Editorial

Salvador Sanchez-Alonso and Enayat Rajabi, Guest editors

Welcome to the Bulletin of the IEEE Technical Committee on Learning Technology, Volume 17, Number 1-2, April 2015 issue. This special issue discusses the current learning technology challenges towards an open discovery space for education. The papers presented in this issue are a collaboration among the partners of an important European project, called Open Discovery Space (http://portal.opendiscoveryspace.eu/). This project aims to serve as an accelerator of the sharing, adoption, usage, and re-purposing of existing educational repositories, by involving thousands of European schools in innovative teaching and learning practices through the effective use of eLearning resources. During this project, lots of technical approaches were applied for collecting, searching, exposing, creating and using a large number of eLearning resources.

The first paper in this issue presents a web-based collaborative learning environment, COLearn, that supports the specification of collaborative learning workflows, their deployment and enactment. It addresses major challenges in modern learning infrastructures to enable contextualization, social constructivism and knowledge-pull. It also offers a shell, easily integrated on top of existing open learning infrastructures (such as LMSs and OER repositories) to enrich their capabilities by offering functionality to design rich collaborative learning activities. COLearn employs an intuitive graphical representation exploiting the Business Process Modeling Notation standard. As an internal representation and interoperability model it uses IMS Learning Design thus offering effective sharing and remixing of learning designs.

In the second paper, the authors describe the current landscape of searching learning resources and show the multiple challenges teachers meet. The search tool of the Open Discovery Space project is presented in this paper. This tool combines several features to answer these issues, including a significant amount of resources (almost one million at the time of writing). The authors also described several issues they faced during the development of search tool along with their proposed approaches.

The third paper presented a vocabulary bank implemented during the Open Discovery Space project lifetime. The authors implemented it as a single point of reference for all authoritative sets of terms, concepts and named entities as well as the network of relations between them, using an innovative, semantic web approach. Their work also provides an overview of the software requirements and the selected implementation approach for developing the vocabulary bank.

In the forth paper, the authors present the aggregation workflow followed in such open educational environment. They describe the procedure in which a large number of eLearning metadata were collected from several learning repositories in the Open Discovery Space project. The collected metadata were later processed, cleaned, evaluated and finally imported into the learning portal. The applied workflow described could be considered as a generic approach for covering most of the learning object’s metadata processing needs.

The fifth paper proposes the usage of an integrated search Moodle in the Open Discovery Space project. The authors describe a method for searching and exposing learning objects from Moodle. Factors which will be taken into consideration include the ease of use from course creators and Moodle administrators point of view, as well as how well the generated metadata match user requests and the impact of the harvesting process to the Moodle installation. The paper also suggests the integrated ODS OAI-PMH Moodle block plugin that automatically checks all available learning objects inside each Moodle course.

We sincerely hope that the issue helps in keeping you abreast of the current research and developments in Learning Technology. We also would like to take the opportunity to invite you to contribute your own work to the Bulletin of the IEEE Technical Committee on Learning Technology if you are involved in research and/or in the implementation of any aspect of advanced learning technology. For more details, please refer to the author guidelines at http://www.ieeetclt.org/content/authors-guidelines .

COLearn: Supporting Collaborative Learning on top of Existing Learning Infrastructures

George Stylianakis, Nektarios Moumoutzis, Polyxeni Arapi, Manolis Mylonakis, Stavros Christodoulakis
Laboratory of Distributed Multimedia Information Systems and Applications
Technical University of Crete (TUC/MUSIC)
Chania, Greece
{gstylianakis, nektar, xenia, manolis, stavros}@ced.tuc.gr

***Abstract*—COLearn is a web-based collaborative learning environment that supports the specification of collaborative learning workflows, their deployment and enactment. It addresses major challenges in modern learning infrastructures to enable contextualization, social constructivism and knowledge-pull. It offers a shell, easily integrated on top of existing open learning infrastructures (such as LMSs and OER repositories) to enrich their capabilities by offering functionality to design rich collaborative learning activities. COLearn run-time environment is used to enact these activities and workflows. This way, COLearn leverages the power of the underlying infrastructures, provides structure to the groups of learners that participate in learning workflows, dynamically adapts the workflows during their enactment, monitors their evolution to facilitate assessment and provides feedback to the learners. COLearn employs an intuitive graphical representation exploiting the Business Process Modeling Notation standard. As an internal representation and interoperability model it uses IMS Learning Design thus offering effective sharing and remixing of learning designs.**

***Index Terms*—Collaborative work, Distance learning, Social Computing, Workflow management software**

I. Introduction

Learning Management Systems (LMS) have been proven successful in the administration, documentation, tracking, reporting and delivery of learning contents of different granularities (learning resources, learning objects and courses). Despite the unquestionable impact of LMSs on the evolution of e-learning and education in general, there is room for improvement with respect to several aspects. COLearn addresses several issues related to the desired improvements on:

Contextualization: A traditional LMS follows a static approach to learning by offering uniform courses to learners. Our approach, in contrast, takes into account the design decisions of educators allowing them to organize learning materials, tools and learners within meaningful contexts (learning activities) that are structured within wider designed learning spaces (learning scenarios).COLearn provides a graphical authoring tool with which educators (acting as educational designers) can describe these learning spaces (learning scenarios) in the form of collaboration scripts[13].Each learning space can be reused or remixed, thus offering to educators the possibility to share their pedagogical strategies with other educators.

Social Constructivism: LMSs restrict opportunities for collaboration in learning and for the promotion of social constructivism. In our approach learners could be engaged in a distributed environment consisting of a network of people, services and resources. To provide such an environment COLearn platform at the run-time level incorporates specific tools for the support of group management and social interaction among participants, tools such as chat, micro-blogging, calendar etc. These tools are employing the Extensible Messaging and Presence Protocol (XMPP)[14]implementing the publish - subscribe communication pattern.

Knowledge-pull: Traditional LMSs adopt a knowledge-push framework. In our approach, however, we permit the educator to adopt a knowledge pull model. Educators could create their customized learning environments -at the authoring level- in the form of organized learning activities and -at the run-time level- those learning activities will guide the participants to pull knowledge that meets their particular needs from a wide range of knowledge sources. If the educator wishes to enable learners become creators of content, it is possible to incorporate content creation tools as external learning tools. In order to cope with the integration of external learning tools in a transparent way, we employ the IMS Learning Tool Interoperability (IMS LTI) [9]specification.

COLearnis designed to be easily integrated with LMSs in order to address the aforementioned challenges and offers a technological solution that enriches the underlying infrastructure and exploit its power. New services are offered (a) to educators to embrace a culture of open educational practices; and (b) to learners to engage in meaningful learning within the context of rich learning activities supported by digital resources and tools.

The rest of this paper is organized as follows: Section 2 presents the related work. Section 3presents the COLearnarchitecture and implementation. Section 4 presents, as a case study, how COLearn has been aligned with the Open Discovery Space portal to enable collaborative workflows on top of a teachers' social network and a repository of Open Educational Resources. Section 5 concludes and presents future plans.

II. Comparison with Related work

The most known platform that combines authoring and run-time aspects for learning scenarios is LAMS [10].LAMS is a complete platform for designing, managing and delivering online collaborative learning activities. LAMS is not intended to complement an existing learning environment as it is the case for COLearn. Activities in LAMS can include a range of individual tasks and group work. LAMS functionality is extended through the integration of external tools. However, to integrate an external tool one should cope with a complex LAMS Tool contract[1]. Despite the fact that activities can be assigned to groups of participants, LAMS lacks synchronous group support. LAMS provides some synchronous communication tools during the learning flow but these tools are initiated as activities as long as the author explicitly defines their use in the scenario. Also, these tools are not constantly active during the scenario execution. A teamwork environment should account the absence of physical interaction by providing synchronous collaboration among participants[11].

SLeD[12] and Reload Player[4]are run-time environments that enable the enactment of IMS LD scenarios. They both lack group coordination and collaboration. In addition, there are no communication tools available during scenario enactment. In contrast, the COLearnrun-time environment supports group-based learning. The learning scenarios are performed by group of users in a collaborative and synchronized manner. To enable synchronization and personalization of the learning process, the CopperCore engine is integrated to the COLearn platform. Furthermore, the adoption of the XMPP real-time protocol offers to the COLearn platform a messaging system to notify group members about the changes in the progress of the scenario in real-time. XMPP-based tools are provided by the COLearn platform to support synchronous and asynchronous collaboration/ communication. Note that XMPP can provide asynchronous communication as well. Messages addressed to absent users are delivered to them when they enter the platform. External learning tools are integrated to the COLearn platform in a transparent way by implementing the IMS LTI specification.

III. COLearn Architecture

COLearn consists of two main components as depicted in figure 1: a) the authoring tool that provides the means to specify the learning workflows and b) the run-time environment where theses workflows are enacted.

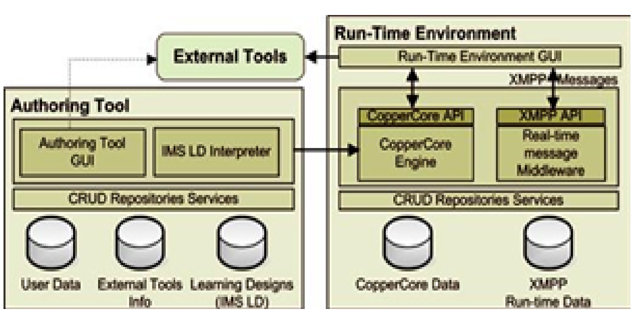

Fig. 1. *COLearn Architecture*

A. Authoring Tool

In order to develop and offer collaborative learning facilities on top of an open learning infrastructure one has to cope with learning content, communication/ collaboration facilities, and run-time execution (i.e. learning process). Collaboration scripts can specify the components as well as the orchestration of the learning process. COLearn offers an authoring tool to describe collaboration scripts as learning workflows, based on the Business Process Modeling Notation(BPMN)[3]. BPMN is the model for the graphical representation of the learning workflows. The COLearn authoring tool, as seen from the user, is presented in Fig. 2.

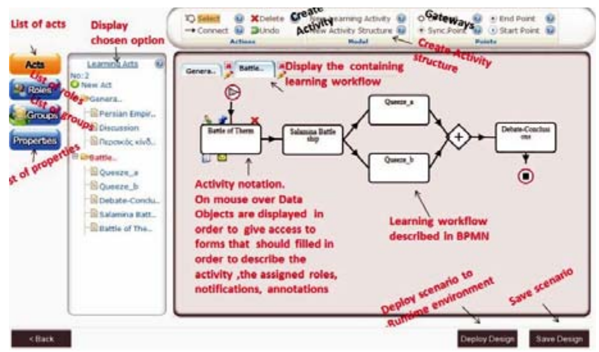

Fig. 2. COLearn Authoring Tool

Learning workflows are parsed and transformed to IMS Learning Design Level C[8]. Details regarding the internals of the Authoring Tool have been given in Stylianakis et al.[6] along with the rationale for differentiating[5] the user interface representation model from the internal representation model and their mapping.

B. Deployment of collaboration scripts

The deployment of a collaboration script refers to the identification of the participants and their roles for enacting it. This is a two-step process: Initially, the IMS LD document is published to the CopperCore engine. Then, the appropriate communication channels among the members of each group are established. The latter, is based on the XMPP protocol and made feasible by integrating the Openfire Server into COLearn run-time environment. The corresponding services and strategy to deploy collaboration scripts which describe learning scenarios, is presented in Stylianakis et al[6].

C. Run-time Environment

Enactment is a term that refers to the actual implementation of a collaboration script through the COLearnrun-time environment after its deployment, i.e. after the assignment of scenario-specific roles to particular users that will participate in the scenario implementation. Consequently, to enact a script it is necessary to deploy it first. Note that each scenario could be deployed many times. Each distinct deployment corresponds to unique implementation of the scenario with specific participants. The COLearnrun-time Environment consists of two main components: a) The CopperCore engine[7] and b) the Real-Time Message Middleware. These components are described in the next paragraphs. Figure 5 presents the COLearnrun-time environment and displays its functionality.

Fig.3. COLearn run-time environment

a) Copper Core Engine

CopperCore engine provides coordination support for learning processes. Checking, synchronization and personalized view by each participant is supported by the CopperCore engine. The engine fully supports IMS LD levels A, B, C and provides the persistence storage and the APIs to manage the administration and delivery of IMS LD. In addition COLearnrun-time environment incorporates the appropriate graphical components to exploit the CopperCore functionality and deliver the IMS LD learning scenarios.

b) Real-time Message Middleware / Social Interactions

COLearn incorporates Openfire server, which is a real-time collaboration server that implements the XMPP protocol. XMPP is an open technology for implementing real-time request response services. XML messages can be transferred from one entity to another in real-time. Additionally it provides the establishment of peer-to-peer media sessions supporting interactions such as voice/video chat or file transfer.

COLearn adopts the following strategy for supporting collaboration: When an author deploys a collaboration script to the run-time environment, a contact list is created for each member of each specific working group. The contact list is essentially, an instant messaging "friend list" providing the option for one-to-one messaging among members of the same group. Additionally, for each group a communication channel implemented as publish-subscribe node is established. Each member of the group is subscribed to this node. The XMPP publish-subscribe extension provides a framework to implement notification, multi-party messaging and file transfer in the COLearn platform. In addition, it offers an extensible approach to integrate tools compliant to instant messaging mechanisms for the organization/ management of working groups as well as the social interaction among group members. Finally, the message middleware interacts with the CopperCore engine, providing information and notifications among the users and triggering actions in the learning workflow process.

c) Using IMS-LTI to integrate external learning tools

COlearn integrates external learning tools by implementing the IMS LTI specification. In COLearn run-time environment, each external tool is launched as a service employing IMS-LTI. In the COLearn authoring tool, an author can choose to use an external tool from a list of available ones. This list is maintained/ updated by the administrator of the system. The administrator provides the URL of the external tool, a secret and a key that are needed in order to establish the connection. The secret and the key parameters are needed because the connection of the COLearn platform with the external tools (at the run-time level) is using the Open Authentication protocol. When the author selects the desired external tool, she additionally submits the IMS-LTI specific attributes described in the specification. At the run-time environment, when an activity which defines external tool usage is activated, then the tool is automatically launched.

IV. Alignment with Open Discovery Space Portal

Following the presentation of the overall architecture for enabling learning scenarios on top of existing open learning infrastructures, this section presents a case study that exemplifies the details of the approach. This is the case of the Open Discovery Space (ODS) community portal targeting teachers in European countries. The portal provides a common access point to many national and international repositories of open educational resources through its federated repository. The main functionalities offered through the alignment of the ODS portal with COLearn are the following:

Automating the navigation between the ODS portal and COLearn to support ODS communities' members in creating, sharing, finding and enacting learning scenarios by the usage of the COLearn authoring tool and run-time environment.

Exploiting ODS harvested resources to enable their inclusion in learning activities managed by COLearn. This is done through a special pane implemented in COLearn that provides search functionality on top of the SOLR search engine.

Assign ODS community/group members with COLearn scenario-specific roles to enable deployment of collaborative learning designs (scenarios or lesson plans in the ODS terminology) using information from ODS portal about its users.

To make the description of these functionalities more concrete, let us present them within the context of five specific use cases that reflect the actual user needs addressed by the aligned ODS-COLearn environment:

1. Creating a new scenario or lesson plan

The creation of a learning scenario is triggered from the ODS portal interface where an ODS user is able to see a list of existing scenarios and create new ones. This list of scenarios is provided through a web service offered by COLearn. When the user selects to create a scenario, the COLearn authoring tool is triggered and the user automatically logs in. This underlying authentication mechanisms is based on OAuth[2].

2. Editing/Cloning an existing scenario or lesson plan

Learning scenarios can be further edited and copied across

ODS communities. To do so, the ODS user can select a specific learning scenario from the list of available scenarios within a specific community and trigger the COLearn authoring tool. Then, within the authoring tool, the user is able to edit the scenario. It is also possible to clone it to make a copy of it to another ODS community that the user participates. Special services are employed for this process that provide information about the ODS communities where the user participates.

3. Searching the ODS repository

During the editing of a learning scenario in the COLearning authoring tool, the user is able to trigger a special pane that exploits the SOLR search engine service of the ODS federated repository to present a list of qualifying resources based on a faceted-search approach. Finally, the user can select the resources that are appropriate for inclusion in a certain learning activity. Resource links are then retrieved from the ODS repository and stored as references within the learning activity elements' metadata.

4. Deploying a scenario or lesson plan

Deploying a scenario or lesson plan means to identify the participants that will take part in an enactment of the scenario or lesson plan and assign roles to each of them so that the enactment can take place in the CoLearnrun-time environment. The deployment process in the aligned CoLearn environment exploits information from the ODS portal regarding the members of the community/group within which the scenario to be deployed is available.

5. Enactment of a scenario of lesson plan

Enactment is a term that refers to the actual implementation of a scenario or lesson plan through the CoLearn runtime environment after its deployment, i.e. after the assignment of specific scenario-specific roles to a set of users that will participate in the scenario implementation. Consequently, to enact a scenario or lesson plan it is necessary to deploy it first. Note that each scenario could be deployed many times corresponding to different implementations with different users participating each time.

V. Conclusions andFuture Work

We presented COLearn, a platform acting on top of an existing open learning infrastructure to enrich its capabilities. An infrastructure (e.g. an LMS) that uses COLearn, could extend its functionality by offering to teachers the ability to design rich learning activities and learning workflows, enact them, and provide structure to the groups of learners that participate in these workflows, dynamically adapt the workflows during their enactment and monitor their evolution to facilitate assessment. The COLearn authoring tool produces collaboration scripts employing the BPMN model for their graphical representation. These scripts are internally represented using the IMS LD specification. The run-time environment consists of the CopperCore engine, which supervises the execution of the design. The group management, social interaction and collaboration among the participants are made feasible by integrating the Openfire server which implements the XMPP real-time protocol. For the integration of external learning tools, COLearn implements the IMS LTI specification. The COLearn platform is an open platform that can collaborate with different types of internal/external learning applications. As a case study we presented the integration of COLearn with the Open Discovery Space (ODS) community portal. The alignment is targeting to support ODS communities' members in creating, sharing, finding and enacting learning scenarios, to exploit ODS harvested resources

Our future plans in enhancing COLearn services focus on supporting mobile learning scenarios including support for augmented reality games. To enable such scenarios it is necessary to model interactions with tools running on mobile devices and incorporating parameters for the specification of physical contexts on which learning activities take place. The learning domains that will be facilitated with these extensions include field trips for science projects and physical inquiries related to history, geography and language learning.

Acknowledgment

This work is partially funded in the scope of the Open Discovery Space project - http://portal.opendiscoveryspace.eu (FP7-ICT-PSP: 297229).

References

[1] C. Alario-Hoyos, M.L. Bote-Lorenzo, E. Gómez-Sánchez, J. I. Asensio-Pérez, G. Vega-Gorgojo, and A. Ruiz-Calleja, "GLUE!: An Architecture for the Integration of External Tools in Virtual Learning Environments." Computers & Education (2012).

[2] Barry Leiba, "OAuth Web Authorization Protocol", IEEE Internet Computing, vol.16, no. 1, pp. 74-77, January/February 2012, doi:10.1109/MIC.2012.11

[3] BPMI.org, OMG: Business Process Modeling Notation Specification. Final Adopted Specification. Object Management Group, 2006, available at: http://www.bpmn.org

[4] D. Griffiths, P. Beauvoir, O. Liber, and M. Barrett Baxendale, "From Reload to ReCourse: learning from IMS Learning Design Implementations." Distance Education, 30, no. 2 (2009): 201-222.

[5] D. Sampson, P. Karampiperis and P. Zervas, "ASK-LDT: A Web-Based Learning Scenarios Authoring Environment based on IMS Learning Design", International Journal on Advanced Technology for Learning (ATL) (ISSN 1710-2251), vol. 2(4), pp. 207-215, ACTA Press, 2005

[6] G. Stylianakis, P. Arapi, N. Moumoutzis,S. Christodoulakis, CoLearn: Real time collaborative learning environment. International Conference on e-Learning and e-Technologies in Education (ICEEE) (2013)

[7] H. Vogten, H. Martens, P. Rosmalen, and E. J. R. Koper, CopperCore. Retrieved January14, 2005, from SourceForge: http://coppercore.org

[8] IMS Learning Design Specification Version 1.0. [Online] Available at:. http://www.imsglobal.org/learningdesign/index.cfm IMS

[9] IMS Learning Tools Interoperability Version 1[Online] Available at http://www.imsglobal.org/lti/blti/bltiv1p0/ltiBLTIimgv1p0.html

[10] J. Dalziel, "Implementing learning design: The learning activity management system (LAMS)." LAMS Teacher's Guide, 2006, V2.0.

[11] M. N. K. Boulos, A. D. Taylor, and A. Breton, "A synchronous communication experiment within an online distance learning program: a case study." Telemedicine Journal & e-Health 11 no. 5(2005)

[12] OUUK. Sled player (2005) (cited April 8th, 2007)

[13] P. Dillenbourg, "Over-Scripting CSCL: The risks of blending collaborative learning with instructional design." Three Worlds of CSCL. Can We Support CSCL? (2002): 61-91.

[14] XMPP Standards Foundation at http://xmpp.org/

Searching Thousands of Learning Resources

Paul Libbrecht, Christoph Bail, and Martin Becker,
Weingarten University of Education, Germany

***Abstract*—Everyone is a publisher nowadays on the web, teachers included. The social web is growing at a fast pace, with more and more needs to *web* the content so as to make it accessible, annotated, and found. While this webbing creates a rich navigation experience, tools to access more resources are needed. Search tools are among the most used tools to discover more learning resources. However, their usage is currently rather limited and frustrating. This paper describes the challenges currently met, and the Open Discovery Space search tool, presenting how it addresses them.**

***Index Terms*—Education, search engines, web search, learning resources.**

I. Introduction

THE world wide web, having started with a model where few were in possession of the publication privilege, has evolved into the largest collaboration space of the humanity, where almost anyone is a publisher. Teachers have also taken part to this exchange: Today, multitudes of learning resources are available published by a multitude of authors, offering hints, materials, or advices to support the teaching or learning process. But how to find orientation among this multitude?

The search challenges of a teacher that “resources himself” are enormous: The learning resources are scattered across multiple sites; they employ slightly different vocabularies to describe themselves and thus to be identified; while a few teachers are able to trust a few resources, many do not. As a result it is common for teachers to spend repeated search sessions in preparing their courses, trying to identify the resources they would be able to adopt.

The Open Discovery Space portal is a large portal that harvests multiple learning object repositories within a single point of access:: with its massive amount of learning resources and indexed with a unified vocabulary, it opens the door for teachers to search through a significant amount of learning resources with a structured support. While text-search remains a central element, a structured drill-down by means of facets and taxonomies allows to remove ambiguities in search terms. Moreover, an ordering of the search result is offered that promotes resources judged relevant to the user by criteria such as the language affinity, the readiness to embrace complex file types, or the recommendation of a friend in one’s network.

Manuscript received May 4th, 2015. This work was supported in part by European Commission CIP PSP 297229.

Paul Libbrecht, Christoph Bail, Martin Becker, Weingarten University of Education, Informatics. Kirchplatz 2. 88250 Weingarten. Germany.

A. Outline

In this paper we first present an overview of existing search engines that are applicable to learning resources, we the survey current challenges that these meet (including ambiguity, multilinguality, and implicitness). The design of the search engine of Open Discovery Space is then presented. An implementation status and testing plan follows in the future works, along with open questions.

II. Search Engines for Learning Resources

Learning resources are *artifacts* that can be employed within learning and teaching processes; this very broad definition is generally endowed with one common restriction: they are *digital documents*, in the sense that they are materialized as files that can be viewed and potentially edited by a computer.

Textbooks form probably the most ubiquitous example of learning resource; they are sometimes digital. Textbooks commonly support the learning processes by guiding teachers and students, supporting their exploration and assessment. Textbooks illustrate well the *resource* nature: they can be *pulled from* in the learning. However, they are often not open.

Open Educational Resources are understood to be digital and digitally exchangeable thanks to their digital natures and thanks a license that allows anyone to receive and redistribute the resource without cost; this definition is that of the Hewlett Foundation [1]. The wave of open educational resources has grown since about a decade and has allowed the production of millions of resources, which could be applied by any teacher of the earth.

The breadth of this availability presents to many of the teachers of the earth a sea of available resources, which teachers can choose from. This makes the identification of learning resources that are relevant for the teachers’ or learners practice and are of sufficient quality, quite a challenge. Among elements of the sea, one can find learning resources which are not complete, lack adaptability, employ outdated tools, use an inappropriate vocabulary, guide the teacher only partially in his attempts, or employ incompatible software. All these issues have to be recognized and coped for and thus it is important to be able to crawl among multiple resources to elect the *most appropriate* or one *requiring the least circumventing actions for it to become appropriate*.

Search tools are among the most important tools to crawl this wealth. They are commonly used to find learning resources but in a way that is not yet fully satisfactory in many cases. We describe a few typical search tools used to date to find learning resources which can be used by teachers in

Europe.

A. *Generic Web Search Engines*

Learning resources are commonly found on the web. Thus, they can generally be fetched by web crawlers and searched for by the search engines behind the crawlers. These search engines search the broad web and thus almost only offer to employ *generic* queries; that is, queries for words in the current language, or in any language. If the normal language had special words for most topics to be learned, just as brand names are unique, the search process for learning resources would be very effective and convergence of users searching within similar topics would be found. However, the *ambiguity* is unavoidable and topic names such as *inflation*, *square*, or *reading* are not able to distinguish learning resources from other artifacts. In general, additions of query terms often bring more noise in the search results.

The teachers searching for learning resources in the broad web thus often have to allocate long repeating periods to search for learning resources, sometimes going through pages and pages of results in the hunt for a more satisfactory resource. While web search engines provide the openness that most teachers want, hoping to find open resources anywhere, and often ready to decrypt resources in a language that is not theirs, they fail to provide an enjoyable search process because of the inability to formulate queries: topic queries are difficult as expressed above, but querying for an educational level is difficult too: There does not exist a uniform query mechanism for the very many school systems and the typical age range is rarely accessible.

To our knowledge, thus far, no web search engine is leveraging the emerging metadata standards LRMI[1] to offer such a refined query mechanism.

Moreover, the broad web is made of multiple web pages that mix several languages. Thus, it is rather common to find pages in a different language than the one queried, even if limiting the results to a single language. This happens most frequently when searching for words common in several languages, the presented results appear unrelated to the query.

Thus we contend that generic web search engine pose an *implicitness problem*: they do not allow to express criteria for learning resources that reflect fine grained expectations of learning resources (such as the quality criteria, the technical affinity, or the community of practices' membership).

B. *Personal Search engines*

Collecting a repertoire of learning resources is among the typical activities of the professional teachers. However, because they are generally not downloaded in a single place, the regular search engine of a teacher's computer is not able to search through complete collections of learning resources that teachers meet. Such a resources' collection would be a good place to collect trusted materials, which are known relevant to the user. This could avoid the implicitness problem, but would still need to be enrich-able through a discovery.

C. *Platform Search engines*

Platform search engines generally offer the search for learning resources contributed by the users of the platform. They leverage an information set that is asked at each contribution, using a vocabulary that is agreed upon at the design of this platform. Such vocabularies are often specific to a platform: While on i2geo.net, a platform to share dynamic geometry, the topics are fine grained mathematical concepts, those of other portals are often coarse. While several portals qualify the didactical function of a resource (e.g. being an exercise, an assignment or a scenario), others do not.

This specificity supports well the community's implicit values but make it difficult for new users to start using a new platform's search engine.

III. Design of the Open Discovery Space Search

The Open Discovery Space resources' search engine attempts to address these issues by several measures, which are made possible by the control it exercises on the learning resources it presents.

The Open Discovery Space search engine is a search tool embedded in the ODS portal, it is expected to be used by the users of the platform within such tasks as the generic search for learning resources, the selection of learning resources to be included in broader scenarios, the browsing of learning resources to explore the set of available resources.

A. *Content being searched*

The ODS resources' search is mandated to search through learning resources harvested from identified repositories, a broad set of repositories relevant to school education. The information about the resources is fetched during the harvesting cycles, which employ the OAI-PMH protocol, which collects Learning Object Metadata records (LOM) of each of the repositories, encoding using the vocabulary of ODS These learning resources cover most domains of school learning with inequalities in size (e.g. within physics, astronomy is richly supported, but ballistics is much less) and in languages. This diversity is somewhat similar to the broad web: for some subjects, there are far too many resources, for some subjects, only a handful.

The search engine is designed to search for resources for text queries as well as for more fine grained metadata facets such as: educational level, typical age, date, language, or learning resource type.

B. *Multilinguality*

The search tool is designed to be multilingual: currently, resources are in more than 23 languages with very different counts.

[1] The Learning Resources Metadata Initiative (LRMI) proposes a definition of a small set of *microdata* annotations, which can be encoded within web-pages for crawlers to consume. This allows information about learning resources to be harvested from any web page.

The multilinguality challenge that teachers meet is resolved by employing content sources whose language is explicitly qualified and by employing a user interface where changing the language is as simple as a click. This allows the application of classical stemming mechanisms at indexing time and at query time, which are generally not applied in a safe fashion otherwise:

- At indexing time, the LOM records being exchanged within the harvesting mechanism make sure that each text that is not a person's name (titles, descriptions, tags…) is surrounded by an element that carries an `xml:lang` attribute. The indexing process converts the words of these texts to tokens in separate fields, which employ different *tokenizers*[2]. For each such text, three versions are converted: the whitespace tokenizer preserves full words, the stemmer converts words to their roots, while the phonetic tokenizer converts words to their phonetic equivalent. It is important to note that learning resources often have multiple languages.
- At query time, the same tokenization processes apply but they are given a different weight. Thus a search for the word *directing*, queried in English, will prefer documents that contain this word, while still brining in the search results, documents containing such words as *direct* or *direct*.

This simple query system allows a fairly tolerant search, as it allows, for example, to query a singular word and still obtain documents containing plural forms, while still avoiding as much as possible the confusion of search matches between different languages (e.g. matching the French *directe*).

Moreover, the search tool supports the multilingual users in a limited fashion: they can easily change language so as perform the same query in a different language. While web browsing tools easily allow the formulation of multiple languages, for example by adding supported languages in the web-browsers' preferences, it has been the experience of the authors that users easily forget about these settings and express poor search results quality feelings, whereas an adjustment of their preferences may have made the difference.

C. *Ambiguity*

To cope for the difficult challenge of imprecise concepts denomination, the Open Discovery Space project has defined a taxonomy of topics relevant to schools, which extends the classification of the EUN LRE.[3] Contrary to this classification, the taxonomy that Open Discovery Space has introduced is a fine grained classification which allows to express fine grained topics as fine as *planet inflation*. Such taxonomy is not always available within resources where some contributors are happy to just indicate that the resource is part of *algebra*, for example. The different granularity of the topics annotations, while they are not completely parallel, can cope with each other, since they use the same vocabulary.

This approach solves the ambiguity problem, because it forces the contributors and searchers to use the same taxonomy terms for topics, that are close to each other. While the search users are sometimes lost in browsing such a taxonomy, for example many teachers have a difficulty to figure out, that *history* is considered to be part of *social sciences*, they can explore the annotations space by interactive searching. A sample strategy to do so can be to employ text search to find typical subjects and observe the topics that could still be used to refine the search (e.g. searching *queen Elizabeth* and observing that *social studies* is among the facets which represent topics of resources which would be relevant).

D. *Implicitness*

The implicit expectations that users have from a search engine are elicited from the users mostly through their user-profile: this set of information about the user is first entered at registration, then incrementally refined as the users progress, e.g. through request prompts. They include the users' language, country, and ICT competencies. Based on this information, one can assert a probable preference for resources of a particular language, or for resources that involve more or less technical competencies in the use of ICT.

This preference is encoded within the same process that converts the searched text to tokens: the query expansion decorates the queries with *preferring queries*, which change the weight of resources matching particular patterns.

The search tool could even employ the social network created by the user in his or her interactions with other users; indeed the platform supports building a network of followers and it would be thinkable to prefer resources of users in one's own network, or prefer less resources, which network members have rated negatively. However, this feature has been left as a plan.

E. *Synthesis Example*

When searching for the words *invade France* using the English language, the query expander converts these words to the mandatory query part:

```
+(title_ws:(invade france)^60 title_en:(invad
   franc)^40 title_phon:(INVD FRNS)^10)
+text_ws:(invade france)^30 + text_en:(invad franc)
   ^25 text_phon:(INVD RFNS)^12 )
```

enriched by the preference parts, in case of a fairly low ICT-technical competency profile:

```
language:en^1.5 (cTyp:image/*
 cTyp:application/ppt^0.9)
```

This example uses the query-parser syntax[4] and indicate the weights (superscript) and wildcards (followed by star).

F. *Implementation Status*

At time of writing, the search engine of Open Discovery Space is in an alpha stage and can be reached at from portal.opendiscoveryspace.eu. Its current weaknesses include

[2] The *tokenization* process is generally understood to be the conversion process from strings of characters to streams of tokens. It is described in [2].

[3] The LRE Thesaurus is a classification of topics for learning resources realized by EUN for the LRE Resource Exchange platform. It aims at representing a broad spectrum of topics without going too much in details.

[4] The Lucene query parser syntax is documented here: https://lucene.apache.org/core/2_9_4/queryparsersyntax.html.

a shallow control on the metadata quality (e.g. mixing of languages or lack of topic metadata), and the adaptivity to technical competencies, as it has not been sufficiently tested to be deployed. The server code is available open-source from http://github.com/OpenDiscoverySpace/.

Basic testing of the search tool has shown simple errors in the multilinguality: among others, searching for simple words such as the German word *verstehen* (understand) yields multiple resources, which do not seem to contain that word. This is explained by multilingual keywords on a mono-lingual resource. Similarly incomplete stemming has been met. The solutions sketched in this paper, once applied strictly, will solve these issues.

This makes the ODS search engine a unique point of access to query for learning resources among a vast pool and using a query vocabulary that is far more precise than that of generic web search engine.

IV. Conclusion

In this paper, we have described the current landscape of searching learning resources and have shown the multiple challenges normal teachers meet. We have described the search tool of Open Discovery Space, a tool which combines several features to answer these features, including a significant amount of resources (more than 900'000 at time of writing). Basic testing is currently showing issues which seem easily solvable, however, it is not yet clear, if more issues will appear. The following sections describe ongoing and proposed future works related to this search engine.

A. Testing

In order for the search tool refinements to be grounded on tangible quality criteria, the project’s last steps will include the involvement of a wide range of search testing experts. Cultivating the diversity of teachers in Europe, it will enroll experts in their fields of teaching, which will propose queries and evaluate the search results’ list. Based on a simple bookmarklet approach which any web user can activate, the voluntary users will provide their feedback while they work, assessing the quality of a search result by simple check-boxes and comment boxes injected within the page. This should allow us to gather test suites and evaluate retrieval in a quantitative manner using approaches such as those described in [3]. The feedback will then be used by developers to identify weaknesses and tune data filtering and weighting.

We expect such feedback as the inappropriate appearance of a resource in the results, the buggy ordering among the results, or the lack of particular resources among the search results. These can be answered by an analysis of the metadata being searched through and the intermediate query processing steps. It may lead to adjustments in the processing of the harvested data, e.g. applying natural language processing techniques to support (semi-)automatic classification, in the relative weight of queries (in particular preferring queries), or in the stemming methods.

Having a significant and culturally diverse test base is probably the only method to refine the search tool in a way that respects, on the long run, the very diverse expectations of teachers in Europe. It might also uncover reasons to adopt (accept to re-use) and to adapt (modify to make more fit to the purpose) which are not yet sufficiently explored transform the platforms into comprehensive exchange marketplaces (see [4] for an early study in this direction). Such a test base might also enrich the vision of cross-lingual search engines, whose pioneering works, as reported in [5] seem not yet appropriate to the world of learning resources.

B. Suggestions

Following the quick search and refine process typical of information retrieval, e.g. [6], a mechanism to suggest *related queries* should be studied and tested, similarly to [7].

C. More detailed user information

While technical competencies are followed through, the personal interests of teachers are not yet collected in the Open Discovery Space platform. Such collection may bring richer preference queries while it runs the risk to add more forgotten context information. A mechanism to make the search process more transparent, allowing the user to understand that a given search result has not respected all of his or her criteria for example, seems not yet explored.

V. Acknowledgment

We thank Ard Lazonder, Thanasis Hadzilacos, Anna Mavroudi, Katarina Riviou, and Noortje Janssen for a productive early design elaboration, which this document sketches. We also thank Panagiotis Zervas, Lamprini Kolovou, Luis Lalueza, Salvador Sanchez, George Thanos, and Maricruz Vallente for their technical collaboration in preparing the Open Discovery Space search engine.

References

[1] Hewlett Foundation, *Open Educational Resources*, available at http://www.hewlett.org/programs/education/open-educational-resources, accessed in May 2015.

[2] Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze:,*Introduction to information retrieval*. Cambridge University Press 2008, ISBN 978-0-521-86571-5.

[3] Kavi Mahesh. *Text retrieval quality: A primer*. Technical report, Oracle Coproration, 2006. See http://www.oracle.com/ technology/products/text/htdocs/imt_quality.htm.

[4] Paul Libbrecht, *Adaptations to a Learning Resource*. To appear in Acta Didactica Napocensia.

[5] Carol Peters, Martin Braschler, Paul D. Clough:,*Multilingual Information Retrieval - From Research To Practice*. Springer 2012, ISBN 978-3-642-23007-3, pp. I-XVII, 1-217.

[6] Cornelis Joost van Rijsbergen. *Information Retrieval*. Butterworths, 1979. Available at http://www.dcs.gla.ac.uk/~iain/keith/.

[7] Paul Libbrecht, Escaping the Trap of too Precise Topic Queries, *Intelligent Computer Mathematics*, Proceedings of CICM 2013, LNCS 7961, Jacques Carette, David Aspinall, Christoph Lange, Petr Sojka, Wolfgang Windsteiger (Eds.), ISBN: 978-3-642-39319-8.

The ODS Vocabulary Bank

Anastasios Koutoumanos and Elina Megalou
Computer Technology Institute & Press "Diophantus", Greece

***Abstract*—The ODS Vocabulary Bank is a component of the Open Discovery Space socially-powered federated architecture for repositories of learning resources. It implements a single point of reference for all authoritative sets of terms, concepts and named entities as well as the network of relations between them, using an innovative, semantic web approach. This paper provides an overview of the ODS Vocabulary Bank, the software requirements, the research of the state of the art, the selected implementation approach and some considerations regarding future work towards the evolution of the system.**

***Index Terms*— ODS, semantic web, taxonomy, vocabulary**

I. Introduction

The ODS Vocabulary Bank is a web-based application that implements a single point of reference for authoritative sets of terms, concepts and named entities used for the metadata descriptions of learning resources, as well as the network of semantic relations between them. It is an essential component of the system's architecture of the Open Discovery Space (ODS) initiative, which implements a community-oriented social platform enabling teachers, pupils, and parents to discover, acquire, exchange, and adapt learning resources from a network of interconnected repositories across Europe. The existence of a single point of reference has been identified as crucial from the early phases of the design of the ODS architecture, to accomplish a unified experience for all users interacting with the ever-increasing wealth of resources, allowing the use of a common language for describing, locating, and browsing for these resources.

The paper is structured in six sections: following this introduction, section 2 provides a short overview of the ODS initiative; section 3 describes the essential requirements of the ODS Vocabulary Bank, based on which an analysis of the state of the art for similar systems is presented in section 4; the implementation of the system is documented in section 5; and section 6 concludes the paper with a discussion of the main outcomes from system's usage and planned future work.

II. The ODS Federated Architecture

The ODS infrastructure and its overall system architecture is described in the following paragraphs, to make apparent the importance and need for a vocabulary bank as a central reference point, enabling the management and usage of common terminology across the different components and facilitating multilingualism for all related services.

The Open Discovery Space is a prominent European initiative that aims in (a) implementing a socially-powered open federated infrastructure for a super-repository on top of existing repositories of learning resources and related federated infrastructures; and (b) providing social features for building and sustaining web-based educational communities and communities of best teaching practices from 2,000 European schools [1]. One of the main objectives of the ODS infrastructure is, thus, to make available a wealth of learning resources from a large network of repositories across Europe, through a community-oriented social platform (web portal) where teachers, pupils, and parents discover, acquire, discuss, and adapt learning resources on their topics of interest. An essential challenge for ODS is, therefore, to achieve interoperability and information sharing between semi-autonomous de-centrally organized repositories. In order to facilitate discovery of resources, all their descriptions need to be aligned to a uniform metadata schema, which includes a set of predefined terms, concepts, and thematic taxonomies for each specific educational discipline. Users can therefore search against selected terms in specific fields or browse for resources in their preferred disciplines.

In order to address this challenge, the metadata of all interconnected repositories are collected, processed, and stored by the central component of the ODS infrastructure, the ODS harvester. This component also executes all necessary transformations to map information from each metadata schema to the native metadata schema. Thus, all information becomes readily accessible through the ODS web portal, which provides the front-end user interface and services. The portal serves an international, multilingual audience of teachers, learners, and parents and facilitates searching, browsing, and accessing the wealth of learning resources across all connected repositories, removing linguistic and cultural barriers. Furthermore, a set of peripheral components allows the interconnection of the ODS infrastructure to external systems, using established protocols and standards.

The ODS Vocabulary Bank provides a specific, yet essential "glue" service to the infrastructure: it implements a central point of reference for all associated vocabulary terms and taxonomies, their descriptions, and their translations. The next section provides an enumeration of the key software requirements and the associated expected benefits from the incorporation of the ODS Vocabulary Bank in the overall infrastructure.

Manuscript received April 30, 2015. This work was supported in part by the European Commission, as part of the "Open Discovery Space: A socially-powered and multilingual open learning infrastructure to boost the adoption of eLearning resources" project, CIP PSP Grant Agr. no. 297229.

Anastasios Koutoumanos is a research associate of the "Strategy and Digital Educational Content" Directorate of the Computer Technology Institute & Press "Diophantus", 26-28 Mitropoleos str., GR-10563, Athens, Greece (phone: +30-210-3350600; e-mail: koutoumanos@cti.gr).

Elina Megalou is the deputy director of the "Strategy and Digital Educational Content" Directorate of the Computer Technology Institute & Press "Diophantus" (phone: +30-210-3350600; e-mail: megalou@cti.gr).

III. FUNCTIONAL REQUIREMENTS

This section presents the key functional requirements for the ODS Vocabulary Bank, to enable the collaborative editing and management of terms using a linked open data approach, while taking into consideration the multilingualism requisites and, at the same time, enabling its integration with other parts of the ODS infrastructure as well as external systems.

While it could be possible to manually manage the set of terms and taxonomies that are used across the ODS infrastructure and the network of interconnected repositories of learning resources, this would be a very demanding task. Curators from each repository would be required to invest many human resources, engage in extensive communication and error-prone processes to exchange information, align terms, provide translations, updates, and so on. It has thus deemed important that a software system is built in order to facilitate this task – this system is the ODS Vocabulary Bank.

The main goal of the ODS Vocabulary Bank is therefore to implement an open vocabulary server, which allows collaborative editing, management, evolution, and publishing of vocabularies, thesauri, and taxonomies. The key requirements for the development of the ODS Vocabulary Bank have been identified as following:

1) In order to be able to harvest resources from all repositories that are harvested within the ODS federated infrastructure, curators from each repository need to provide a mapping of the vocabulary terms and thematic taxonomy used in their repository to the terms and taxonomy of the ODS metadata schema.

2) The ODS central repository needs to expose the metadata terms and taxonomies as well-defined ontologies, using a linked open data approach, allowing all learning resources to be linked and described against them.

3) A set of different formats is used for the exchange of terminologies (terms and taxonomies) from related repositories: (a) Spreadsheet: a simple spreadsheet file using different columns to categorize terms' classifications; (b) OWL: a file following the Web Ontology Language specification; (c) SKOS: the W3C recommendation for representation of thesauri, taxonomies, or any other type of structured controlled vocabulary (Simple Knowledge Organization System). It has been identified as essential for the ODS Vocabulary Bank to be able to import terminologies from files that follow each of these three formats.

4) The ODS Vocabulary Bank needs to address the multilingual requirements of ODS, by facilitating translations of all managed terms and taxonomies. To this end, all necessary relations between relevant terms in different languages need to be easily identified, managed and exposed to the other components of the ODS infrastructure. Moreover, the system itself ought to provide a multilingual user interface, to allow curators from repositories across Europe to use it in their preferred language.

5) As a single point of reference, the ODS Vocabulary Bank needs to provide access to all managed terms and their relationships through a SPARQL endpoint, following the Linked Open Data approach and enabling integration with other applications and services of the semantic web.

6) The system should feature a user-friendly, multilingual user interface that facilitates an intuitive overview and search through the managed vocabularies and a specific web page with a well-defined URL holding the detailed description of each term and all its relations to other terms.

IV. RELATED WORK

After the initial requirements specification, the development team engaged in a research of the state of the art in terms of technology, standards, and protocols, as well as the associated choices for the implementation of the vocabulary server. The results of this activity are briefly described in the following paragraphs.

The *Vocabulary Bank for Education* (VBE) [2] is a terminology management system for publishing vocabularies, concept schemes, data definitions, taxonomies, and thesauri. The VBE is designed to support a range of exchange formats and exposes all historical information about the terminology. The VBE is an outcome of the ASPECT (Adopting Standards and Specifications for Educational Content) EU-funded project and is further developed by Knowledge Integration with the brand name "Lexaurus Bank" [3].

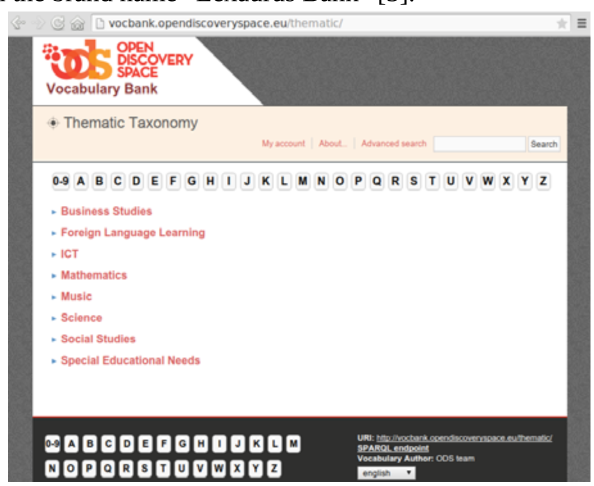

Fig. 1. The main page of a taxonomy managed by the ODS Vocabulary Bank

VocBench [4] is a web-based, multilingual, editing tool that manages thesauri, authority lists, and glossaries using SKOS. It is developed by the Food Agriculture Organization (FAO) of the United Nations (UN) and its partners. VocBench supports a growing set of communities, including the distributed group of terminologists who manage AGROVOC, a multi-lingual thesaurus of 40,000 terms in over 20 languages. Developed in Java, VocBench's latest production release has undergone a major rewriting, rewiring the RDF backend on Semantic Turkey, an RDF framework for Knowledge Acquisition and Management developed by the ART team of the University of Rome Tor Vergata.

iQvoc [5] is a vocabulary management tool that combines easy-to-use graphical user interfaces with semantic web interoperability. It supports a number of features including: import of existing vocabularies from a SKOS representation;

multilingual browser-based display and navigation; editorial features for registered users; publishing in the semantic web. iQvoc is implemented with the Ruby on Rails web framework and is licensed under the Apache License, Version 2.0.

TemaTres [6] is an open source vocabulary server to manage, maintain, and exploit formal representations of knowledge. It is able to manage relations between terms across different vocabularies. It supports the lifecycle of vocabulary management and terms can be marked as candidate, accepted or rejected. TemaTres enables export of the vocabularies and classifications in RDF format (SKOS-Core) and XML format (Zthes, TopicMaps, Dublin Core, and more). TemaTres is built with the PHP web server scripting language, uses the MySQL database engine and is licensed under the GNU GPL.

MoKi [7] is a collaborative tool for modeling ontological and procedural knowledge in an integrated manner, based on the MediaWiki open source wiki package that was originally written for Wikipedia. The main idea behind MoKi is to associate a wiki page, containing both unstructured and structured information, to each entity of an ontology and process model. MoKi's units have been implemented upon a LAMP (Linux-Apache-MySQL-PHP) platform, with the addition of the Java language that has been used for implementing a part of the Ontology Import unit.

The *PoolParty Thesaurus Server* [8] is a tool for the creation of taxonomies, thesauri, and knowledge graphs on the market. It combines methods based on linked data, text mining, and knowledge engineering. It is available in three versions, all of which are offered under a commercial license with an increasing, four-figured price tag.

After our state of the art research and taking into consideration the requirement for open source software as the basis for the implementation of the ODS Vocabulary Bank, it was clear that we need to rule out the options of the VBE, as well as the PoolParty Thesaurus Server. MoKi has been successfully used in the EU-funded "Organic.Edunet" and "Organic.Lingua" projects [9], but the source code was not distributed at the time of our research and it was not directly reusable for implementing the planned functionality. The same held true for the Java-based VocBench. TemaTres and iQvoc both offer similar functionality and their source code is freely available. Our final choice has been TemaTres, based on the following criteria: (i) it uses the PHP scripting language, which was more familiar to our development team than the Ruby language; (ii) it has been successfully deployed in many projects; (iii) it offers responsive feedback from its developer community; and (iv) it features a well-documented source code.

V. Implementation Approach

The ODS Vocabulary Bank is based on the TemaTres open source software and implements a web-based system that serves all tiers and workflows for the management of the terms and taxonomies of the metadata schemes related to the ODS infrastructure. Meeting the requirements set in section 3, it facilitates authorized users for editing and organizing the terms, their descriptions, translations and relations. Additionally, it assists in linking terms between associated vocabularies and supports Semantic Web technologies for exporting the ODS terms in Linked Data formats. Through the ODS Vocabulary Bank, all terms of relevant vocabularies are publicly available to every interested party, readily accessible with any modern web browser, as well as via well-documented web services. Furthermore, the ODS Vocabulary Bank enables a complete export of the vocabularies and taxonomies using well-established specifications and file formats (SKOS-Core, Zthes, TopicMaps, Dublin Core, etc.).

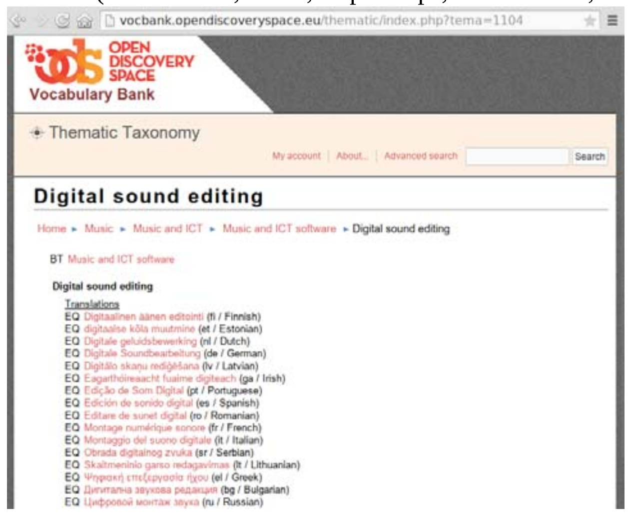

Fig. 2. The page of the term "Digital sound editing"

TemaTres has been adapted and extended, including:

- new functionality for the management terms' translations
- enhancement of the user interface for a more user-friendly look & feel and a responsive, mobile-friendly experience;
- support of multilingualism at the front-end to allow usage from an international audience of interested users;
- support export of vocabularies' terms using the JavaScript Object Notation for Linked Data (JSON LD);
- enhancement of the SKOS-Core export;
- enhanced quick navigation by letter and language;
- web services for incremental harvesting of terms, including support for reporting of last updated terms;
- web services for getting translations of any term.

For every vocabulary that needs to be managed, and for each language that needs to be supported, a new instance of the software is being setup. All instances share a common source code and differ only in the configuration of the database, the set of users, and related access rights, and, possibly, the layout of the user interface. In order to indicate translations of terms, authorized users can associate definitions of terms, across different instances of the Vocabulary Bank, using an "equality" relation. It is not an overstatement to describe the overall system as a "federation of vocabularies to support the ODS federated infrastructure".

The front page of the ODS Vocabulary Bank, available at http://vocbank.opendiscoveryspace.eu, presents the index of

```
<rdf:RDF
 xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
 xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
 xmlns:skos="http://www.w3.org/2004/02/skos/core#"
 xmlns:map="http://www.w3c.rl.ac.uk/2003/11/21-skos-
mapping#"
 xmlns:dct="http://purl.org/dc/terms/"
 xmlns:dc="http://purl.org/dc/elements/1.1/">
<skos:ConceptScheme rdf:about="http://vb.ods.eu/thema/">
 <dc:title>Thematic Taxonomy</dc:title>
 <dc:creator>ODS team</dc:creator>
 <dc:description><![CDATA[  ]]></dc:description>
 <dc:date>2013-11-28</dc:date>
 <dct:modified>2014-05-22 18:47:41</dct:modified>
 <dc:language>en</dc:language>
</skos:ConceptScheme>
<skos:Concept rdf:about="http://vb.ods.eu/thematic/1104">
 <skos:prefLabel xml:lang="en">
   Digital sound editing</skos:prefLabel>
 <skos:prefLabel xml:lang="es">
   Edición de sonido digital</skos:prefLabel>
...
 <skos:inScheme rdf:resource="http://vb.ods.eu/thematic/"/>
 <skos:broader
rdf:resource="http://vb.ods.eu/thematic/1103"/>
 <dct:created>2013-11-28 16:32:34</dct:created>
</skos:Concept>
</rdf:RDF>
```


Fig. 3. SKOS-Core export of a selected term from the ODS Vocabulary Bank

all the taxonomies and vocabularies hosted in the Vocabulary Bank. The “Thematic Taxonomy” item holds the set of terms that comprise the thematic, tree-structured taxonomy for describing learning resources (Fig. 1). Other items, include the “Vocabularies”, which holds all the sets of controlled vocabularies with the terms that are used for specific elements of the metadata description of learning resources and the “National Taxonomies”, which includes all taxonomies from each nation holding the terms for the disciplines of “History”, “Religion”, and “Language & Literature”. The “Repositories' Vocabularies” item contains vocabularies and taxonomies as they are specific for certain repositories aggregated by ODS.

Selected metadata experts from each associated repository have access to the management back-end, where they can use the provided functionality to manage the terms of their metadata application profile. The description of each term is available through a user-friendly web page (Fig. 2), or can be directly exported in one of the supported formats so that it can be handled and digested by other software systems (Fig. 3). In particular, the ODS Vocabulary Bank offers many alternatives to be integrated with other systems. Each taxonomy and/or vocabulary is accessible through web services over HTTP, it can be made available through batch export in a single file, and a SPARQL endpoint allows for querying the vocabulary bank and requesting for the representation of each term in machine readable format. Regarding the web services, these are listed in the web page “/services.php”, which holds the index of the “verbs” of the Application Programming Interface (API), along with their description, the expected input, and the output. They are pretty well self-descriptive and the output can be either in XML or JSON format. In terms of the batch export, the main and fully-supported format is SKOS-Core RDF/XML. This way a single file can hold all the information regarding a specific taxonomy and/or vocabulary. The SPARQL endpoint, provided at “/sparql.php”, can respond to complex queries structured in SPARQL syntax and provide specific results. Finally, each term is provided in two machine readable format, SKOS-Core RDF/XML, and JSON-LD.

VI. Conclusions And Future Work

The usage of the ODS Vocabulary Bank has served well its objective to facilitate the management and maintenance of a large set of vocabularies and taxonomies associated with the metadata schemas of the repositories participating in the ODS federated infrastructure. Moreover it as allowed for additional added-value services to be developed. For example, a visual terminology browser can provide an intuitive overview of the hierarchical structure of a thematic taxonomy, so that users can directly navigate to their topic of interest and browse for resources in each selected topic. Or users can search for resources using keywords and terms in their preferred language, whereas the system can use the information of the ODS Vocabulary Bank to search for resources with equivalent terms in their description and, thus, create a much richer result-set for the original query. A powerful autocomplete service has allowed to step-up the effectiveness of user searches. Last, but not least, other systems can readily use the offered services through the well documented API in order to create new services or enhance existing ones.

In terms of future work, the development team at CTI Diophantus is planning to enhance the system in the following areas: better support for versioning of vocabularies; integration of a service for providing globally addressable persistent unique identifiers; and support for more human-friendly “clean” URLs. Moreover, the source code of the existing and further modifications to the TemaTres source code shall be contributed back to the Free Open Source Software (FOSS) community through a publicly available source code repository, at http://git.dschool.edu.gr.

References

[1] N. Athanasiadis, S. Sotiriou, P. Zervas, and D. G. Sampson, “The Open Discovery Space Portal: A Socially-Powered and Open Federated Infrastructure,” in *Digital Systems for Open Access to Formal and Informal Learning*, Springer, 2014, pp. 11–23.

[2] R. Tice, “Aspect Vocabulary Bank for Education,” in *Digital Libraries: interoperability, Best Practices and Modeling Foundations, 2009*.

[3] "Lexaurus Bank." Knowledge Integration Ltd, 2013. *Accessed online (Feb 2015)* http://www.k-int.com/products/lexaurusbank

[4] A. Stellato, S. Rajbhandari, A. Turbati, C. Caracciolo, M. Fiorelli, J. Keizer, and M. Pazienza, “VocBench 2.0: a Collaborative Web Application for the Development of Thesauri and Concept Schemes,” in *LOD2014*.

[5] T. Bandholtz, J. Fock, A. Wolff, and H. Schentz, *“LOD-ready Environmental Terminology with iQvoc,”* 2011.

[6] A. Gonzales-Aguilar, M. Ramirez-Posada, and D. Ferreyra, “TemaTres: software to implement thesauri,” *Professional de la Infromaction*, vol. 21, no. 3, pp. 319–325, 2012.

[7] M. Dragoni, C. Ghidini, and A. Bosca, “Multilingual MoKi: How to Manage Multilingual Ontologies in a Wiki,” in *The Semantic Web: ESWC 2013 Satellite Events*, vol. 7955, P. Cimiano, M. Fernández, V. Lopez, S. Schlobach, and J. Völker, Eds. Springer, 2013, pp. 162–166.

[8] M. Bastiano, . Blankman, M. Frenzel, U. Grandin, M. Kertesz, and A. Oggioni, *“EnvThes-interlinked thesaurus for long term ecological research, monitoring, and experiments,”* 2013.

[9] A. Bosca, M. Dragoni, C. Di Francescomarino, and C. Ghidini, “Collaborative management of multilingual ontologies,” in *Towards the Multilingual Semantic Web*, Springer, 2014, pp. 175–192

e-Learning Object Ingestion in an Open Educational Environment

Enayat Rajabi, Kostas Vogias, Salvador Sanchez-Alonso, Ilias Hatzakis

***Abstract*— A large number of learning objects' metadata are available on the Web. Published by different sources, these metadata cannot be represented to the end users in their original formats, as they require some technical steps to be filtered, checked and cleaned due to several issues e.g. broken links. In this paper, we present an aggregation workflow followed in an open educational environment (the Open Discovery Space project) in which a large amount of metadata passed through several technical steps as a pre-filter to be integrated into this educational Web portal.**

***Index Terms*— Repository, Aggregation, Metadata, Open Discovery Space.**

I. INTRODUCTION

In the past decades, educational objects have been published by various suppliers of educational services on the Web [1]. These resources targeted to certain categories of learners, which can be students, teachers, employees, etc. On the other hand, researchers and repository owners in the educational domain have developed various e-learning systems to aggregate, publish and consume plenty of these e-learning resources, so that they can be discovered, navigated and reused by different kinds of applications and users on the Web [2]. However, low quality metadata can render a library or repository almost unusable, while ingesting metadata with high quality can lead to higher user satisfaction [3]. To this aim, data publishers utilize different approaches for filtering, enriching, and checking metadata. In this paper, we describe an aggregation approach in which a large number of metadata were harvested from several eLearning repositories, validated to be aligned to the defined metadata schema, and transferred to the proper application profile. In this experimentation, we filtered out the useless and low quality metadata as well. The final resources were imported to an eLearning repository as well.

The reminder of this paper is organized as follows. Section 2 describes the metadata aggregation in open learning environment and introduce the Open Discovery Space project as the case study we selected for this research. The methodology of the present study is outlined in Section 3. Section 4 presents the challenges we faced in the aggregations of metadata within the project. The final remarks are presented in section 5.

II. BACKGROUND

The exponential growth in the amount of digital learning objects is forcing architects, engineers and developers involved in creating digital repositories to face the harsh reality that their solutions need to handle an amount of e-objects that is orders of magnitude larger than originally intended. Optimizing, tuning, and tweaking the existing repository infrastructure can initially alleviate performance problems, but eventually limits are reached. At that point, a major redesign of the repository solution is an obvious option. An alternative is to move towards an environment that consists of parallel instances of the existing repository solution and to glue those together into a repository federation that behaves as if it were a single repository [4]. The desire to federate repositories in such a way typically emerges as a result of the understanding that no single repository hosts all e-learning objects that are relevant for a specific subject domain. Generally speaking, federation is a decentralized approach that emphasizes partial, controlled sharing among repositories and provides a means to share data and transactions using some protocols such as OAI-PMH [5] and providing the coordination of data exchange among them. A federation supports interoperability among the registered repositories and reduces dependency on any single metadata collection. Figure 1 illustrates in a simple way a federation of repositories.

The work presented in this paper has been part-funded by the European Commission under the ICT Policy Support Programme CIP-ICT-PSP.2011.2.4-e-learning with project No. 297229 "Open Discovery Space (ODS)".

Enayat Rajabi is a PhD researcher at the Computer Science Department of University of Alcalá, Spain. email: enayat.rajabi@uah.es

Kostas Vogias is a Software Engineer on metadata aggregation, processing and exposure at GRnet. email: gvog84@gmail.com

Salvador Sanchez-Alonso is an associate professor at the Computer Science Department of University of Alcalá, Spain. email: salvador.sanchez@uah.es

Ilias Hatzakis is a metadata aggregation related project manager, at GRnet. email: hatzakis@grnet.gr

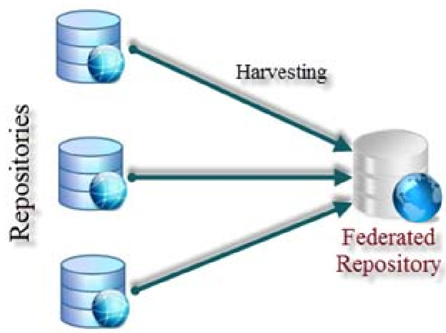

Fig. 1. Harvesting a federation of repositories

In a federated repository metadata information is collected from many contributors or repositories to create, on top of all of them, a search service supporting simultaneous discovery of information resources residing in the collections of all the repositories. The way in which the metadata are collected is usually referred to as "harvesting", a computer software technique of extracting metadata information from external data sources by periodically accessing them, using a standard protocol agreed by the 2 parties (client and server). Collecting metadata through such standard protocol (e.g., OAI-PMH) has been utilized by a wide variety of projects including Open Discovery Project (ODS) [6].

The ODS project aims to support open access to digital educational resources and practices from members of school communities (that is teachers, students and parents) in Europe. This project, which is the result of collaboration between 51 partners from 23 European countries, exploits the elements of eLearning resources (i.e., educational objectives, pedagogical models, learners' personal characteristics and needs, etc) collected from many educational repositories and federations across the Europe. This ongoing project has promoted community building between numerous schools of Europe (At the time of this research 2,833 schools with around 7,600 teachers participated) and empowered them to use, share and exploit unique resources from a wealth of educational repositories. ODS uses an Open Linked Learning Content Infrastructure and has recently exposed around 800,000 eLearning metadata (from 25 eLearning repositories) through a Web portal.

In ODS, most of the harvested repositories conform to an Application Profile, which is based on an IEEE LOM schema and created within the project, and provided their eLearning metadata according to it. However, some repositories have their own custom schema and a mapping between their schema and ODS Application profile was carried out after harvesting by means of producing a XSLT file (specific of each repository) and running a transformation on the harvesting infrastructure side by using transformation/alignment tools. The production of these XSLT files was the responsibility of each repository assisted by some technical experts in alignment and integration mentioned in the description of work (Figure 2).

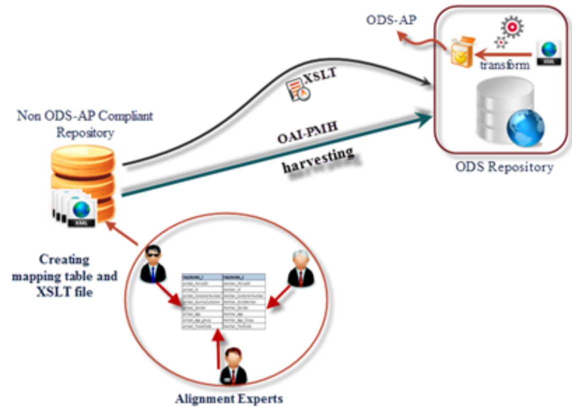

Fig. 2. harvesting process

III. METHODOLGY

The harvested metadata in ODS have passed the following steps to be imported in the ODS portal (consider Figure 3):

A. Repository Catalog

The Repository Catalog is a CRUD software developed for maintaining repositories' related information. This information is used by the aggregation workflow software in order to configure the various constituent steps.

B. Harvesting

This is the first step in the metadata aggregation workflow. The harvesting protocol used is OAI-PMH.

C. Transforming

In this step, each repository metadata were transformed to the ODS AP using the XSLT file they sent. Each repository generated its own XSLT file and sent it to the ODS repository owner to perform the transformation step. As a result of this step, all the repositories metadata were transformed to the ODS schema as well.

D. Identification

In this step, each learning object and its metadata were identified by a global identification approach.

E. Validation

In the validation step, each repository data were validated against the ODS AP and thus the invalid records were filtered out. The main reasons for a record to be considered as invalid are the following:

- Mandatory elements absence
- Incorrect vocabulary mapping

F. Filtering

Those metadata that did not include any text in the mandatory elements (e.g learning object's title and location) were filtered out.

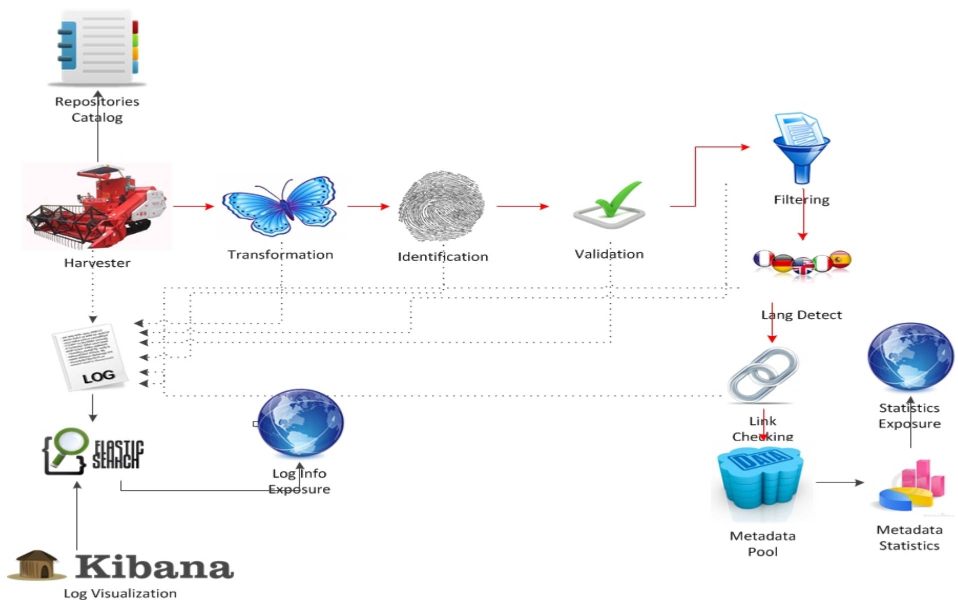

Fig. 3. Aggregation workflow

G. Language enrichment

In this step, if there are some metadata elements that it is important from ODS Portal perspective to contain a language attribute(title, description, keywords etc), then they are enriched with it using a language detection software [7].

H. Link checking

Figure 4 represents the link checking process in ODS, which is part of the metadata aggregation system. Overall, the ODS harvester (1) collects repositories' metadata based on ARIADNE-powered infrastructure. The harvested metadata are stored in a file system, separated by their harvested source in different folders. The link checker engine checks (2) each learning resource individually by testing the URL contained in the respective resource metadata, as mentioned in the previous paragraphs. Information regarding to broken links (e.g. File name, file path, status, timestamp) is stored in a log file (3). Non-broken links, what we call usable resources, are moved in different folders (4). Resources that contain broken links are checked (7) periodically (definite periodicity is still under discussion) and they will be moved to live folder (8) for joining to usable resources, if the link is live again. The recovered resources will be re-indexed in the next harvesting cycle (5). The link checking results are also exposed with a variety of REST APIs. The whole idea is these APIs to be consumed by ODS Portal in order this to be informed about the link checking results and perform the necessary updates to its data base.

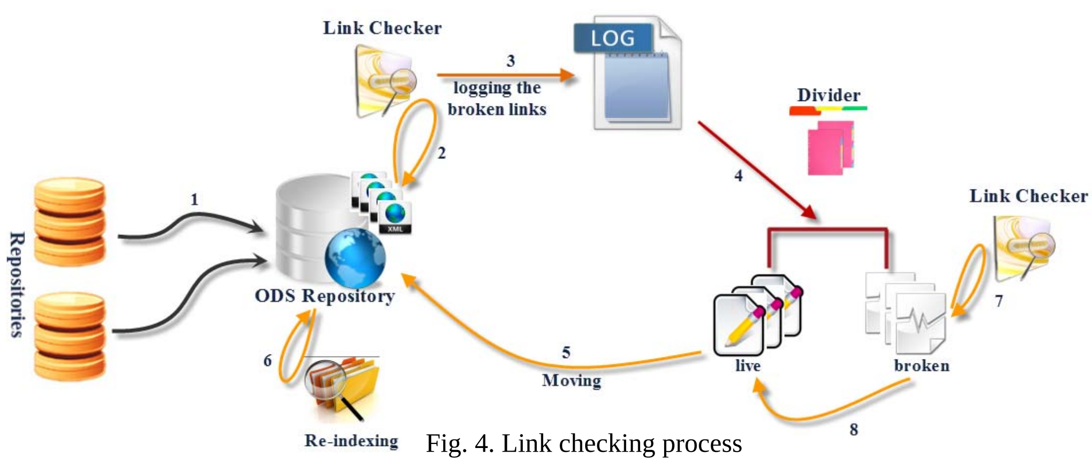

Fig. 4. Link checking process

I. Metadata Pool

The metadata pool is a file system folder where all the usable metadata records are placed.(The name is an abstraction of the file system folder that contains the final usable metadata.)

J. Metadata Statistical Analysis and visualizations

The aggregation workflow is accompanied with a metadata analysis tool. This tool was developed using JAVA and its purpose is to analyze XML documents and export the results to a human readable format (CSV) [8]. s a first step, it performs a per repository analysis and as a second step it performs the analysis of the repositories at the aggregation level. The statistical measures that are being calculated are: element frequency, element completeness, element dimensionality and element content relative entropy [9]. It is also possible for the user to additionally choose a specific element for vocabulary usage analysis (frequency). Finally, an attribute based value analysis (attribute value frequency) is also implemented that can be used to study the multilinguality of the free text metadata elements. This tool proved to be very helpful for aggregation specialists and data providers to gain a deeper knowledge of the quality and content anomalies of the aggregated metadata.

K. Aggregation Results Visualization

A need that came up after the first harvesting cycle was the metadata aggregation results and statistics visualization. In this way the aggregation analysts, data providers and project managers became able to supervise the whole aggregation process.

L. Ingesting in the portal

Using an updater the finalized metadata are read and inserted into a relational database which is later used for the portal.

Fig. 5. Aggregation workflow results visualization

IV. CHALLENGES

The challenges we faced in the aggregation steps are described below:

Harvesting: At this point the most usual problems occur on data providers' side and are mostly OAI-PMH protocol implementation issues:

- Bad timestamp implementation. This disables the harvester's ability to recognize the new and also the updated records thus incremental harvesting approach cannot work.
- Bad resumption token implementation. This makes the harvester unable to harvest all the metadata records exposed by data providers. The harvester freezes to a certain record and can't harvest the rest metadata.
- No deleted records policy: The absence of a deleted records policy makes the harvester and the whole aggregation workflow unable to recognize what records should be deleted from the respective repository.
- Identification: If a metadata record doesn't use any element that describes the learning objects' existence (in LOM the technical.location and the general.identifier elements) then the specific metadata describes nothing, therefore the respective metadata record is discarded.
- Language Detection: The language detection mechanism could be considered as second filtering step since it filters out all metadata records that contain elements the content of which can't be language detected. The most usual case here is a metadata element that contains various symbols as text (this record passed successfully validation and filtering steps but should not be presented to ODS Portal though).
- Link Checking: A very common issue at this step is the case when a Learning Object's link although it is live, accessible and well formed however it points to a login page or a web representation of the metadata that describes it instead of the actual Learning Object.

V. CONCLUSION

In this paper, we described a workflow in which a large number of eLearning metadata, were collected from several repositories, processed, cleaned, evaluated and finally imported into a learning portal. After defining a schema for structuring the metadata, a large number of metadata were harvested by ARIADNE harvester, and then we separated the healthy metadata by checking their schema and contents. Particularly, we filtered out the metadata with broken links and empty titles, and finally we imported the cleaned data into the portal. The aggregation workflow described above could be considered as generic enough, covering most of the learning object's metadata processing needs. Thus with very small adaptations (mostly at metadata schema level) the specific workflow could be used as a ruler for future learning object metadata repositories.

REFERENCES

[1] M. Fernandez, M. d' Aquin, and E. Motta, "Linking data across universities: an integrated video lectures dataset," in *Proceedings of the 10th international conference on The semantic web - Volume Part II*, Berlin, Heidelberg, 2011, pp. 49–64.

[2] M.-C. Valiente, M.-A. Sicilia, E. Garcia-Barriocanal, and E. Rajabi, "Adopting the metadata approach to improve the search and analysis of educational resources for online learning," *Comput. Hum. Behav.*

[3] B. Stvilia, L. Gasser, M. B. Twidale, S. L. Shreeves, and T. W. Cole, "Metadata Quality for Federated Collections," Nov. 2004.

[4] H. Van de Sompel, R. Chute, and P. Hochstenbach, "The aDORe Federation Architecture," *ArXiv08034511 Cs*, Mar. 2008.

[5] "Open Archives Initiative: Protocol for Metadata Harvesting." [Online]. Available: https://www.openarchives.org/pmh/. [Accessed: 29-Apr-2015].

[6] "Open Discovery Space Project." [Online]. Available: http://www.opendiscoveryspace.eu/. [Accessed: 28-Apr-2015].

[7] "Language Detection Library for Java." [Online]. Available: https://code.google.com/p/language-detection/. [Accessed: 05-Jan-2015].

[8] K. Vogias, I. Hatzakis, N. Manouselis, and P. Szegedi, "Extraction and Visualization of Metadata Analytics for Multimedia Learning Repositories: the case of Terena TF-media network," in *Proceedings of the LACRO 2013 Workshop*, 2013.

[9] "GLOBE Metadata Analysis." [Online]. Available: https://sites.google.com/site/globemetadata/home.

Searching and exposing learning objects from Moodle: The ODS experience

Siniša Tomić, Vlatka Paunović, and Kristijan Zimmer
Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia
sinisa.tomic@fer.hr; vlatka.paunovic@fer.hr; kristijan.zimmer@fer.hr

***Abstract*—Learning Management Systems (LMS's) play an important role in technology-enhanced learning. Open source LMS Moodle is currently one of the most popular, due to its rich set of features, flexibility and compliance with IT and pedagogical standards and user needs. In order to adopt the exchange of learning materials with as little intervention from the course creators, two new Moodle plugins were developed. Their purpose is to enable teachers to search but also configure what is exposed, how and when, by providing mechanisms for automatic generation of validated LOM-based metadata for harvesting by the OAI-PMH harvesters. This research paper describes the method for searching and exposing learning objects from Moodle. Factors which will be taken into consideration include the ease of use from course creator's and Moodle administrator's point of view and how well the generated metadata match user requests as well as impact of the harvesting process to the Moodle installation.**

***Index Terms*—E-learning, LMS, LOR, OER, OAI-PMH, automatic metadata generation, matching problem**

I. Introduction

With the growing usage of Learning Management Systems (LMS's) they are becoming an essential part of learning. More and more teachers are accepting this new way of disseminating knowledge and even more students are learning how to use this kind of technology. In Croatia these platforms were at first recognized at the university level as a way of providing additional knowledge to students. However, in the last few years they are becoming an important part of learning process in both primary and secondary schools. E-Learning has overgrown its initial usage of just an additional or specific knowledge provider and it is becoming an integral part of school curricula [1].

Today teachers are not only using already made materials but they are creating a lot of their own content as part of their teaching. This led to the creation of similar e-Learning materials by the teachers who use the same curricula. The same problem occurred in programming in the late 1960s when lots of programmers began writing similar parts of code for different projects [2]. At that time a new programming paradigm was defined. It was called object-oriented programming. The same idea is now becoming a part of e-Learning. This new approach to e-Learning enables teachers to reuse parts of already existing materials (learning objects) in their newly created courses [3, 4].

A learning object can be any digital chunk of information. It can be text, audio, video, model, but it can also contain additional information such as simulations and assessment. It can be as small as one definition, but also as large as a complete lesson module. These chunks of information can be stored centrally or distributed between many different systems. The system that holds that kind of information is called Learning Object Repository (LOR).

There are many LORs available today, some of which are used more in comparison to others. One method of choosing the right LOR for a teacher is the number of indexed learning objects. Repository that indexes more objects can usually offer better search results and is therefore regarded as better. An example of such a large repository is Open Discovery Space, which currently indexes more than 800.000 individual learning objects [5].

Once found, learning objects can become a part of a new course that is created in LMS. Today, Moodle [6] is widely accepted as one of the top LMS, mostly because it, among other things, enables its users to easily create content by reusing the existing one and additionally customize it with built-in tools [7].

The idea behind the work described in this paper is to give Moodle users two additional enhancements: (1) Search the already available resources on the Open Discovery Space and (2) Share their own learning objects created or customized in Moodle with other users using OAI-PMH open standard [8].

For those reasons, two separate Moodle block plugins were designed. Both of them are designed in a way that even teachers who are not technically adept can use them. This approach should result in higher usage of developed plugins since more teachers will be able to use them.

The reminder of this paper is organized as follows. Section 2 describes the ODS Learning Object Repository with the possibilities it offers to teachers as well as the reasons for creating this Moodle integration. Section 3 presents the methodology of creating and using the proposed solution. The

Siniša Tomić is with the Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia (e-mail: sinisa.tomic@fer.hr)

Vlatka Paunović is with the Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia (e-mail: vlatka.paunovic@fer.hr)

Kristijan Zimmer is with the Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia (e-mail: kristijan.zimmer@fer.hr)

final remarks are presented in section 4.

II. Background

Open Discovery Space is a socially-powered and multilingual open learning infrastructure to boost the adoption of e-Learning resources. One of the parts of the Open Discovery Space infrastructure is Learning Object Repository. Access to this repository is fully open for all teachers and students. The Open Discovery Space engages teachers, students and parents in a first of a kind effort to create a pan-European e-Learning environment to promote more flexible and creative ways of learning. It is also the Europe’s largest collection of e-Learning resources, scenarios and activities [5].

The main problem is that many teachers who use Moodle are not aware of this repository. In some cases, even when acquainted with it, they are sometimes not completely sure how to use the data in their Moodle course.

Currently, if a teacher uses Moodle and wants to search LOR for some learning objects to use in a course, they should first recall the information they are looking for, then open a new browser window, enter the correspondent keywords and try to find the learning object that would be the most appropriate for their course.

The main problem in this scenario is that the teacher must constantly mentally switch between two completely different contexts: first the Moodle, where new content is being created, and then LOR, where the existing content is being searched for. Visually and contextually, these two environments are completely different. This results in teacher using the search only as a last resort or when they know exactly what to look for. The correct approach would be to use learning objects not as a solution but as a part of design process. That way they would provide new, fresh ideas and solutions based on experiences from other teachers.

The idea behind the proposed solution is to provide the teacher with an easy and accessible tool for making search queries to the LOR. This brings LOR into a known environment and teacher can use it as a part of design process. By using existing learning objects new course can be designed and created much faster.

Once the course is created the teacher may like to expose his work to others. They can show a complete course or only a part of it. This decision can also be a part of school policy in which all created courses must follow the same exposure and licensing rules.

Default Moodle installation does not have a method for easy exchange of information with other systems. In order to expose learning objects, the teacher must first export the course and then repackage it by using additional tools. This task requires a lot of technical knowledge and can be considered as a complex one.

If a teacher is not able to easily expose the created content to others, they might decide not to do it at all. Therefore some valuable information might not be exposed only due to lack of technical skills.

Besides the tools needed for exposing the data, the teacher should also be supported by an infrastructure that can handle this kind of information. ODS LOR can include results from remote locations if they expose their data using OAI-PMH standard, which is an Open Archives standard based on XML and six defined verbs. It enables remote services to query the data stored on data provider and to retrieve all required metadata. By using internationally recognized standards ODS LOR enables all learning object providers to be harvested and indexed in the same manner after which exposed data is made available to all other users.

The developed plugin enables the teacher to easily expose all content from their Moodle course just by specifying certain data. This action requires no technical knowledge or background, and gives the teacher a complete sharing control and view over exposed data. Moodle system administrator is able to enforce any school policy by defining global settings applicable for all courses, not allowing teachers to change those settings but only to view them. These settings must be available to all learning object Repositories able to harvest data provided through OAI-PMH interface.

Currently, there are more than 1000 plugins available [9]. Many of them are distributed from the official Moodle plugin site. Using the officially distributed plugin ensures users that this product is independently tested for compatibility with their system and no information loss or leak should occur. Official plugins repository also provides means of notifying the system administrator of all updates to the software.

The idea behind the proposed solution is to be available as a part of official Moodle plugins repository and thus to allow full verification by anyone. This is the only way for the solution to be completely safe for both teachers and students to use.

III. Approach

Moodle has several types of plugins that developer can create [10]. One of the available types used to expand Moodle functionality is the block plugin. The block plugin was chosen for this solution because once added in a course it is available on all its main pages. With the block plugin developer can present short and simple information for teachers such as status of the exposed data or simple search field.

This plugin supports a full screen view so the teacher can have full attention on the content, which is most notable in the case of search results. First, user enters the query in a search box that is a part of the course page, and then results are displayed in a larger view.

The block plugin allows two types of configuration: global and local. A system administrator can configure the plugin globally and thus enforce school policy regarding the exposure of learning objects. Using local configuration, the teacher can set sharing options for each course separately.

Two separate Moodle block plugins were developed as a part of the proposed solution. First one provides teacher with easy search access to large LOR and second one enables teacher and system administrators to easily expose the content.

A. ODS Search block plugin

Search plugin enables teachers to make direct search queries from Moodle to ODS Learning Objects Repository. ODS LOR will reply with a list of all found learning objects containing links to them.

System administrator can install the search block using the standard Moodle installation of block plugins. A teacher can add the block plugin using the standard Moodle method of enabling the block plugins.

Once plugin is installed and enabled, the teacher can access the Search Moodle block from any course page as presented in Fig. 1.

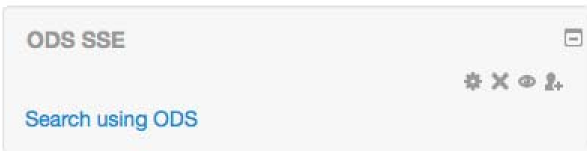

Fig. 1. Example of Moodle Search block plugin when part of page

After selecting the link "Search using ODS", teacher can enter any term and the ODS LOR search response will be displayed inside Moodle. By using search plugin the user does not leave content creation context since all the operation is done using the known Moodle environment as presented in Fig. 2.

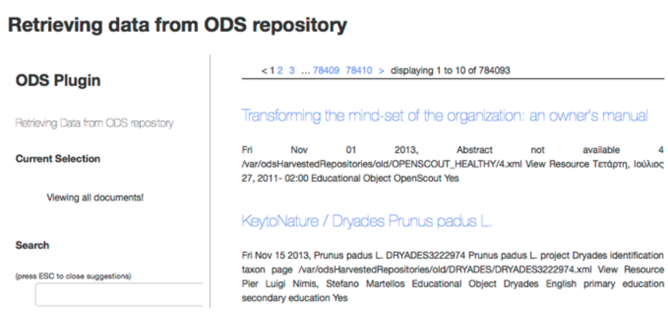

Fig. 2. Example of search results page of Moodle Search block

Search plugin is aware of teacher's course and course content since it is the integral part of Moodle. This allows for plugin to use course title, keywords, description and metadata while searching for learning objects. A profile for the teacher can be made by using that additional course data and metadata. That profile supports better matching between the teacher and the searched content. For example: if the teacher is searching with just the keyword "network", the search plugin can use the available course metadata and decide if the course is about IT or social science. In the first case the correct learning objects would be about computer network, and in the second case it should retrieve objects related to social networks. This is an important part of the matching process between what a teacher requested and automatically generated metadata regarding the learning object from another teacher.

The complete message exchange between user and ODS LOR using Moodle and search block plugin is presented in Fig. 3.

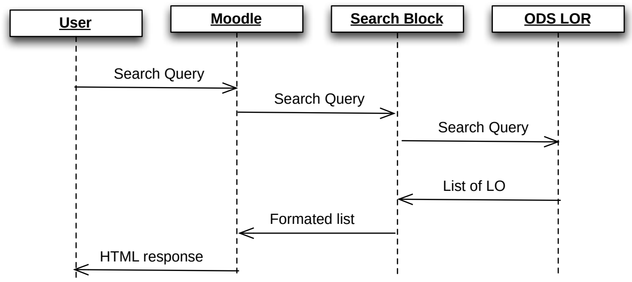

Fig. 3. Search query message exchange

By using the ODS search block plugin, the teacher does not leave the course environment. This enables the teacher to look for new course materials more often since it is much easier to access search.

Some teachers were granted access to the plugin during its development phase. This encouraged them to be more open to the idea of exposing their own content, since they were able to see how useful the plugin is and wanted to contribute to its growth by adding more searchable content. In some cases, the teacher who exposed learning objects got matched with other teachers who used a similar learning methodology, which opened a new communication between them.

B. OAI-PMH sharing block plugin

OAI-PMH sharing plugin enables teachers and system administrators to define rules for exposing all or some of the course metadata. By using this plugin it is not necessary for the teacher to know anything about the method of exposing data to ODS LOR or any other LOR that can harvest OAI-PMH structured data, or anything about creating the learning objects [11]. All the work is done automatically in background by this block plugin.

Sharing block plugin was designed to automatically detect many formats used internally by Moodle. During the process of indexing, it scans through each course instance. The plugin checks if the course metadata can be exposed to other users. If a teacher allows the data to be shared, the OAI-PMH sharing block plugin is going to expose all found learning objects and create all required metadata for all of them. All parts of the process are done automatically and with no required additional information provided by the teacher. At any moment, the teacher can check what metadata is exposed and change sharing settings.

One of the main features of the plugin is the type of contents license, which can be specified for each course separately by a teacher or for all courses on Moodle by a system administrator. All metadata to be harvested is cached so the process of harvesting by a remote system is not noticeable to normal Moodle operation. A system administrator can also limit the access of OAI-PMH target to the selected systems so even high number of OAI-PMH requests can be handled with virtually no impact on server performance.

Teachers can select which Moodle plugins and modules are harvested for each course. This information is available to them at any time together with licensing information. ODS OAI-PMH block plugin has predefined values for all Creative Commons licenses. However, the teacher can also choose a commercial or custom license. The plugin allows the teacher to specify his own custom license that will be provided with all course learning object metadata.

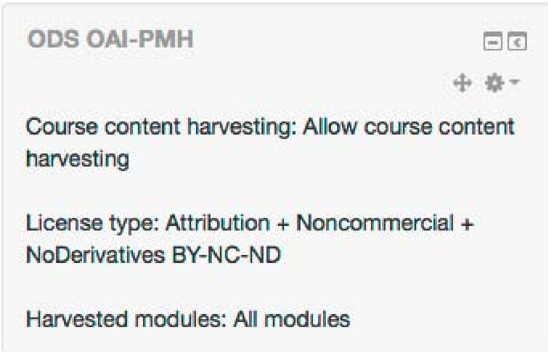

Fig. 4. Example of Moodle Search block plugin as a part of page

The ODS OAI-PMH block is available through official Moodle Plugins repository. This enables all users to confirm the validity of developed block plugin and to be able to easily install and update plugin. Third parties can check the source code since it is fully available on GitHub which allows all developers to contribute to the plugin. Currently, this is the only actively supported OAI-PMH Moodle extension and the only one that supports IMS/IEEE LOM standard for defining data for harvesting.

IV. Conclusion

Currently, teachers are still not using as many learning objects from LORs as possible. There may be plenty of reasons for this, with the main one being that search is still too hard for them to perform because of the constant need to switch between two contexts: creating content and searching for content. Both of these contexts are currently executed in two completely different environments with different design patterns and user experience. Therefore teachers are using Learning Object Repositories only if they know exactly what kind of learning object they want to find, and they are still required to know exact keywords in order to find the desired content.

We propose the usage of an integrated ODS Search Moodle block plugin which resolves this problem and can be further expanded with information about the existing course structure which contains information the teacher is looking for. This extended information can be further used to define more queries and provide teacher with better matching instead of simple search that uses keywords.

The sharing of created content is still not simple enough for many teachers. As a result, even the content that teacher would like to expose is not shared since they lack the required technical knowledge to do that.

We suggest the integrated ODS OAI-PMH Moodle block plugin that automatically checks all available learning objects inside each Moodle course and, as per teacher's instruction, exposes the used learning objects metadata to any OAI-PMH compatible harvester. Thus the teacher has a complete control over this process and can check what data is exposed at any given moment.

This work described two different Moodle block plugins: one that enables teachers to search for new content directly from Moodle and then allows them to use that information as a base for new ideas and teaching methods; the other that enables teachers to expose their work to other teachers in a secure and highly accessible way that has virtually no impact on Moodle performance.

Future work will focus on enhancing queries to the ODS LOR with knowledge on Moodle metadata information of the user, simplifying the process of exposing courses, and also on enhancing the process of automatically generated metadata for learning objects that is provided to OAI-PMH harvesters.

References

[1] Capay, M.; Magdin, M., "Lifelong learning of teachers at primary and secondary schools with the support of e-courses,", *Interactive Collaborative Learning (ICL), 2012 15th International Conference on, 26-28 Sept.2012*, pp. 1-6.

[2] O'Regan H, (2012), "A Brief History of Computing", Springer

[3] Polsani, Pithamber R. (2003), "Use and Abuse of Reusable Learning Objects", Journal of digital information, Volume 3, Number 4, Article No. 164, 2003-02-19.

[4] Noor, S.F.M.; Yusof, N.; Hashim, S.Z.M., "Creating granular learning object towards reusability of Learning Object in e-learning context", *Electrical Engineering and Informatics (ICEEI), 2011 International Conference on, 17-19 July 2011*, pp. 1-5.

[5] ODS, (2015) Open Discovery Space [Online]. Available: http://www.opendiscoveryspace.eu

[6] Moodle, (2015) Moodle [Online]. Available: http://www.moodle.org

[7] Arteaga Sanchey, R.; Duarte Hueros A. (2010), "Motivational factors that influence the acceptance of Moodle using TAM", *Computers in Human Behavior, November 2010, Vol. 26(6),* pp. 1632-1640

[8] Open Archives Initiative, (2015), Open Archives Initiative Protocol for Metadata Harvesting [Online], http://www.openarchives.org

[9] Moodle, (2015) Moodle Plugins [Online]. Available: http://plugins.moodle.org

[10] Moodle, (2015) Moodle Plugin Types [Online]. Available: https://docs.moodle.org/dev/Plugin_types

[11] Najjar J.; Duval E, (2006) "Actual use of learning objects and metadata- an empirical analysis," *TCDL Bulletin, vol. 2, no. 2*.