Integration of Heterogeneous Information Sources into a Knowledge Resource Management System for Lifelong Learning

Elena Demidova[1], Stefaan Ternier[2], Daniel Olmedilla[1], Erik Duval[2], Michele Dicerto[3], Krassen Stefanov[4] and Naiara Sacristán[5]

[1]L3S Research Center and University of Hannover, Hannover, Germany, E-mail: {demidova, olmedilla}@L3S.de

[2]Katholieke Universiteit Leuven (K.U.Leuven), Leuven, Belgium, E-mail: {stefaan.ternier, erik.duval}@cs.kuleuven.be

[3]GIUNTI ILABS, Sestri Levante, Italy, E-mail: m.dicerto@giuntilabs.it

[4]Sofia University, Sofia, Bulgaria, E-mail: krassen@fmi.uni-sofia.bg

[5]Altran SDB, Madrid, Spain, E-mail: nsacristan@altransdb.com

Abstract: Access to learning information is still restricted due to the lack of technical and semantic interoperability, locking in knowledge resources in disconnected islands. A successful environment for competence development and lifelong learning must be built on top of an infrastructure that maximizes the amount of information available, integrating centralized repositories and user desktop resources as well as emergent new applications. This paper describes the vision and current efforts of the TENCompetence project towards this goal as well as current collaboration with other initiatives like those being performed by ARIADNE foundation, describing the requirements and challenges towards developing such an integrative knowledge resource management system on the information source layer as well as on the service layer.

Keywords: Knowledge Management issues related to competence development, lifelong learning; Simple Query Interface; Learning Object Repositories; Learning Object Metadata; interoperability.

1 Introduction

Traditionally, instructional learning has covered the needs of people searching for new labour opportunities, be it in the form of improvements or upgrades in current jobs or as a way to be able to apply for new ones. Traditionally, the learning process would start with some formal curriculum (e.g., at the university level) and then be followed by some specific courses related to the area of work (e.g., an MBA). Nowadays, there is a growing need for more flexible and cost-effective solutions for learners in order to provide lifelong competence development. People require the possibility to learn at different locations (e.g., at home) and at times that better fit their work hours

but still keeping most of the benefits of face-to-face learning like, for example, the availability of tutors to help them if a question arises. TENCompetence [1] addresses this need and aims in supporting individuals, groups and organizations in Europe in lifelong competence development by establishing the most appropriate technical and organizational infrastructure, using open-source, standards-based, sustainable and innovative technology. Such an infrastructure requires a knowledge resource management system as basis in order to ensure the availability of knowledge resources as well as the ability to exchange and (re-)use them. Moreover, learners should be able to share their resources and access resources from other learners as well. On the one hand there exist the centralized and monolithic repositories on which e.g. Learning Management Systems (LMSs) rely while on the other extreme there are completely decentralized networks like Peer-to-Peer (P2P) networks in which users may share content without loosing control over it. In addition, emergent Web 2.0 applications (e.g., Flickr [2] or YouTube [3]) provide new means for users to share information with a higher level of motivation than the more old-fashioned repository-like. Unfortunately, the lack of standards makes the integration of all these information sources[1] difficult so that the benefit of such a rich set of knowledge resources is not always fully exploited.

Among the key activities of TENCompetence are the integration models and tools for the creation, storage and exchange of knowledge resources[2]. In order to achieve this goal we are currently implementing a base infrastructure in which information is made accessible in order to better support lifelong learning and at the same time enhance the learning experience. Such an infrastructure would bring together information stored for example in institutional servers and LMSs (centralized repositories), locally on learner desktops (by means of P2P networks) and online community-sharing systems like online-storage applications, wikis or blogs.

In this paper we identify the main challenges that must be addressed and solved before our goals can be realized and present the architecture we are currently implementing towards such goals.

2 TENCompetence Knowledge Resource Management System: Challenges & Architecture

Integration of heterogeneous information sources like learning object repositories, user's desktops and resources in Web 2.0 applications within a single architecture raises the challenge of interoperability, both technical and semantic. Due to the limitations (both in metadata and query capabilities) of some of the information sources we are considering (especially in online web systems) we have decided to start with a basic core set of metadata (a subset of Dublin Core and therefore of LOM) and rely on keyword based search therefore simplifying considerably the problem of semantic interoperability.

However, technical interoperability still remains a challenge. Our KRMS must not only provide sharing and search capabilities but also management of knowledge resources (storage, update and deletion) as well as other advanced services like for example rating. The lack of a standard interface for these services makes this integration more difficult and costly. Furthermore, TENCompetence KRMS must provide services for storage, search and retrieval of knowledge resources in both the local system or remotely, in other repositories or web storage systems. In addition, the number of remote repositories may evolve over time. Therefore the architecture must be flexible enough to deal with these issues. For this reason, we adopted a Service Oriented Architecture and rely on an online registry in which information sources can be added or removed dynamically (see Figure 1).

[1] With information sources we mean all kinds of information container e.g. repositories, P2P networks and Web 2.0 applications.

[2] Knowledge resources are the containers that store the explicit knowledge for sharing purposes. Examples are learning objects, articles, books, software programs, informal messages, etc.

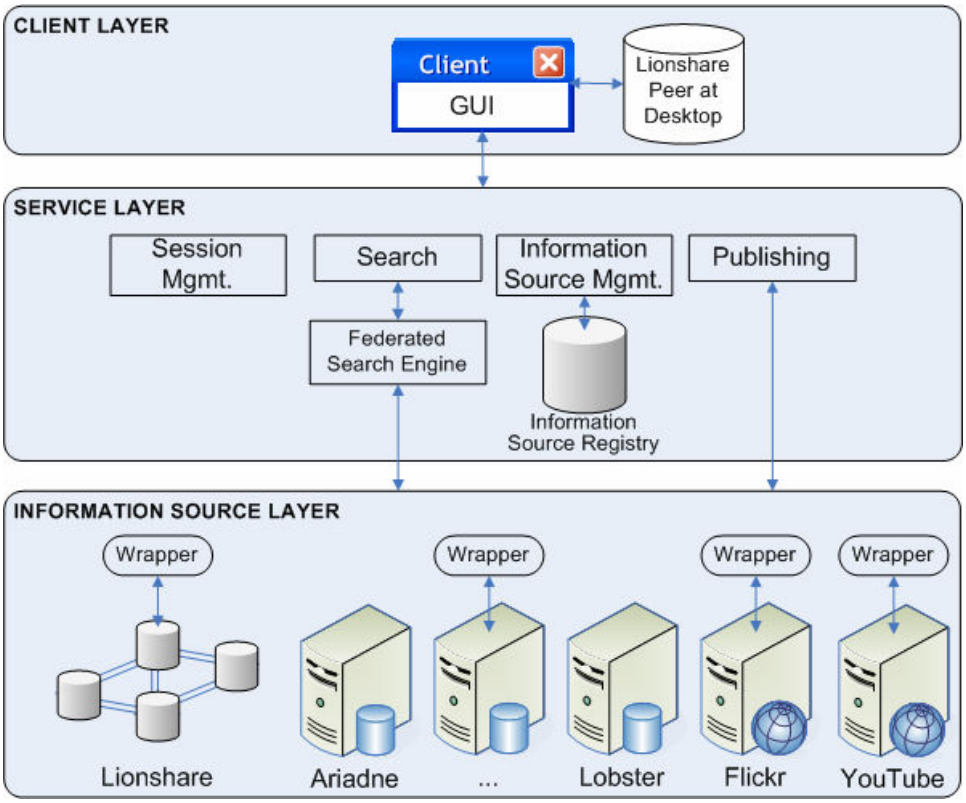

Figure 1: TENCompetence KRMS Architecture

The KRMS is built as a 3-layered architecture incorporating information source, service and client layers.

2.1 Information Source Layer

The information source layer consists of the set of repositories, P2P networks and Web 2.0 applications to be integrated in our system. Each information source to be connected must implement a common interface for the services provided. In case that the information source does not support such an interface natively, a wrapper module is required.

Interfaces for Search and Publishing

TENCompetence KRMS is aimed to include knowledge resources contained in different kinds of information sources independent of the concrete information source implementation. In order to achieve this aim, information sources should implement common interfaces. Regarding to search, there exist different standards with varying complexity [4]. We decided to base our solution on the Simple Query Interface (SQI) [5] due to its simplicity and adaptivity to different scenarios (e.g., synchronous and asynchronous) as well as independence of query language, which in fact made possible that some of the information sources we are targeting natively support it. Otherwise, an SQI wrapper installed on the top of an information source allows performing search over the network in a platform independent manner. However, there are no standards that allow for comparable homogenization of resource publishing operations. On this direction, ARIADNE [6] already initiated the work for the specification of such a publishing interface following similar

design guidelines as the ones used during the development of SQI. TENCompetence has joined this effort and collaborates on its refinement and validation in order to use it as common interface in the TENCompetence KRMS.

Components

One of the interesting approaches of our architecture is the integration of centralized repositories (like ARIADNE and DSpace [7]), distributed networks (like the LionShare P2P [8]) and online web systems. Such an approach gives flexibility to users and reduces the effort they need in order to share or find information they are interested in [9] allowing at the same time its use in different scenarios.

For example, the core of the ARIADNE infrastructure is a distributed library of digital, reusable educational components called the Knowledge Pool System (KPS) now actively used in both academic and corporate contexts. This ARIADNE KPS is a Learning Object Repository that offers a web service based API through which learning objects and their metadata can be transparently modified. The goal of having such an API is to enable loosely coupled integrations in third party applications such as VLE's, authoring tools or federated search engines.

On the other hand, LionShare is a P2P network which primary goal is to facilitate the distribution of localized content found on the personal computers of educators and researchers not having an easy way to publish this content in popular learning object repositories or preferring to keep control over their resources. The LionShare P2P network now comes with a SQI to LionShare gateway. This SQI target allows users to treat the LionShare's P2P network collectively as just another big and very distributed learning object repository [10].

2.2 Service Layer

The service layer provides search and publishing services as well as services for information source and session management. This layer contains the services that will be accessed by client applications (in the client layer). Services in this layer may also make use of other services from the same layer.

Combination of P2P and Centralized Repositories

A number of centralized repositories, like ARIADNE, DSpace, etc. were created over the last years. One of the aims of the KRMS architecture is to allow inclusion of knowledge resources already contained in these repositories. On the other hand, a lot of knowledge resources that are available on the user's desktops are unlikely to be shared by means of inserting them in a central repository. P2P technologies provide a common framework for sharing of knowledge resources stored on the user's desktops. Combination of both, P2P and centralized repositories allows for integration of the knowledge resources independently of their storage place.

Information Source and Session Management Services

In order to combine publishing services provided by different information sources within a single application, information sources should publish information about supported metadata schemas, constraints on resource formats and connection details in a registry. Our UDDI registry provides homogenized descriptions about the services of the information sources connected to it. Based on this information, a client application can select one or more information sources to work with.

Federated Search Engine

Our KRMS currently relies on the engine developed by ARIADNE in order to perform federated search. It is conceived such that it leverages standards at 2 levels. At the top level, this federated search engine offers search functionality through an SQI target. Thus from a standards perspective,

sending a query to a federation of repositories is no different than sending a query to a single repository. At the second level, this engine leverages SQI at the level of the information sources. As we want to lower the threshold of adding a new information source to the federation, adding a new information source requires no more than registering its capabilities by our information source management service. The federated search engine consults this service for the new information sources dynamically.

2.3 Client Layer

The KRMS client layer provides the graphical user interface. It may be implemented as a stand-alone application or web-based. The former has the advantage that it may allow local management and sharing of knowledge resources becoming a LionShare peer. The latter has the advantage that it may be accessed anywhere from any computer and without need for any extra software installation.

2.4 Implementation

The architecture described in this chapter is currently being implemented. At present it allows performing search in all integrated information sources like ARIADNE repositories, eXact Lobster [11] and LionShare P2P network. Ongoing work is directed towards the implementation of the SQI-based search functionality for Flickr and YouTube. Apart of that we integrated a preliminary version of the publishing interface into the ARIADNE repository.

3 Conclusions and Further Work

This paper identifies the requirements a KRMS system should satisfy in order to integrate heterogeneous information sources and provide homogenized search and publishing services for knowledge resources. This is a basic backbone over which a full environment for competence development and lifelong learning will be built. Architecture for such an integration that has been presented makes use of existing standard interfaces for search (SQI). However, lack of standards for publishing and information source management services led the TENCompetence project to join efforts initiated by ARIADNE in this direction and to collaborate on the refinement and validation.

Although the results presented here are promising, we are still at the beginning of the process. Refinement and validation of the interfaces as well as the integration of emergent web online storage applications must be performed. Finally, the whole infrastructure will need to be validated as part of the overall infrastructure for the competence development and lifelong learning.

Acknowledgements

The work on this paper has been sponsored by the TENCompetence Integrated Project that is funded by the European Commission's 6th Framework Programme, priority IST/Technology Enhanced Learning. Contract 027087 (www.tencompetence.org)

References

[1] TENCompetence, http://www.tencompetence.org

[2] Flickr, http://www.flickr.com/

[3] YouTube, http://www.youtube.com/

[4] IMS query services white paper.
http://www.imsglobal.org/query/imsQueryServices.html, June 2005.

[5] F. van Assche, E. Duval, D. Massart, D. Olmedilla, B. Simon, S. Sobernig, S. Ternier and F. Wild. 'Spinning interoperable applications for teaching & learning using the simple query interface', Educational Technology & Society. Special Issue (April 2006) on Interoperability of Educational Systems, 9(2):51–67, 2006.

[6] Ariadne Foundation, http://www.ariadne-eu.org/

[7] DSpace, http://www.dspace.org/

[8] LionShare, http://lionshare.its.psu.edu/

[9] I. Brunkhorst and D. Olmedilla. 'Interoperability for peer-to-peer networks: Opening p2p to the rest of the world', in Innovative Approaches for Learning and Knowledge Sharing, First European Conference on Technology Enhanced Learning (EC-TEL), volume 4227 of Lecture Notes in Computer Science, pages 45-60, Heraklion, Greece, Oct 2006. Springer.

[10] S. Ternier, B. Bosman, E. Duval, L. Metzger, M. Halm, S. Thorne, and J. Kahn, 'Connecting OKI and SQI: one small piece of code, a giant leap for reusing learning objects', in proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2006 (Pearson, E. and Bohman, P., eds.), pp. 825-831, 2006.

[11] eXact Lobster, http://www.giuntilabs.com/info.php?vvu=31