Campbell, L.M., Barker, P., Currier, S., & Syrotiuk, N., (2013). The Learning Registry: social networking for open educational resources? In Proceedings of OER13: Creating a Virtuous Circle. Nottingham, England. Available from http://www.medev.ac.uk/oer13/108/view/ *and* http://publications.cetis.ac.uk/2013/770

The Learning Registry: social networking for open educational resources?

Lorna M. Campbell and Phil Barker, Cetis

lmc@strath.ac.uk, phil.barker@hw.ac.uk

Sarah Currier and Nick Syrotiuk, Mimas

sarah.currier@manchester.ac.uk, syrotiuk@manchester.ac.uk

Abstract

This presentation will reflect on Cetis' involvement with the Learning Registry and Jisc's Learning Registry Node Experiment at Mimas (The JLeRN Experiment), and their application to UKOER initiatives. Initially funded by the US Departments of Education and Defense, the Learning Registry (LR) is an open source network for storing and distributing metadata and curriculum, activity and social usage data about learning resources across diverse educational systems.

The LR's innovative technical methodology applies a new approach to the perennial problems of describing and managing OERs. Rather than mandating specific standards, the LR is metadata agnostic; it ingests all kinds of resource descriptions and data into a document-oriented, schema-free database. Described as "social networking for metadata", the LR should make it possible to mine networks to build useful services based on the context in which educational resources are used, and the conversations users have around them.

Since its inception, there has been sustained interest in the LR from a number of innovative technologists and developers within the UKOER community. Part of Cetis' and JLeRN's brief was to engage this group to explore the applicability of the LR approach to UKOER initiatives, and to scope the type of services that could usefully be built on top of an open LR node. As a result, a small number of projects and developers have engaged with JLeRN and the LR and have established the technical feasibility of implementing the LR architecture and ingesting and extracting data, and have also demonstrated a range of innovative services that can be built on top of an LR node. However, the network effect currently remains unproven, as there are only a handful of nodes in existence.

In exploring the benefits and drawbacks of the LR approach, JLeRN and CETIS surfaced a number of issues. These include managing expectations of the LR's promise in light of its early stage of development; examining the technical skills and capacity available in the sector; identifying the benefits of adopting the LR versus other technologies (e.g. basic schema-free noSQL databases, RDF triple stores); and exploring the value of the LR approach to subject areas with clearly defined curricula and learning outcomes.

The problem of sharing the educational context and value of resources is one that the sector has struggled with for many years. The LR may not conclusively solve this messy problem, but by taking a new approach to the challenge, it certainly merits further attention.

Keywords

activity streams, Cetis, CouchDB, curriculum data, ENGrich, Jisc, JLeRN, Jorum, learningreg, metadata, networks, nodes, OER, OERRI, OERs, open educational resources, paradata, Pgogy, RIDLR, schema-free databases, SPAWS, The Learning Registry, UKOER, usage data.

Introduction

In this paper we reflect on Cetis' involvement with the Learning Registry and Jisc's Learning Registry Node Experiment at Mimas (The JLeRN Experiment), and their potential application to UKOER initiatives. Initially funded by the US Departments of Education and Defense, the Learning Registry (LR) is an open source network for storing and distributing metadata and curriculum, activity and social usage data about learning resources across diverse educational systems. The JLeRN Experiment was commissioned by Jisc to explore the affordances of the Learning Registry for the UK F/HE community within the context of the HEFCE funded UKOER programmes.

Most initiatives aimed at managing and sharing information about learning resources rely on agreeing a metadata schema to describe those characteristics of the resources that are of relevance to their user community. Learning resources represent a diverse class of objects, and identifying and describing the characteristics that are likely to be relevant to a wide range of stakeholders is a non-trivial task; as a result, a plethora of educational metadata standards, application profiles, schema and vocabularies have emerged (Barker and Campbell, 2010). In addition to describing the educational characteristics of learning resources, many of these schema also attempt to describe general and technical characteristics. However, they are not particularly successful in this regard and are rarely adequate for describing media specific characteristics of various resource types. Consequently it is often necessary to use format-specific and resource-type specific metadata (e.g. MARC for bibliographic records; ID3 tags for audio, EXIF for images, etc.) alongside educational metadata schema. Furthermore, social tools and applications that enable users to share comments and opinions about resources and the contexts in which they are used are increasingly being employed to discover learning resources. Capturing this secondary usage data adds yet another level of complexity. In short, the whole environment for sharing information about learning resources is characterised by diversity: diversity of resource types, resource descriptions, user requirements and educational practice. In such an environment it is little wonder that creating a single metadata schema that meets all requirements is problematic.

The Learning Registry approach is to sidestep this problem by allowing metadata of any kind to be stored in a schema-free database. Schema-free databases, also referred to as NoSQL databases or document-oriented databases, are a class of database that have been adopted increasingly widely in recent years. In contrast to relational databases, they do not rely on data conforming to a pre-set schema; instead documents are stored as a collection of key-value pairs in JSON format. The "key" identifies a property or characteristic that is described by the value, e.g. name: "Jo Bloggs". The allowed content of documents is not predefined, and once the data is accumulated, it can be interrogated on the basis of the keys that are present and the content of the documents. This approach is increasingly widely used for managing heterogeneous data sets and large scale distributed applications.

The Learning Registry

The Learning Registry describes itself as:

> "a new approach to capturing, sharing, and analysing learning resource data to broaden the usefulness of digital content to benefit educators and learners".

It is not a repository or a search engine, instead it seeks to complement such services by providing the technical means to exchange and aggregate information about learning resources and how they are used, including formal metadata descriptions and informal comments on resources and their use. Potentially, the Learning Registry may also capture information about

the educational context in which resources were used, e.g. course name, curriculum alignment, etc., from learning environments and course management systems. The vision is that over time, a stream of data about the resource builds up, forming a timeline similar to the timelines used by social networking sites. While some of the data forming these timelines is formal metadata, much of it is subjective or contextual information, a class of information that has become known as paradata.

The Learning Registry was designed to work as a large scale network with no one central point of control. It is implemented as nodes, each based on an instance of the CouchDB schema-free database, on top of which is an Application Programming Interface (API) that allows nodes to exchange data with other nodes and with external services. The exchange of data between nodes may be partial, depending on the policies of each node. The support provided for external services, e.g. the ability to provide customised subsets of the data, may also vary between nodes. It is the external services that provide user-facing functionality to facilitate resource discovery; the Learning Registry is merely the “plumbing” that allows the data to flow.

The JLeRN Experiment

As part of its remit to advise Jisc on innovative learning technologies, CETIS maintained a watching brief on the Learning Registry and liaised closely with the development team in the US. As the Learning Registry approach appeared to be well aligned with requirements for developing a technical infrastructure to support the release and management of OERs arising from the UKOER Programmes, Cetis recommended that Jisc establish an experimental node. As a result, Jisc’s Learning Registry Node Experiment (JLeRN) was set up at Mimas to explore the feasibility of setting up and running an LR node, to contribute and analyse data, and support the development of use cases and applications relevant to UK F/HE. The emphasis of the project was on evaluating the affordances of the LR approach and building expertise to inform future decisions, rather than on developing a sustainable service. Cetis worked closely with the JLeRN team to facilitate liaison with the LR developers in the US and to build a special interest community in the UK of projects and institutional partners who were keen to explore the potential of the LR approach.

Projects and partners that engaged with the JLeRN experiment included:

- Sharing Paradata Across Widget Stores (SPAWS): aimed to use the Learning Registry to share usage data, such as reviews, ratings, and download statistics between web app stores of widgets and gadgets for educators.
- Rapid Innovation Dynamic Learning Maps—Learning Registry (RIDLR): set out to test the release of contextually rich paradata via the JLeRN node to the Learning Registry and to harvest back paradata to provide resource discovery linked to specific topics displayed within the context of the curriculum and personal learning maps.
- EngRICH: aimed to design, develop and evaluate a customised search engine for visual media relevant to engineering education. Information about student ratings and recommendations are stored in their own Learning Registry node and used to enhance customised Google searches.
- Independent developer Pat Lockley of Pgogy, developed several tools for interacting with Learning Registry nodes including: Ramanathan, which submits information from an RSS feed to the Learning Registry; and Pliny, which submits Google Analytics data to the Learning Registry.
- Jorum worked with JLeRN to test the ingest of the repository’s metadata, via an OAI-PMH feed.

The JLeRN Experiment also commissioned a Wider Potential Report (Kay, 2012) that focused on the potential of the Learning Registry architecture and conceptual approach, looking beyond its core educational technology and focusing on the broader information environment.

Issues and Observations

The first observation to be made is that all the projects found that, on a technical level at least, the Learning Registry worked well. Setting up and configuring node software presented no difficulties once developers understood the basic requirements of CouchDB and the other technologies used, and the LR API met and supported their needs. The technical guides and documentation available from the main Learning Registry site, though occasionally oblique, were helpful in this respect. There are however two significant aspects of the Learning Registry that remain untested.

The Learning Registry was originally conceived as a network of nodes automatically sharing data but, to date, there is little evidence of this happening. The JLeRN Experiment made no attempt to share data between nodes as there was no immediate requirement to do so from the community the project was working with. This means that the functionality of the APIs for distributing data between nodes is less well tested than the APIs for interfacing with services external to the Learning Registry. Stand-alone nodes do have benefits, both as a data store for a single service (e.g. EngRICH) or as a means of sharing data between services (e.g. the JLeRN node, SPAWS), however they cannot exploit the benefits that could potentially result from sharing data between nodes, while maintaining community specific features of individual nodes. For example a UK HE node might selectively obtain information about resources suitable for HE from multiple sources, including nodes that covered all levels of education; the EngRICH node might selectively obtain information about resources relevant to Engineering from a generic UK HE node. The UK HE node and the EngRICH node could then filter and process this data in ways that are specific to the services that use these nodes. As this functionality has not been tested, it is difficult to judge how effectively the Learning Registry will work at network scale.

A more fundamental caveat is that few projects have yet to build production level services for end users on top of an LR node. This is significant because the LR approach is based on the assumption that by ingesting data from multiple sources, regardless of schema, a critical mass of data will build up, and by providing APIs to this data, developers will build services and applications to process the data and provide meaningful information to end users. While this approach potentially opens the door for developers to start building new and innovative services for educators, it also pushes the overhead of processing the data from the data providers to the service developers. Though these features remain untested at present, work is ongoing among LR developers and adopters in the US.

Another issue explored by the JLeRN community was whether there are real benefits to be gained from working with the Learning Registry, as opposed to using “vanilla” schema-free databases such as Mongo and CouchDB. One obvious benefit is that in addition to defined APIs, the LR provides a level of support and documentation that is valuable for developers who may be new to working with schema-free databases. However there is also a potential drawback in that the LR locks developers into using a specific database, CouchDB, and it is not clear if the architecture is sufficiently flexible to swap CouchDB for other schema-free databases, such as Mongo, should they become more widely used for data storage.

Similarly, are there any advantages to adopting the Learning Registry rather than semantic technology such as RDF triple stores? Triple stores have been the innovative technology of choice for sharing data on a web-wide scale for a decade or more, but their uptake in the

education domain has been slow, possibly due to the steep learning curve associated with such technologies. The general consensus from the JLeRN community suggested that the Learning Registry's open approach to dealing with messy educational data seemed to fit the ethos of the teaching and learning sector better.

Conclusions

Although the JLeRN Experiment met with considerable technical success, and a small community of developers in the UK have demonstrated the type of tools and services that can be built around LR nodes, the overall impact on the UK F/HE sector has been negligible to date. Learning Registry deployment and implementation is on-going in the US and appears to be gathering some momentum, particularly in the K-12 sector. It appears, from initial developments, that the Learning Registry approach may be most readily applicable to education sectors and domains that have a well-defined shared curriculum, e.g. K-12, medical education, but it remains to be seen whether there will be significant uptake in the UK F/HE sector more widely. However the technical solutions adopted by the Learning Registry still represent a genuinely innovative approach to the thorny problem of managing and sharing learning resource descriptions and contextual data. Even if the Learning Registry itself is not widely adopted in the UK, it seems likely that the innovative technology approaches employed, particularly the use of schema-free databases and the focus on managing and sharing paradata and usage data, will have some impact on the education technology landscape in the longer term.

References

Barker, P. (2008) *Learning Materials Application Profile Scoping Study, final report*. Retrieved from *http://publications.cetis.ac.uk/2008/376*

Barker, P., & Campbell L.M. (2010). Metadata for learning materials: an overview of existing standards and current developments. *Technology, Instruction, Cognition, and Learning* 7(3-4), 225-243. Also available from *http://publications.cetis.ac.uk/2010/79*

Cetis (n.d.). Retrieved 21st February, from *http://jisc.cetis.ac.uk/*

CouchDB (2012), Retrieved 21st February, from *http://couchdb.apache.org/*

ENGRich (n.d.). Retrieved 20th February 2013, from *http://engrich.liv.ac.uk/*

Jisc (n.d.). Academy / JISC Open Educational Resources Programme. Retrieved 20th February 2013, from *http://www.jisc.ac.uk/oer*

JLeRN Experiment (n.d.). Retrieved 20th February 2013, *http://jlernexperiment.wordpress.com/*

Jorum (n.d.). Retrieved 20th February 2013, *http://www.jorum.ac.uk/*

Kay, D., (2012), *The Learning Registry – Wider potential*. Retrieved from *http://jlernexperiment.files.wordpress.com/2012/11/121109-jlern-wider-potential-report-dk1.pdf*

Learning Registry (2012). Retrieved 20th February 2013, from *http://www.learningregistry.org/*

Learning Registry (2011). Technical Guides. Retrieved 20th February 2013, from *http://www.learningregistry.org/documents*

Pgogy (n.d.). Pliny - Analytics to Paradata. Retrieved 20th February 2013, from *http://www.pgogy.com/pliny/*

Pgogy (n.d.). Ramanathan. Retrieved 20th February 2013, from *http://www.pgogy.com/ramanathan/*

RIDLR (n.d.). OER Rapid Innovation. Retrieved on 21st February, from *http://www.medev.ac.uk/blog/oer-rapid-innovation-ridlr/*

Wilson, S. (2012). Sharing usage data about web apps between widget stores. Retrieved 20th February 2013, from *http://scottbw.wordpress.com/2012/10/19/sharing-usage-data-about-web-apps-between-stores/*

License and Citation