communities in a semantic fashion. FOAFRealm[2] is one of the projects based on FOAF metadata. One of its interesting features is the social semantic collaborative filtering (SSCF) [1] that incorporates solutions known from collaborative filtering in Web 2.0. Apart from the given example other projects also aim to utilize social semantic information sources defined by the emerging Semantic Web 2.0 standards.

A digital library called JeromeDL is one of the first systems that binds together Semantic Web and Web 2.0 efforts. The idea of social semantic information sources has been implemented in a 2-layer metadata enrichment architecture. The lower layer is responsible for lifting up concepts of legacy metadata like MARC21, BibTeX or DublinCore to the semantic level. This allows the interoperability with already existing legacy digital library systems. However the legacy metadata (especially MARC21) is usually hard to understand by an average user of a digital library. Therefore, JeromeDL delivers a second layer of metadata enrichment that is community oriented. A community of users (and authors) can interact in the Web 2.0 fashion by tagging resources through the Social Semantic Collaborative Filtering (SSCF) interface. SSCF allows users to annotate resources (and share those annotations with their friends) reflecting the way they perceive the world. Semantic information managed by both layers of metadata enrichment is later used by the semantic query expansion algorithm that takes user interests into account.

In this article we will present how a similar 2-layer approach can be applied to e-learning solutions.

1.2. Use case scenario.

This section describes some sample real life scenarios pointing out the difficulties that our research is trying to cope with.

---

[1] http://www.wikipedia.org/
[2] http://www.foafrealm.org/

Contemporary e-learning solutions become more and more popular. The industry grows very quickly. Along with it, user expectations increase as well. When a user approaches an e-learning system, the materials he/she will see and the courses that he/she will have access to must be selected based on some of his/her individual characteristics. When a student starts a course with a fixed curriculum constructed according to producers' assumptions, it is very likely that a considerable part of the course will not be adequate to his/her actual requirements. This will lead to dissatisfaction and confusion which makes the whole learning process a disaster. For example if a user does not learn as fast as the courses' producer assumed, successive materials will be too hard. Sooner or later that person will resign from using the e-learning solutions. On the other hand, if some parts of the course are too easy or concern subjects that user is not interested in, he/she will get bored very quickly. A possible resolution for this problem could be user personalisation and an on-demand dynamic course composition.

While pushing research towards these directions certain facts have to be taken into account. During the learning process people often tend to ask their friends about advice or knowledge that could aid them to choose proper sources like books or ways of learning the material. Usually in e-learning a user is simply given a fixed course assigned to a subject he/she selected to learn. The shape of the course is not always appropriate for the user. E-learning solutions should incorporate an advising mechanism based on similarities between users e.g. friendship relations or some characteristics found in semantic annotations.

1.3. Contribution.

The solutions described in this paper focus on improving the quality of e-learning systems by proposing a method of assembling an on-demand curriculum from existing learning objects provided by e-learning services suppliers. In order to improve user satisfaction, the selection and the work-flow scheduling of learning objects is based on the semantically annotated specification of the current user's skills, anticipated resulting knowledge and technical details of the clients platform.

1.4. Outline.

This article describes ways of introducing Semantic Web technologies to e-Learning services in order to achieve the goal of composing on-demand courses. As research and extensive talks with major Irish companies dealing with e-learning have shown, the key parts of this solution are to provide machine processable semantic descriptions of:

• A user's profile: knowledge level in different domains, goals or expectations from the course, technical availability specifications, previous evaluations;

• Learning objects: required/optional knowledge / skills / evaluations/ requirements, value added, cross-references with other services, pre/post-conditions, technical requirements specifications;

• Assessments: kinds of evaluations/ test/ exams provided, users' evaluations on the learning, technical requirements specifications;

2. Theoretical analysis.

Several key services can be identified in the generalized model of LMS (see Fig.1).

The content Repository stores educational digital data in the form of assets and learning objects. Both can be reused in different contexts to create learning units with different

objectives and purposes. To facilitate a management of the sets of resources, descriptive metadata is used. Although the Learning Object Metadata (LOM) specification tends to be the most popular among different systems, many formats and variations were developed.
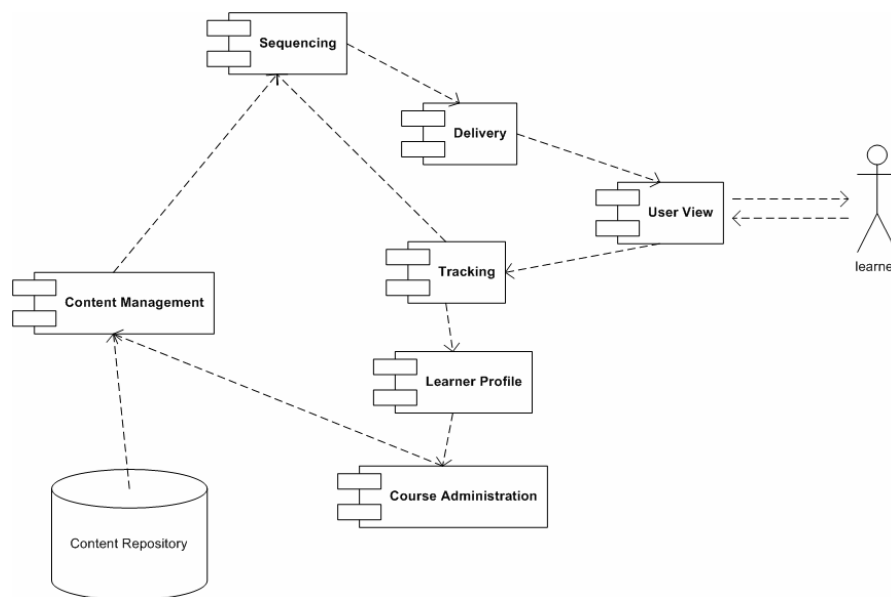

Figure 1. Current LMS architecture.

The content Management component is responsible for preparation of educational units. Many input parameters are taken into consideration during the process of creation: user capabilities, preferences, hardware, course subject etc. Finally, a ready to use instructional unit which fulfils requested learning objectives is formed.

Sequencing and Delivery modules display a learning unit to a learner according to some rules which determine the order of learning activities.

Tracking, Learner Profile and Course Administration services track user progress. Information about accomplishing particular parts are stored in the user profile and decisions where to go are made.

2.1. Problems of existing LMSs.

It is important to realize the role of a typical learner. Although the user profile is used to deliver appropriate, personalized learning units, actions performed during the instructional process are not used in any other way but to manage curriculum scheduling. Moreover, the fixed path of learning activities is prepared for a learner and he/she is not able to interact with this aspect of processing. The learner is considered as a person who only accomplishes static steps prepared in advance by an LMS; thereby, the user plays a very limited and narrow role in the learning process.

Current LMSs use different metadata to describe their content objects stored in repositories. This leads to another problem related to interoperability. There are a lot of diverse learning objects which can be helpful in many contexts. However, they are not described in a uniform way, which makes them useless despite a lot of effort that was put in to create them.

A similar problem with using different description formats concerns user profiles. Learners who use LMS are separated from each other. Their experiences are only used in their own environment and therefore are not reused.

In the following sections we will present weaknesses and will suggest necessary changes to LMS architecture.

2.2. Observations and suggestions.

We believe that user interaction and activity carry a lot of useful information which can be used in variety of ways to improve the quality of learning experiences. To utilize this potential as much as possible the information from all learners should be gathered, shared and reused. Moreover, the learner should be able to adjust and refine the process of instruction and make his/her own annotations and bookmarks.

Choices made by a particular learner can indicate what the prerequisites are for a given learning object. We know the user profile: capabilities, preferences, history, courses taken etc., and we can assume with some probability that these properties are needed to start this course. Collecting similar information from other learners, we can provide more and more accurate assumptions. After reaching a particular line of certainty, the metadata of a considered learning object can be updated and new information reused in the future to better suit a learner's abilities.

The same can be applied to acquire information of learning unit objectives. Observing the choices of a learner after finishing a given learning object can give us clues and hints about new skills gained by the user.

Moreover, the choices made by a learner during a course can be stored and reused to propose a similar path for another similar person interested in the same topic. This way, experiences of other people can help to teach new pupils in better, more adjusted fashion.

Bookmarks and annotations of individuals also carry important information [1]. People create their own classifications and hierarchies which are of use for others who try to find interesting materials from particular domain of interest. We believe that learners should form more collaborative, open communities and share their knowledge. This would improve acquiring new information.

To facilitate searching, discovering and learning, bookmarks and annotations of other people are freely available. Our solution introduces this feature with the possibility to make some restrictions. We propose to utilize the Distributed Friend of a Friend network [2, 3] which stores and manages user profiles. It enables us to exchange the bookmarks and other information between different people. Moreover, groups of friends, who share more data (e.g. their entire user profiles), can be created. A new user who joins the network can add his/her friends to the list of known people. Using this information in their profiles, we can conclude some initial knowledge about this new learner.

We also propose how to deal with the wide diversity of metadata formats that exist for describing learning objects and user profiles. To address this problem we suggest using the Semantic Web technologies, especially an ontology approach for modelling problem domains.

3. E-Learning on the social semantic information sources.

3.1. Architecture of semantic e-learning system.

To deal with the problems presented in 2.1 and to address all observations in 2.2, we introduce two main elements:

 - Ontology for e-learning content.
 - Ontology for user profile.

These are the basis of the architecture of an e-learning system based on the social semantic information sources (see Fig.2).
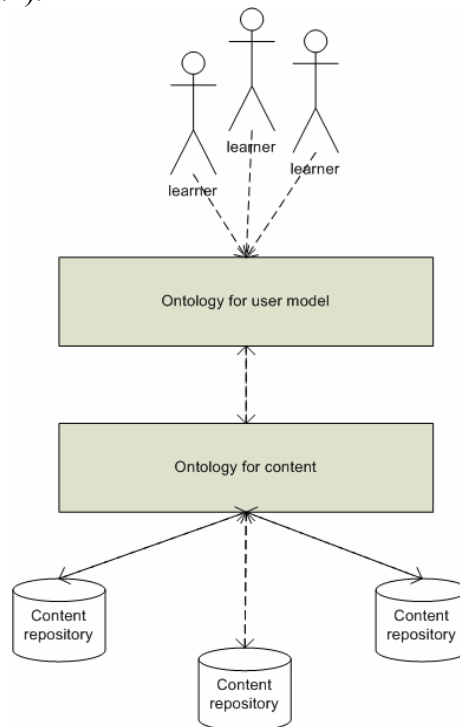

Figure 2. Architecture of the future e-learning system

3.2. Ontology for legacy content description.

Contemporary learning systems describe their resources using mainly Learning Object Metadata, Dublin Core, BibTeX and many other formats which are created to fulfil the needs of a particular product or company. Although the mentioned specifications are mature and have existed for a long time, they miss some key concepts which are needed to utilize our ideas.

We distinguish two goals for resource description:
- express common concepts from different formats
- preserve the information connecting user to e-Learning object content (e.g. object prerequisites, objectives that learning object fulfils, user skills required)

To address the first problem we propose using an ontology approach. A common conceptual level will preserve the semantics of different descriptions and ease the mediation between them. This will ensure cooperation of heterogeneous environments which use different formats to accomplish their tasks [4,5].

The resolution of the second goal is reached by proposing certain extensions to the ontology that make it e-Learning content more user aware (see Fig. 3).
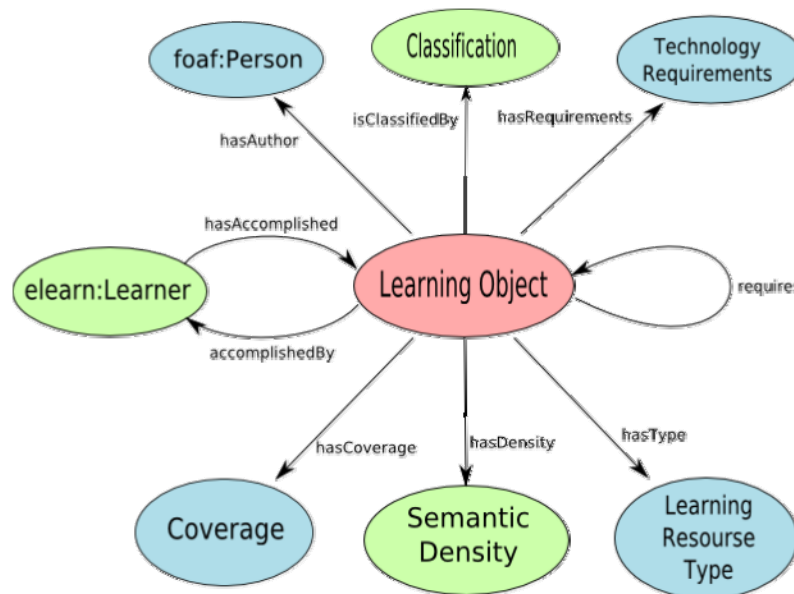

Figure 3. Structure of Learning Object ontology

Learning Object Ontology delivers necessary information about given Learning Object (LO):

- Classification - classifies a learning object and determines the category it belongs to. It defines prerequisites, required skills, competency, educational and security levels and accessibility restrictions.
- Technology Requirements - technical requirements that must be fulfilled by student's devices in order to use given LO. For instance this part concerns aspects like supported operating systems and internet browsers.
- Coverage - the time, culture, geography or region to which this LO applies. It gives the extent or scope of the content of the learning object. Coverage will typically include spatial location (a place name or geographic coordinates), temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity).
- Semantic Density – semantic density of a given LO
- Learning Resource Type – defines the type of a learning resource to which this LO refers. Most common types are: diagram, exam, exercise, experiment, figure, graph, index, narrative text, problem statement, questionnaire, self assessment, simulation, slide and table.
- elearn:Learner – a learner who is using given LO during his learning path. He is described by separate ontology (see Point 3.3).
- foaf:Person – an author of a given LO. One LO can have more than one author. He is described by separate ontology.

Information obtained from all these parts makes possible to choose a proper LO for each user. It allows to reason about necessary skills and requirements that must be fulfilled by a student in order to complete given LO.

3.3. Ontology for user learning profile.

In order to deliver a personalised content, the system gathers as much information about the user as possible (see Fig. 4).
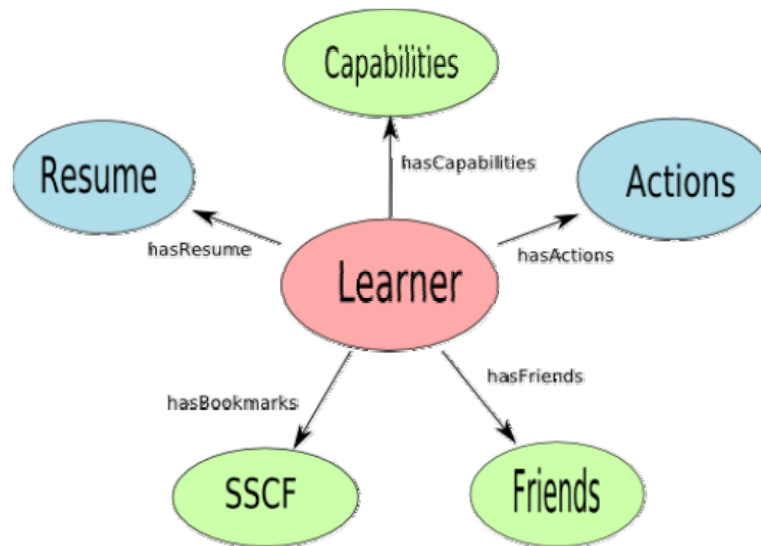

Figure 4. Structure of learner profile ontology

The FOAFRealm ontology used to store this information covers a wide area of different aspects of a learner's profile:

- resume – personal description of a user, including education, areas of expertise, work experience, career level, etc.
- capabilities – circumstances that may affect the learning process (e.g. user's disabilities) and also a description of equipment used (e.g. a mobile phone with a limited display).
- actions – history of user's choices about courses and learning objects (e.g. which of given alternative course parts were picked). This part of the ontology covers all different paths and scenarios chosen by the user. Such information is used to help the user with his future choices and also is a source of advice to other people who are figured out to have similar preferences.
- SSCF (social semantic collaborative filtering) – bookmarks with courses and objects that the user finds valuable or interesting; this information is reused as suggestions for his friends and people with similar interests.
- friends – data about friendships from FOAF profiles. If there is not enough detailed information about the user (for example the user has not fully filled his/her profile), preferences of his/her friends might be used for personalisation (assuming that generally friends have common interests).

Information from each of these parts of the ontology influences the learning process and the materials composition in different ways. Resume provides a hint on user's interests and helps to decide which materials are inappropriate for a particular user (a seasoned engineer should not be taught basic maths, etc.). The capabilities section obviously determines the presentation techniques which might be used with the additional device.

Information deduced from user actions has to be stored in the resource description. Contemporary metadata formats do not provide sufficient vocabulary for such purposes and thus have to be extended. In order to efficiently compose courses we need to express information about prerequisites and learning objectives more precisely. Learner choices will help to determine exact abilities that are needed to start certain learning object and which are gained after finishing it. The learner's ontology branch called "actions" is the connection point between the user profile and the previously described learning object ontology.

3.4. Ontologies utilisation in the learning objects composition process.

This section describes the usage of ontologies discussed previously in order to create user oriented courses. E-learning object ontology and user profile description are key elements for a mechanism that composes courses according to individual user abilities and preferences.

The subject of objects orchestration can be examined from many points of view. The one that is reasoned and researched often is composing based on workflow of specific information. Such techniques can be very clearly seen in research concerning web services composition. During the search for components that will fulfil a user's task, the most important goal for the system is to match compatible services within the flow in a way that will enable it the transfer of data from one service to another in the chain. The solution of this problem allows the reuse of the previously processed data and compose a bigger functionality from smaller parts. Although similar attempts have been made in the e-learning domain by introducing tailoring of an object's size and level based on some input parameters [6], this paper takes into account user characteristics and requirements concerning e-learning content and extends the discussed meaning of composition. In order to bring users greater satisfaction from the system on-the-fly proposed course, composition techniques consider proper object ordering within the flow. This is based on technical aspects of learning object connectivity as well as on the descriptions of object content and its relation to user goal and profile description.

The main idea which could lead to achieving our goal is by taking advantage of benefits brought by initiatives like FOAF user profiling [1, 2, 3]. Collaborative filtering technologies already developed by our group [1] allow to find people with similar interests and extract information stored by them. We propose to use this mechanism during course composition to extract data describing other people choices regarding composition of their courses. When at some point user tries to compose a course system searches for similar users and analyses their courses. If the goal and the subject of courses match, their structure is used to suggest the newly created course shape. In the final solution, a user can select the track that he/she wants to follow by choosing components suggested by the system. At each point system recommendations should aim to fit best to user requirements. In order to analyse the composed courses structure and choose their elements for the new course system needs each learning object description to hold some pre and post conditions. Basing on those it can be identified what a given object requires and gives.

According to our current research there are two main sources of preconditions. The first source is based on described previously user profiling and community information. While crafting a course system should dynamically create a list of possible objects that the user could select from at each course step. The contents of the list should be based on choices of people of similar interests made in the similar courses. In this case the preconditions for a particular object are its position in the analysed course and the objects that other users defined as required to complete before one. In order to limit the amount of possible objects to check, the required strength of similarity between users and courses has to be defined. Depending on how much two users are similar to each other their courses should be used for aiding

composition of new ones. The definition of the "similar interests" concept can be understood in various ways depending on the amount of knowledge the system has about current user. In the best case, the system would analyse choices of people that user him/herself declared as of similar interest (for example by utilizing the functionality of the previously mentioned FOAF technology). However, if we are dealing with a totally new user it is possible that he/she has not described people of similar interests or assigned him/herself some kind of user group. In that case system should search for the similarities based on information such as nationality, occupation or other filled in by the user and stored in user ontology proposed in section 3.3.

The second source of preconditions is a less innovative idea but also very important, when no prior information about the user is available. The ultimate goal of composing the course can be aided by analyzing definitions found in an object's description ontology (see section 3.2) that contains some predefined suggestions about the required users' experience and specifies the context of an object in a given domain. The predefined descriptions of required user level are a good guide for the system at its bootstrapping. Taking notice of object context is very important even when the system has wide knowledge about the users and their learning choices. By analysing the paths users have taken during the course composition, the system can find objects that are popular at some point of learning in a given subject but do not necessarily concern the subject directly. For example, while learning Spanish, at some point many foreign users might have chosen to learn about Spanish history. That fact should not result with the system proposing Spanish history lessons to a new foreign user that is only interested in learning the language and grammar. By comparing object context with user expectations automatic course generation can be controlled in a way to give better results and satisfy user more.

In order to maintain the quality and correctness of a proposed courses the system should be able to track post conditions of the objects that user selects and modify further parts of the course. Post conditions are information defining what benefits a user gains by completing an e-learning object and what level of knowledge he/she will hold. At this point of our research, to achieve the goals described at the beginning of the article, we extract this information only from fixed descriptions stored in the proposed learning object ontology. Ideally this information could also be created dynamically on-demand by analysing what similar users have learned after completing the course. This solution however assumes some additional input from the user to rate objects and is rather meant for future research.

4. Conclusions and Future Work.

In this article we presented the vision of an e-learning system based on the Semantic Web 2.0. The concept of the social semantic information sources is yet to be clearly defined. However, we showed that by utilizing existing Semantic Web and Web 2.0 solutions, the e-learning system can build a dynamic course from learning objects provided by various LMSs.

This theoretical model of an e-learning system is currently being built in the eLITE project. The Semantic Web 2.0 components, like: JeromeDL, FOAFRealm and Didaskon, delivered by the Corrib.org initiative, are the building blocks of the future e-learning system based on the social semantic information sources.

References
[1] S. R. Kruk, S. Decker. Semantic Social Collaborative Filtering with FOAFRealm. Semantic Desktop Workshop, ISWC 2005
[2] S. R. Kruk. FOAF-Realm - control your friends' access to resource. In Proceedings of the 1st Workshop on Friend of a Friend, Social Networking and the (Semantic) Web (FOAF Galway), Galway, Ireland, pages 1–9, September 2004.

[3] S. Grzonkowski, A. Gzella, H. Krawczyk, S. R. Kruk, F. J. Martin-Recuerda Moyano, T. Woroniecki. D-FOAF - Security aspects in distributed user management system. In Proceedings of TEHOSS 2005.

[4] S. R. Kruk, M. Synak, K. Zimmermann. MarcOnt Initiative - Mediation Services for Digital Libraries, In Proceedings of ECDL 2005.

[5] S. R. Kruk, M. Synak, K. Zimmermann. MarcOnt - Integration Ontology for Bibliographic Description Formats. In Proceedings of the International Conference on Dublin Core and Metadata Applications (DC-2005), Madrid, Spain, September 2005.

[6] Jovanović, J., Gašević, D. Devedžić, TANGRAM: An Ontology-based Learning Environment for Intelligent Information Systems. Paper presented at the E-Learn 2005 Conference, October 24-28, 2005, Vancouver, Canada.

[7] Reigeluth, C.M.: A New Paradigm of ISD, 1996.