Semantic Technologies for Learning and Teaching in the Web 2.0 Era: A survey of UK Higher Education

Thanassis Tiropanis, Hugh Davis, Dave Millard, Mark Weal
LSL, School of Electronics and Computer Science, University of Southampton, UK
{tt2, hcd, dem, mjw}@ecs.soton.ac.uk

Abstract

This paper reports the results of a survey of semantic tools and services relevant to UK higher education. The survey establishes the availability of a number of semantic tools and services for collaborative authoring and annotation, searching and matching, repository and semantic infrastructure development. The relevance of semantic tools and services to learning, teaching and institutional challenges for higher education is discussed. Current developments show that an emerging field of ontology-less linked can provide the ground for added value tools and services for education in the future. These semantic tools and services can initially rely on linked data; mapping linked data to application specific or community agreed ontologies could subsequently provide further added value.

Introduction

The strengths of semantic technologies for learning and teaching and their benefits in the areas of digital libraries, virtual communities and e-learning have been an item of discussion during recent years. The case for semantic technologies in education has been made on the expressive power of metadata to describe learning content, people, and services, and on how these could be intelligently matched for added value services and an advanced learning experience. However, certain concerns on the feasibility of ontology consensus and of the annotation of the enormous amount of content currently available on the Web have arisen making globally available and interoperable semantic-rich metadata for learning resources a long term vision.

The recent emergence of Web 2.0 systems has enabled a lightweight knowledge modelling approach (sometimes called folksonomies) based around techniques such as community tagging, clustering, and community authoring. Such Web 2.0 systems are already used in education [1]. In this respect, it can be argued that semantic technologies are already impacting on the way in which we learn and teach and this will have an increasing impact as the sophistication of the tools increases (as identified at the recent JISC CETIS Conference 2007[1]).

The Joint Information Systems Committee (JISC) in the UK commissioned the SemTech[2] (Semantic Technologies for learning and teaching) project to conduct a survey on the use of semantic technologies in higher education today taking into account these recent developments. One of the first tasks in SemTech was to establish the relevance of semantic technologies to higher education learning, teaching and support challenges. Subsequently, a survey of relevant semantic technologies was conducted. The current adoption of such technologies in the UK higher education sector was researched and is discussed in this paper together with the open challenges and issues that require further attention.

Hard and soft semantic technologies

For this survey we distinguished between hard and soft semantic technologies. *Hard semantic technologies* provide for expressing the meaning of resources and their relationships in machine processable ways and for drawing conclusions (reasoning) based on this meaning with

---

[1] http://ochre.wordpress.com/2007/11/21/semantic-web-session-at-cetis/
[2] http://www.semtech.ecs.soton.ac.uk

mechanisms that are independent of meaning. *Soft semantic technologies* provide for the expression of meaning for resources in ways that can be understood and processed by humans but not by machines. Examples of hard semantic technologies are RDF, FOAF, SKOS or Triple Stores, while examples of soft semantic technologies are traditional tagging tools and topic maps.

Many of the Web 2.0 technologies for education seem to employ soft semantic technologies such as wikis, tagging and topic maps. At the same time, requirements for interoperability across datasources and for more advanced and efficient resource discovery seem to encourage a transition from soft semantic technology to hard semantic technology use. An example of this trend is the transition from the soft semantic knowledge structure in Wikipedia to the hard semantic knowledge structure (in RDF) in *DBpedia*[3] or *Freebase*[4].

Establishing the relevance of semantic technologies to education

The SemTech project engaged with the JISC CETIS Semantic Technology working group and with a number of UK universities and organised a workshop to in London in January 2009 to identify the challenges from a learning and teaching and from the higher education institutions' perspective. Learning and teaching challenges to which semantic tools and services can be relevant include:

❑ Assisting the workflow of course creation delivery and revision by recommending relevant content and people in the context of the course and the institution.
❑ Assisting students by recommending resources that match the topics of their assignments and people that may be able to support their activities.
❑ Group formation for collaborative work based on students' background, personal preferences and successful prior collaboration.
❑ Support for critical thinking and argumentation by visualising arguments and linking relevant discussions.
❑ Efficient support for cross-curricular activities in emerging areas by matching people and resources across schools or departments of the same or different institutions.
❑ More efficient personalised knowledge construction to assist parties involved in learning and teaching.
❑ More efficient support of group knowledge construction as part of teaching and learning activities to encourage innovation.

From the viewpoint of a higher education institution a number of additional challenges were identified that could be addressed by semantic technologies:

❑ Student retention with more effective student support and student progress monitoring. Data currently scattered across different databases in each institution could be linked with semantic technologies to support advanced data analysis.
❑ Transparency of data held by institutions could be supported by exposing their various databases to relevant parties in semantic, interoperable formats such as RDF.
❑ The knowledge capital of universities and information on the organisational resources is often in different formats, which makes it often inaccessible to funding bodies looking for expertise. Semantic technologies could enable information integration, searching and matching to this end.
❑ Collaboration across institutions cannot easily be facilitated since the relevant information systems of universities are not interoperable and the deployment of linked data repositories in each institution involves additional costs. Large repositories, like triple stores, to which information can be efficiently stored, searched and managed, could be of benefit.

---

[3] http://dbpedia.org
[4] http://www.freebase.com/

❑ There is lack of a framework to enable each institution to state the IPR the resources they are otherwise eager to expose in order to attract funding and students. Declarative semantic formats could unambiguously state this information and encourage knowledge dissemination.

Categorising semantic technologies for education

The survey that was performed on the availability of tools and services that relate to learning, teaching and institutional challenges resulted in the identification of over thirty relevant tools and services[5], which can be coarsely classified into four main categories based on their main types of functionality:

❑ Collaborative authoring and annotation tools, including semantic wikis, argumentation tools, like *AceWiki*[6], *Cicero*[7], *Mymory*[8], *Kiwi*, *Compendium*[9], *Debategraph*[10], *PROWE*[11].

❑ Searching and matching tools using semantic technologies, like *ArnetMiner*[12],*Twine*[13], *Watson*[14].

❑ Repositories and VLEs that import/export their data using semantic technologies, include *Freebase*, *DBpedia*, *CIP*[15], *Project Gutenberg*[16], *MyExperiment*[17]. Repositories for scientific resources which can expose metadata in RDF include *DSpace*[18] and *EPrints*[19].

❑ Infrastructural tools and services that enable exposing databases or integrating datasources inside or across organisations in interoperable semantic formats, include: *D2R Server*[20], *TALIS*[21], *Virtuoso*[22], *RKBexplorer*[23], *Yahoo! SearchMonkey*[24].

In response to the challenges identified in the previous section, collaborative content authoring and annotation tools and services seem to provide:

❑ Representation of shared knowledge with precision and recommendation of related content and people for collaborative activities. Semantic wikis like *AceWiki* and *Kiwi* and tools like *ArnetMiner* are of relevance.

❑ Documentation and support of collaboration workflows on a larger or smaller scale for teaching and learning activities and for collaboration across departments and institutions. *MyExperiment* provide for the documentation of experiment workflows.

❑ Support for argumentation and visualisation of arguments and relevant resources to enable critical thinking. Tools like *Debategraph* and *Cicero* support argumentation.

❑ Representation of the shared knowledge capital of higher education institutions in ways that can be accessed by different faculties, schools, institutions and the public. It can be argued that

---

[5] http://semtech-survey.ecs.soton.ac.uk/
[6] http://attempto.ifi.uzh.ch/acewiki/
[7] http://cicero.uni-koblenz.de/wiki/index.php/Main_Page
[8] http://www.dfki.uni-kl.de/mymory/
[9] http://www.aktors.org/technologies/compendium/
[10] http://debategraph.org/
[11] http://www.prowe.ac.uk/
[12] http://www.arnetminer.org/
[13] http://www.twine.com/
[14] http://watson.kmi.open.ac.uk/WatsonWUI/
[15] http://www.ilrt.bris.ac.uk/whatwedo/projectsaz/project?search=CIP
[16] http://www.gutenberg.org
[17] http://www.myexperiment.org/
[18] http://www.dspace.org/
[19] http://www.eprints.org/
[20] http://www4.wiwiss.fu-berlin.de/bizer/d2r-server/
[21] http://www.talis.com/platform
[22] http://virtuoso.openlinksw.com/
[23] http://www.rkbexplorer.com
[24] http://developer.yahoo.com/searchmonkey/

semantic wikis can be employed to this end. At the same time, collaborative knowledge modelling as per *Freebase* can also be of relevance.

Searching and matching tools can support:

❑ Contextualised queries and searches, searches across repositories potentially in different departments or institutions, and matching of people for collaborative activities. Best example of the surveyed technologies to this end is *ArnetMiner*.

❑ More efficient Question and Answer systems and knowledge bases for learning and teaching support. Semantic search framework like *Yahoo! SearchMonkey* could enable the development of such systems.

❑ Exposing the expertise of the institution to the outside world in order to attract funding and students. *ArnetMiner* is the most representative example of such tools at the moment.

❑ Combination of information from different sources within institutions to enable better monitoring of student progress and to provide recommendations based on declarative statements that can be validated. None of the surveyed tools seem to address this challenge at the moment.

Repositories, VLEs and authoring tools can provide for:

❑ Semantic annotation of content to support more precise knowledge construction, interoperability and integration of repositories across institutions. Representative examples are *EPrints*, *DSpace*, *DBPedia*, *Freebase*, *Project Gutenberg*.

❑ Semantic enrichment of classifications of repositories to enable more efficient resource discovery and interoperability. This is happening to an extent in *Freebase* and *DBpedia*.

Infrastructural technologies can offer:

❑ Large repositories for efficient storage and search of data from different sources in different institutions and repositories. An example of such a repository is *RKBexplorer*.

❑ Exposing part of organisational data to partners or the Web, Formatting data in interoperable machine-processable formats and integration of data from different sources. A number of infrastructural tools such as *TALIS*, *Virtuoso* and *D2R Server* that can enable semantic enrichment and exposure in semantic formats.

Semantic technology adoption in the UK higher education

Between ten and twenty UK universities appear to have adopted the use of wikis on an institutional level to support learning and teaching. Current activities on the development of semantic wikis indicate awareness of the advantages of adding meaning to the relationships among wiki resources. In addition, reasoning to support argumentation, where the relevance between arguments can be precisely identified and used for navigation and visualisation of discussions. This indicates a trend for institutional adoption of semantic wikis and argumentation tools in the near future.

Searching and matching making use of semantic technologies is featured by repositories that already employ semantic metadata. Expert matching as provided by the *ArnetMiner* has been deployed by the University of Tsinghua in China [2]. The JISC funded *AWESOME*[25] project provides software that combines semantic wiki and pedagogy-aware inline recommendations to empower academic writing and is used by a number of schools at the University of Leeds, University Coventry and University of Bangor.

Repositories are the most widely adopted type of infrastructure by the UK education institutions. Over forty universities are reported to employ repositories in the UK to publish their research

---

[25] http://awesome.leeds.ac.uk

results, conference and journal articles, presentations or course material. Commonly used platforms are DSpace and EPrints. The fact that they are both adding support for RDF [3, 4] shows potential adoption of semantic technologies for educational repositories.

A handful of Universities expose SPARQL endpoints (for example the University of Southampton[26] and the University of Oxford[27]) but for the moment these concern research projects. Requirements for repository integration and queries on larger datasets might encourage the use and further development of infrastructural semantic tools and services like large linked data repositories and semantic enrichment and data integration platforms.

Open challenges and responses
The Semantic Web or Web 3.0 vision has inspired significant research output and there is an emerging consensus that some form of Semantic Web is an inevitable development of existing technologies [5]. The view that the adoption of semantic aware applications for education is placed in a horizon of four to five years [6] but we believe that this activity will just commence during this period.

Anticipating the emergence of a linked data [7] field and the population of this field with interoperable semantic data before agreement on ontologies and advanced pedagogically meaningful applications will become available is a key conclusion from our survey. We also foresee an emergence of advanced reasoning applications that will not be built on a well-established linked data field and will most likely rely on a limited number of repositories and on ad-hoc mapping of unstructured data to specific ontologies.

Further work and research to identify and address the issues related to fostering the development of a global linked data filed based on optimised metadata repositories across education institutions is critical. Development and availability of tools that will assist collaborative ontology building and efficient mapping of linked data to those ontologies will be decisive in the development of advanced pedagogically meaningful semantic tools and services.

References
[1] Anderson, P. (2007) What is Web 2.0? Ideas, technologies and implications for education, JISC Technology & Standards Watch, Feb. 2007 http://www.jisc.ac.uk/media/documents/techwatch/tsw0701b.pdf

[2] Tang, J., Zhang, J., Yao, L., Li, J., Zhang, L., and Su, Z. 2008. ArnetMiner: extraction and mining of academic social networks. In Proceedings of the 14th ACM SIGKDD international Conference on Knowledge Discovery and Data Mining (Las Vegas, Nevada, USA, August 24 - 27, 2008). KDD '08. ACM, New York, NY, 990-998. DOI= http://doi.acm.org/10.1145/1401890.1402008

[3] Huynh, D., Mazzocchi, S., and Karger, D. (2005) Piggy Bank: Experience the Semantic Web Inside Your Web Browser. Semantic Web Conference (ISWC) 2005, Galway, Ireland.

[4] Harris, S. and Gibbins, N. (2003) 3store: Efficient Bulk RDF Storage. In Proceedings 1st International Workshop on Practical and Scalable Semantic Web Systems, Sanibel Island, Florida, USA.

[5] Ohler, J. (2008); The Semantic Web in Education – What happens when the read-write web gets smart enough to help us organize and evaluate the information it provides? EDUCAUSE Quarterly Magazine, Volume 31, Number 4

[6] Johnson, L., Levine, A., & Smith, R. (2009). The 2009 Horizon Report. Austin, Texas: The New Media Consortium.

[7] O'Hara, K. and Hall, W. (2009) Semantic Web. In: Marcia J. Bates; Mary Niles Maack; Miriam Drake (eds), Encyclopedia of Library and Information Science Second Edition, Taylor & Francis. ISBN 978-0-8247-2075-9 (In Press) http://eprints.ecs.soton.ac.uk/17126/1/hall-ohara-elis-semantic-web.pdf

---

[26] http://imageweb.zoo.ox.ac.uk/wiki/index.php/DefiningImageAccess/Project/SCULPTEUR
[27] http://zoo-garos.zoo.ox.ac.uk/ibrg/index.php/Image_Bioinformatics_Research_Group_home_page