mEducator 3.0: combining semantic and social web approaches in sharing and retrieving medical education resources

Stathis Th. Konstantinidis, Lazaros Ioannidis, Dimitris Spachos, Charalampos Bratsas, Panagiotis D. Bamidis
Medical Informatics Laboratory, School of Medicine, Aristotle University of Thessaloniki, Greece
cs@med.auth.gr, larjohn@gmail.com, dimitris.spachos@gmail.com, cbratsas@math.auth.gr, bamidis@med.auth.gr

Abstract

Sharing of educational resources over the web has been a key development for both educators and learners in recent years. Pivotal roles in these developments have been played by following principles of the social/collaborative web, and more recently by exploiting advances in the semantic web front. In this paper an architecture that allows the sharing and retrieving of educational resources is proposed. In contrast to similar past attempts, this is now done without the need of copying or storing metadata in a centralised dataset/repository. Linked Data principles were used for exploring the metadata, while collaborative techniques were used in the systems that take advance of this architecture in order to enhance the end user experience, thereby offering a contemporary way to achieve resource adaptation and personalisation.

1. Introduction

During the last decade the scene for online educational resources has dramatically changed. Institutions and organization developed and incorporated e-learning environments to enable the sharing of educational resources across their members in an easy and friendly way. Numerous collaborative tools have been deployed, usually exploiting web and social media (web 2.0) capacities in order for collaborative resources to be created.

At the same time centralized repositories appeared, in an attempt to collaboratively gather resources from different users of the same scientific field[1]. Finding, retrieving and sharing resources were enabled by means of resource metadata.

The intrusion of semantic web and linked data into the Web brought into the foreground the connectivity of knowledge, information and resources. Those resources from different repositories are connected and related and, therefore, a provision for different information retrieval techniques to retrieve educational recourses arises.

Yet, only a few efforts so far have attempted to incorporate the sharing of educational resources through the exploitation of semantic web and linked data principles. mEducator, an EU co-funded project (www.meducator.net), has managed to fill this gap by shifting resource sharing, retrieving, re-using and repurposing to the era of semantic web while maintaining key assets of the social web. mEducator, has compared two alternative solutions for sharing and retrieving educational resources, namely, mEducator2.0 based completely on Web2.0 principles and mEducator3.0, which exploits the richness of web3.0 approaches.

The aim of this paper is to present the latter of the two solutions, that is mEducator 3.0. mEducator 3.0 is based on semantic web services and linked data. Different instantiations have been developed along this endeavour, in an attempt to both experiment with different (well known) systems and collaborative techniques and utilities, but also to secure a solid and wide sustainability plan.

The remainder of this paper is structured as follows. In section 2, we provide a brief account on current trends and approaches in medical repositories, while the era of web 2.0 and 3.0 is briefly exploited. In the following section, the focus is on the mEducator 3.0 architecture, while some practical implementations including web 2.0 technologies are also described in

[1] See for example MedEdPortal, www.mededportal.org

978-0-7695-4887-6/12 $26.00 © 2012 IEEE
DOI 10.1109/SMAP.2012.27

there. In the final section, a discussion on key issues of concern and future work is provided.

2. Setting the scene

Sharing of educational resources has been assigned a catalytic role for online education. Moreover, re-using or re-purposing state-of-the-art educational resources could overall improve training and education, especially in the medical domain.

In order for sharing and retrieving to be feasible, the need of educational standards is imperative. This need drove the educational informatics community to the establishment of numerous standards [1]. A pivotal initiative in the medical domain, has lead to the release of an XML ANSI standard, namely Healthcare LOM [2] which provided a formal way of describing educational resources. However, Healthcare LOM lacks a detailed description when it comes to educational aspects of the description; in addition, it is based on XML without any consideration (or exploitation) of contemporary notions like the "Web of Data"[3].

The mEducator consortium [4] has developed a scheme/ontology, which was originally based upon Healthcare-LOM, but incorporated ten (10) mandatory fields (identifier, title, creator, IPR license, language of resource, language of metadata, creation date, metadata creation data, keywords and description) and a number of optional fields (educational objectives, educational outcomes, assessment methods, educational context, technical description, discipline, etc.) [5]. The mEducator scheme/ontology provides the description of resources in RDF, thereby allowing their interconnections in the Linked Data Cloud.

Repositories for educational resources in the era of Web2.0/3.0 may provide new experiences of sharing resources across federated institutions, if the use of contemporary schemes or ontologies is considered.

2.1. Repositories for educational resources

Some of the repositories contain medically oriented content, while others like Multimedia Educational Resource for Learning and Online Teaching (MERLOT)[2] provide educational content of various fields with medicine being only one of those.

Most of repositories address Intellectual Property Rights (IPR) like iSEEK Medical[3], while a few include repurposing support like Medicine[4]. Since Taxonomies are essential in the medical domain due to the fact that they allow users to find relevant content easily, they are frequently encountered in many repositories, such as EBSCO PUBLISHING[5], COREMINE[6]. However, different taxonomy systems are utilised among the different systems. For example, MERLOT and CurrMIT[7] use grouped taxonomies based on topic and material type, while MedEdPORTAL[8] use taxonomies grouped based on Health Profession, Intended Audience, Instructional Methodology, Resource Type and Academic Focus. [4]

Multilinguality is an issue that most repositories do not address while very few of them make a valuable effort (e.g. EBSCO PUBLISHING).

The systems themselves may be proprietary or they may be distributed under a license. This is a key aspect since it's often that a medical education organization may be interested in having a separate instantiation of a specific system for its own use.

Another discrimination between repositories is whether the repositories hold at least the metadata of the resources in a local dataset, or they harvest different LCMSs that hold and maintain the metadata descriptions of their own resources. Along the latter notion, LOP2P [6] and EduLearn [7] follow a peer-to-peer architecture, while Talis Aspire, being a management system of digital lists of learning resources, facilitates resource linking with other resources from the same or different systems [8].

2.2. The era of Web2.0 and 3.0

Social media and collaborative technologies (web2.0) have recently become part of educational procedures [9][10]. A collaborative wiki or a blog constitutes an educational resource [11] while Social Media like YouTube videos and others are now becoming routinely used in educational processes [12], [13]. While the creation of such resources continues to expand, their discoverability is questionable [14], [15].

On the other hand, end users are becoming more and more literate with social media techniques and collaborative technologies altogether, and, therefore, capable of seeking for them in many different systems [16], [17].

[2] http://www.merlot.org/merlot/index.htm
[3] http://medical.iseek.com/iseek/home.page
[4] http://www.medicinejournal.co.uk/home
[5] http://www.ebscohost.com/schools/multilanguage-searching
[6] http://www.coremine.com/medical/#search
[7] https://www.aamc.org/services/currmit/183178/about_currmit.html
[8] https://www.mededportal.org/

At the same time, the World Wide Web has been changing towards linked data and semantified information to lead to the "Web of Data" [3].

In 2006 Tim Berners-Lee [18] outlined a set of best practices for publishing data on the web, known as "Linked Data principles". These principles HAVE changed the scene in sharing and retrieving knowledge and resources. The Linked Open Data (LOD) cloud holds an enormous number of triples and connects together different datasets across many disciplines.

3. The fruit in the basket

In mEducator, developments have focused on the creation of mechanisms to federate learning content management systems (LCMS's) together. Despite initial attempts for sharing the metadata of the resources through semantic web services, the recent availability of SPARQL 1.1 has allowed for a more versatile architecture that was at the end followed and implemented in mEducator3.0. The architectures of the latter is described below together with the requirements for capitalising it in real systems.

3.1. mEducator3.0 sharing via Semantic Web services

The first architecture developed in mEducator 3.0 consisted of 3 layers [19]: (i) The (Web) data and service layer, (ii) The data and service integration layer, (iii) The application and presentation layer.

The first layer consists of available educational resource metadata, Web services and data sources such as the ones forming part of the Linked Data cloud [19]. The second layer consist of an RDF repository and a REST API that allow different system to store and retrieve data from the RDF repository. It also consists of iServe [20] & SmartLink [21] repositories that are used to gather information from different LCMSs through APIs or WEB services coupling each LCMS [19] . The last layer provides different interfaces to the end user by the use of APIs from the data and service integration layer [19].

3.2. A versatile architecture for federating mEducator 3.0 instantiations

The limitations of the first approach for sharing and retrieving metadata is that it does not combine different repositories of metadata of educational resources, but instead it gathers the metadata into a single RDF store, by means of the developed API. However, combining metadata for educational resources of different LCMSs requires custom development for exposure of metadata at each LCMS; the provided API for SmartLink repository (that gathers those metadata) is not as flexible as the one used at the RDF store. To this extent, there is no connection between the RDF store and the SmartLink repository even if though this could be doable.

For the above reasons the necessity of a global way for sharing and retrieving metadata across different LCMSs was imperative.

To integrate the search functionality of the deployed LCMSs into a single access point, the search federation service was developed [22]. The service exploits the SPARQL 1.1 query-federation functionality to request information from multiple SPARQL endpoints. So each LCMS holds its metadata for the educational resources and the responsibility for maintaining them and exposes them through a SPARQL endpoint that follows the mEducator scheme/ontology. A unique way is used to search the metadata of the educational resources.

The architecture of the search functionality followed in this approach is depicted in figure 1.

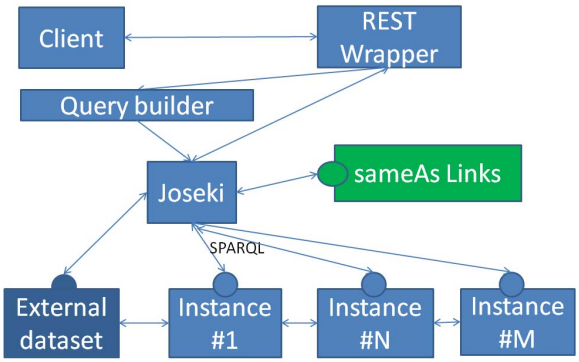

Figure 1: the versatile mEducator 3.0 architecture

The core of the service is based on a Joseki RDF server, which, in turn, is wrapped by an HTTP interface, developed in PHP. The service clients can use a simple syntax to define filters and returned fields, as well as, common Boolean operators and other parameters.

To this extent, this approach includes the SILK framework [23],[24], so as to discover relationships between the metadata of the educational resources and other data sources on the web and to create the RDF links to them. Exact matches between the metadata of educational resources existing in the various platforms and other datasets on the LOD cloud enrich these metadata and create (hopefully useful) interlinks among them.

Queries across distributed educational repositories are conducted in the same way, without the need for creation of services and APIs for each repository. Metadata of educational resources are exposed through SPARQL endpoints. Before the initiation of SPARQL 1.1 the working solution was to query each endpoint, get the results and merge them into a single result set. By the use of "UNION" and "SERVICE" existence in SPARQL 1.1 efficient, optimized federation of the results could be realized. In this case, one creates the same query, but by contacting a single meta-endpoint that contacts the instances and merges the results itself. The implementation of the merging algorithm is guaranteed to be optimized for this use and probably performs better than the previous mEducator solution [19]. To this extent, a RESTful API is provided in order to ease the connection with interfaces and systems.

3.3. mEducator 3.0 in action

Three different systems have been tested within the aforementioned mEducator3.0 architecture: (i) Medical Inter-Linked Educational Space (MILES+) as an extension of the Moodle 2.0 LCMS, (ii) Medical Educational LINked Arena (MELINA+) as an extension of the Drupal 7 CMS, and (iii) Linked Labyrinth (LL+) as an extension of the OpenLabyrinth Virtual Patient platform [25].

In each of the aforementioned systems the end user, upon creating or uploading an educational resource, is obliged to describe it with metadata compliant with mEducator scheme/ontology. Those metadata are stored initially to their local rational database and then they are exposed either by the use of D2R server (or ARC server) to RDF triples, so as the search functionality can retrieve them. In addition functionalities for searching and retrieving the metadata by the use of the RESTful API, as well as, the suitable interface for each platform were created. The end user can find the needed resource without the need to leave the local system.

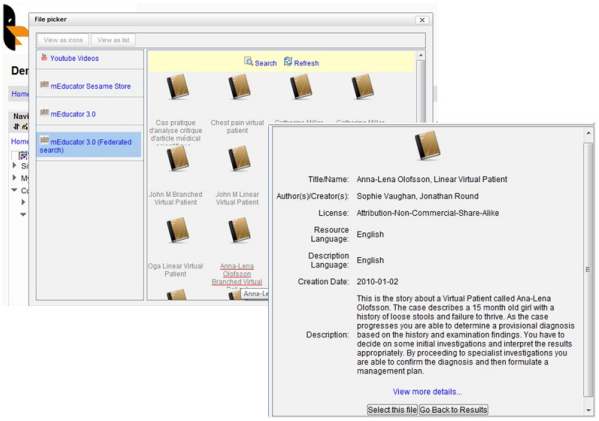

Figure 2: Federated search in MILES+

Moreover, collaborative technologies (web 2.0) where applied into these systems, so as to foster the user sharing experience. The notion of a social network of users where each user can view educational resources created by other users is being exploited. FOAF is used in order to semantify this information. Collaborative creation of the metadata for an educational resource is applied in the case of MELINA+ while the provision of these relationships over the resources themselves may in fact reflect a semantic adaptation of social media principles thereby reflecting a map of connections between educational resources.

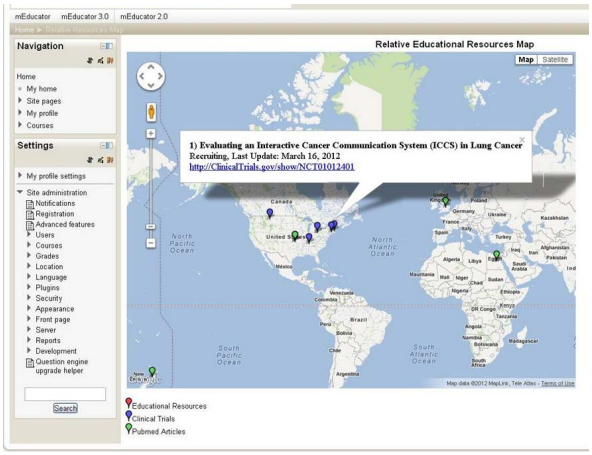

Figure 3: Geotagging of medical educational resources

Educemiology, a newly introduced term in [26], stands for the creation and study of "epidemiological" maps, displaying how medical learning resources from one Institution or academic teacher or expert are distributed to other places and connected to other resources and people, so as to be used in other educational contexts. In addition, mash-up tools representing geographical maps of learning resources with relative material (e.g a PubMed article, a Clinical

Case, etc.) that exist in the web are provided in an effort to enrich the educational experience.

MELINA+ was also extended to allow automatic importing of social media content (YouTube videos) transforming social media metadata to mEducator metadata and consider them as educational resources.

Functionalities like rating an educational resource in 4 different aspects allow the end user to control the quality of the educational resources in a social media way and exclude the "noisy" resources from later uses. However, a proper mechanism for quality review in some systems is also available in case resource peer-review is required.

The repurposing functionality is supported by all systems. A user can repurpose an educational resource (that holds the proper intellectual property license) and describe it within the systems thereby allowing for the display of the evolution and the use of educational resources.

4. Discussion and future work

One of the main limitations associated with the sharing and retrieval of information across different LCMSs is the long delays in SPARQL queries responds. This problem increases while the dataset is growing. However, techniques for caching the results of querying triple stores and compound application objects containing such queries [27] [9] are being developed, while join processing and grouping techniques to minimize the number of remote requests [28] [10] are being investigated.

Providing a solution for sharing, searching and retrieving of medical educational resources across federated LCMSs fosters the exchange of knowledge between health professionals and eases the re-use and repurposing of medical educational resources. At the same time sharing and retrieving resources by the use of collaborative technologies fosters users sharing experience and transfers the "human" or "traditional" way of exchanging educational resources into a virtual sharing sphere.

In this paper we have proposed an architecture that allows sharing and retrieving of educational resources, without the need of copying or storing metadata in a centralized dataset. Linked Data principles were used for exploring the metadata, while collaborative techniques were used in the systems that take advance of this architecture in order to enhance end user experiences. This offers a unique way to adapt resources in a semantically meaningful way, thereby allowing for a contemporary idea for personalization. To this extent, these technologies act as complementary in order to provide the end user all means of searching and retrieving educational resources. Resources from the linked data cloud can be included in user portfolios and be exchanged through collaborative techniques.

Future work will aim to improve the querying results with respect to delays, followed by an evaluation of the sharing mechanism performance.

Placing these techniques in era of the Future Internet or Internet of Things (IoT) looks very promising. Physical educational resources may easily be retrieved through the aforementioned techniques and learning context from virtual resources in medical education could engage physical points and actions.

5. Acknowledgement

This work is supported by "mEducator: Multi-type Content Repurposing and Sharing in Medical Education" (www.meducator.net), a project funded by the eContentplus 2008 program, Information Society and Media Directorate-General, European Commission (ECP 2008 EDU 418006).

6. References

[1] S.Th. Konstantinidis, E. Kaldoudi, P.D. Bamidis, "Enabling Content Sharing in Contemporary Medical Education: A Review of Technical Standards", *The Journal on Information Technology in Healthcare*, vol. 7(6), pp. 363–375, 2009.

[2] V. Smothers. "Healthcare Learning Object Metadata Specifications and Description Document, v0.9", MedBiquitous Consortium 2007, http://www.medbiq.org/working_groups/learning_objects/HealthcareLOMSpecification.pdf, last visited on Dec 2008.

[3] C. Bizer, T. Heath, K. Idehen, T. Berners-Lee (2008) "Linked Data on the Web (LDOW2008)", *In Proceeding of the 17th international conference on World Wide Web*, Beijing, China, 2008, pp. 1265-1266.

[4] PD Bamidis, A. Antoniades, E. Mitsopoulou, M. Nikolaidou, D. Spachos, S.Konstantinidis, C. Bratsas, E. Dafli, C. Pattichis, D. Giordano, e-sharing of multi-type content in medical education: glimpses of best practices and sustainability plans as drawn from the meducator network. *In 'E-EDUCATION & E-SCIENCE", L Toleva-Despotova & PD Bamidis (Eds), ISBN 978-960-243-682-0, 2011.*

[5] Mitsopoulou E, Taibi D, Giordano D, Dietze S; Yu HQ; Bamidis, et al. Connecting medical educational resources to the Linked Data cloud: the mEducator RDF Schema, store and API. Proceeding of the 1st International Workshop on eLearning Approaches for the Linked Data Age within 8th Extended Semantic Web Conference (ESWC2011); 2011 May 29; Heraklion, Greece;2011.

[6] Rafael de Santiago, Andre L.A. Raabe, "Architecture for Learning Objects Sharing amongLearning Institutions-

LOP2P," IEEE Transactions on Learning Technologies, pp. 91-95,April-June, 2010.

[7] Prakash, L. S., Saini, D. K., and Kutti. N. S., 2009. Integrating EduLearn learning content management system (LCMS) with cooperating learning object repositories (LORs) in a peer to peer (P2P) architectural framework. SIGSOFT Softw. Eng. Notes 34, 3 (May 2009), 1-7. DOI=10.1145/1527202.1527212 http://doi.acm.org/10.1145/1527202.1527212

[8] Clarke C. (2009) A Resource List Management Tool for Undergraduate Students based on Linked Open Data Principles. In The Semantic Web: Research and Applications. Lecture Notes in Computer Science, Volume 5554/2009, pp. 697-707.

[9] Kaldoudi E, Konstantinidis S, Bamidis P. Web 2.0 Approaches for Active, Collaborative Learning in Medicine and Health. In: Mohammed S and Fiaidhi J, editors. Ubiquitous Health and Medical Informatics: The Ubiquity 2.0 Trend and Beyond. Hershey, PA, USA: IGI Global; 2010 [chapter 7].

[10] Paton C, Bamidis PD, Eysenbach G, Hansen M, Cabrer M. Experience in the Use of Social Media in Medical and Health Education Contribution of the IMIA Social Media Working Group. Yearb Med Inform. 2011;6(1):21-9.

[11] Boulos, M., Maramba, I., & Wheeler, S. Wikis, blogs and podcasts: A new generation of web-based tools for virtual collaborative clinical practice and education. BMC Medical Education, 6: 41. (2006).

[12] A. Clifton and C. Mann, "Can YouTube enhance student nurse learning?," *Nurse education today*, vol. 31, no. 4, pp. 311-3, May 2011.

[13] L. Sharoff, "Integrating YouTube into the nursing curriculum.," *Online journal of issues in nursing*, vol. 16, no. 3, p. 6, Sep. 2011.

[14] Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 2002 May; 287(20):2691-2700.

[15] Agichtein E, Castillo C, Donato D, Gionis A, Mishne G. Finding high-quality content in social media. Proceedings of the international conference on Web search and web data mining (WSDM '08), 2008 Feb 11-12; Palo Alto, California, USA. 2008: p. 183-94.

[16] K.-J. Lin, "Building Web 2.0," *Computer*, vol. 40, no. 5, pp. 101-102, May 2007.

[17] C. Greenhow, B. Robelia, and J. E. Hughes, "Learning, Teaching, and Scholarship in a Digital Age: Web 2.0 and Classroom Research: What Path Should We Take Now?," *Educational Researcher*, vol. 38, no. 4, pp. 246-259, Jun. 2009.

[18] T. Berners-Lee, (2006). "Linked Data", Retrieved on 10/11/2010, www.w3.org/DesignIssues/LinkedData.html.

[19] H. Q. Yu et al. A Linked Data-driven & Service-oriented Architecture for Sharing Educational Resources A Linked Data-driven & Service-oriented Architecture for Sharing Educational Resources. *Workshop: eLearning Approaches for the Linked Data Age at 8th Extended Semantic Web Conference* June, 2011.

[20] C. Pedrinaci, D. Liu , et al. iServe: a Linked Services Publishing Platform, *Workshop: Ontology Repositories and Editors for the Semantic Web at 7th Extended Semantic Web Conference*. 2010.

[21] S .Dietze, H.Q. Yu, Pedrinaci, C. Liu, D. and J. Domingue. SmartLink: a Webbased editor and search environment for Linked Services, *8th Extended Semantic Web Conference (ESWC)*, Heraklion, Greece. 2011.

[22] L.Ioannidis, C. Bratsas, P.D. Bamidis. A versatile architecture for federating mEducator 3.0 instantiations [Abstract]. In proc. of 1st International Conference on Medical Education Informatics (MEI 2012). Thessaloniki, Greece, 6-7 April, 2012.

[23] C. Bizer, J. Volz, G. Kobilarov, and M. Gaedke, "Silk – A Link Discovery Framework for the Web of Data," 2009.

[24] J. Volz, C. Bizer, and M. Gaedke, "Discovering and maintaining links on the web of data," *The Semantic Web-ISWC 2009*, pp. 650-665, 2009.

[25] Bamidis PD, Konstantinidis ST, Bratsas C, Iyengar M S. Federating Learning Management Systems for Medical Education: a persuasive technologies perspective. Proceedings of the 24th IEEE International Symposium on COMPUTER-BASED MEDICAL SYSTEMS (CBMS 2010);2011 June 27 – 30; University of the West of England, Bristol, UK; 2011.

[26] PD Bamidis. Educemiology: the "Epidemiology" of Medical Learning Resources. [Abstract]. In proc. of 1st International Conference on Medical Education Informatics (MEI 2012). Thessaloniki, Greece, 6-7 April, 2012.

[27] M. Martin and J. Unbehauen, "Improving the performance of semantic web applications with SPARQL query caching," *Web: Research and Applications*, pp. 304-318, 2010.

[28] A. Schwarte, P. Haase, K. Hose, R. Schenkel, and M. Schmidt, "FedX : Optimization Techniques for Federated Query Processing on Linked Data," pp. 601-616, 2011.