



> Semantic Web Use Cases and Case Studies

Case Study: A Linked Open Data Resource List Management Tool for Undergraduate Students

Chris Clarke, Talis Information Limited and Fiona Greig, University of Plymouth, United Kingdom

January 2009



University of Plymouth

Context

Resource Lists, or Course Reserves as they are known in the US, are collections of text books, journal articles, Web pages and/or audio visual content defined by instructors, intended to be companions for students to degree courses, modules or assignments. Traditionally, these were provided to the student as a paper handout comprising a list of citations, the intention being for the student to visit to the campus library or bookshop to obtain the resources.

Increasingly, over the last 10 to 15 years, these materials have become available over the Web, with publishers and content owners moving to an e-delivery model for both monographs and serials. Instructors have been making increasing use of blogging, podcasting and multimedia learning objects in their syllabuses. Resource List Management (RLM) tools emerged over this period, moving paper lists to online electronic bulletin board.

Traditional RLM tools have gone a long way to solve some of the problems of paper-based lists, most notably in the area of stock management. Electronic manifestations of lists that link to the resources in the University's library catalogue can be used to indicate to library acquisition staff in advance that one hundred first year physics students will be vying for that single copy of the core text at the start of term.

Limitations of the traditional RLM tools

One could be forgiven for assuming that the move to electronic resources in itself has improved the access landscape somewhat, but it has also brought an era where the same single journal subscription can be fulfilled from a myriad of different publisher or aggregator platforms. As an institution shifts subscription agents from year to year, the landscape for what access rights a University has purchased becomes complex. Links to content harvested by the instructor into the RLM to e-content may work this year, but expire or move to a different platform the next.

Institutional link resolver solutions go some way to solving these issues, but often represent a poor experience for the student, and are potentially forgotten about when harvesting. Instead of searching the resolver and harvesting the resource's resolver page, many instructors naturally harvest directly from the source Web page to which the resolver points.

In addition, the interoperability of data between publishing platforms, the University library catalogue, the University Virtual Learning Environment (VLE) and RLM tools themselves has been poor.

On the subject of the user experience for the student, RLM tools have often been no more than electronic replicas of the paper based list. There has traditionally been little attempt to allow students to annotate items with regard to their intended use, or interact with content. Students could not feedback whether they have found items useful in attaining their learning goals, or add personal study notes. There has been no facility to allow the students themselves to form new collections or bibliographies of resources for a particular essay or group exercise.

For instructors, the task of discovering new or revised content to update existing or create new lists was largely unaddressed by existing systems above offering a simple bookmarking system or search interface over the local University library catalogue. With hundreds if not thousands of instructors facing a similar task of creating that first year physics resource list worldwide, surely there must be a more collaborative solution?

General Description

In partnership with 15 UK and Eire Universities (the focus group), Talis embarked on a project in 2007 to replace their existing RLM tool and seek to address some of these issues. It was Talis' belief that the use of Semantic Web technology could be used in the following areas:

- Unify the descriptions of resources using existing ontologies
- Improve the interoperability of the data using linked open data principles
- Design for (data) appropriation—using existing schemas and ontologies such as FOAF and SIOC to describe the relationships means that the open data in the application is more likely to be used elsewhere in other contexts

- Encourage semantic enrichment of the data by students and instructors to enable context-aware recommendation functionality

As well as re-using existing ontologies, Talis developed and published two new ontologies as part of the project. The Resource List ontology [1] underpins the semantics of the relationships between resources and intended uses. The Academic Institution Internal Structure ontology (AIISO) [2] describes the courses, modules, departments and schools that make up an institution, which was required to enable instructors to link lists to the relevant module or course, to enable students to find lists.

On the issue of semantic enrichment, Talis developed several features designed not only to improve the experience for the student, but also increase the richness of the data (see Figure 1). For example, using the Resource List ontology, the new system allows the student to annotate each item on the list, augmenting the instructor’s class notes with their own, and allowing to categorize each of the resources by intention (e.g. “Planning to read”, “Reading now”). The student can then use the tool to manage their library of resources, typically running into the thousands, much more effectively. Additionally, these intentions can be used to enhance feedback given to instructors as to the popularity of the resources, by overlaying the raw usage statistics gathered from Web server logs with the intentions of the students in the class.

The screenshot shows the University of Plymouth website interface. At the top, there is a navigation bar with 'Home', 'My Resources', and 'Sign out'. The main content area is titled 'Financial accounting and reporting' and includes a description: 'The 12th edition of Elliott and Elliott is the recommended text to buy for this module.' Below this is a book entry for 'FINANCIAL ACCOUNTING AND REPORTING' by Barry Elliott and Jamie Elliott. The entry includes a 'Get this item' button and a table of metadata: Item type (Book), Author(s) (Elliott, Barry., Elliott, Jamie.), Publisher (Financial Times/Prentice Hall), Edition (12th ed.), ISBN 10 (0273712314), ISBN 13 (9780273712312), and Date (2008.). To the right of the book entry is a 'My Notes' section with a rich text editor containing the text: 'Chapters 2 and 3 are useful for the 3rd seminar - more as an example of a poor argument than a good one!'. Further to the right is a 'Have you read this?' section with radio buttons for 'Decided not to Read', 'Have Read' (selected), 'Planning to Read', and 'Reading Now', along with a 'Set...' button. At the bottom right, there is a 'Library availability' section.

Figure 1: The student experience (individual item view)

The lists themselves can be constructed using sections (see Figure 2), as defined in the Resource List ontology, to group resources according to topics, time periods, or importance. For example an instructor may group items into a time-based section (e.g. “Lecture 1”) or use a topic-based (e.g. “Accounting for Intangibles”) arrangement. This, along with any class notes the instructors add, enhance the semantics of the list, and indicate intended uses for a group of resources. This makes the lists much more semantically-rich, and provides high-quality contextual data for the construction of recommendation services.



Advanced Financial Accounting and Reporting

Used by [ABF311](#)Group by: [Section](#) | [Format](#)

Table of contents [hide]
1. Module Webpage
2. Recommended Textbook
3. Supplementary Resources
4. Assessment
5. Problem Based Learning Resources
5.1. Problem Statement 1
5.2. Problem Statement 2
5.3. Problem Statement 3
6. Lecture Resources
6.1. Welcome Lecture
6.2. Consolidation 1 and 2
6.3. Consolidation 3 and 4
6.4. Accounting for leases
6.5. Accounting for taxation
6.6. Capital transactions and distributable profit
6.7. Accounting for changing prices 1 and 2
6.8. Accounting for changing prices 3
6.9. Foreign currency translation

(48 items)

Module Webpage (1 item)	
There are useful resources for all topics here	
ABF311 Homepage Get this item Availability, buying options and notes	Web Page
Recommended Textbook (5 items)	
Financial accounting and reporting - Elliott, Barry., Elliott, Jamie. , 2008. Get this item Availability, buying options and notes This the most recent edition of Elliott and the recommended text to buy for this module unless you already have the 11th edition (please see below)	Book
Companion website Elliott and Elliott 12th edition Get this item Availability, buying options and notes	Web Page

Figure 2: Outline view of the list showing sections

For the instructor, the system includes several other enhancements aimed at making harvesting resources and the creation of lists much easier.

Using a bookmarklet, and in a similar pattern to sites such as del.icio.us, instructors can harvest resources from a multitude of Web sites, including the library catalog. The difference between the system and most bookmarking services is that the system attempts to identify the resource the page is describing, rather than just recording the location of the page itself. A series of recognisers are applied to the page content to determine the particulars of the item being harvested. Recognisers use a variety of techniques, including querying of other sources, based on facts gleaned from the page content, or evaluating microformat-style markup such as COinS [\[3\]](#) to pull metadata from the page. Only if a resource cannot be identified by automatic means is the user prompted to supply metadata and classify the type of resource being harvested.

Once obtained, the metadata is stored in the instructor's library as RDF using the Bibliographic Ontology [\[4\]](#) increasing the interoperability of the harvested data with other systems and workflows. Bibliographic Ontology is already in use by a number of other groups, including Zotero [\[5\]](#) a tool to help collect research sources, and the LIBRIS project at the National Library of Sweden [\[6\]](#) which is a linked open data online library catalogue.

By storing metadata about the resource being described, rather than the page describing it, more resilient strategies can be employed to ensure content links do not break if the library decides in future to change supplier. Much of this more sophisticated logic is simply invoked by the user by clicking the "Get this item" button (see [Figure 1](#)). The philosophy of this feature is to get the user as close to the resource as possible. Talis is working to ensure that in future, as much content as possible is embedded in the page without linking to an external location, in the style of YouTube and the Google Books services, which display in-line multimedia viewers for content in the

page rather than the user having to download binary files via a hyperlink and view them in an external application.

The system uses the Talis Platform [9] to store, index and query across its RDF data. The application makes use of a group of distinct stores to separate graphs with different concerns and levels of privacy (see Figure 3). For example, publicly linkable and discoverable data is stored in a different store from private user content or details, such as preferences, passwords or private study notes. Data from these stores can be merged within the system at run time when required. Using this approach encourages re-use of the public data by other applications. The full set of Talis Platform APIs, such as SPARQL and search capabilities, can be offered externally without exposing sensitive or private data.

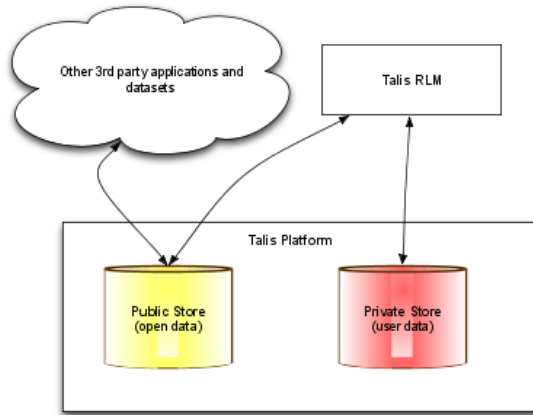


Figure 3: Partitioning of data stores

The interface to build or edit lists uses a WYSIWYG metaphor implemented in Javascript operating over RDFa markup, allowing the user to drag and drop resources and edit data quickly, without the need to round trip back to the server on completion of each operation. The user's actions of moving, adding, grouping or editing resources directly manipulate the RDFa model within the page. When the user has finished editing, they hit a save button which serialises the RDFa model in the page into an RDF/XML model which is submitted back to the server. The server then performs a delta on the incoming model with that in the persistent store. Any changes identified are applied to the store, and the next view of the list will reflect the user's updates.

Using the approach of direct manipulation of RDFa, which is then posted this back to the server has two advantages. Firstly, in elegance simplicity of implementation—rather than having to keep track of multiple operations or edits to data in the persistent store via complex form submission or multiple AJAX calls, we are simply allowing the user to directly edit an RDF model locally (albeit via an appropriately familiar and usable metaphor) and simply send one form field containing the new model back to the server to update the stored copy. Secondly, by embedding RDFa in the HTML view of the entity, we are providing systems integrators or even indexing agents (such as Yahoo's SearchMonkey) another route to discovering the semantics of the data represented in the page.

Interoperability of data

As well as standardising the description of resources by using ontologies in use by other groups to enhance linking, the RLM tool encourages re-use of that data by adhering to linked open data principles.

Unlike many Web applications, the system is architected to use a noun-based URL structure to address entities, rather than a verb-based approach designed to describe operations (see Table 1). Operations on these entities can then be described by using the 4 HTTP verbs—GET (view), POST (create or modify), PUT (replace), DELETE (remove). This method allows us to embrace and work with the design of the web and its underlying protocols, rather than against, and makes for a more predicable and consistent behavior across the application. More discussion of this approach can be found at [7].

This approach lends itself well to publishing linked open data, with URLs representing entities rather than actions, and the system uses content negotiation approaches [8] to allow each URL in the system to be served as XHTML (for humans) and RDF/XML or JSON for software agents.

URL Structure

Noun-based	<code>http://lists.lib.plymouth.ac.uk/lists/abf203</code>
Verb-based	<code>http://example.org/view?courseCode=ABF203</code>

Table 1: URL Examples

Using this approach makes it very easy for system integrators to either simply link to the system's human-consumable interpretation of the user interface, or create their own interpretation within their respective VLE or institutional portal environments. Additionally, other services on the Web can link to resources described within the system to augment their own descriptions, and mine rich relationships to other related topics or resources.

By using existing ontologies and publishing the data within the RLM as linked open data, the tool not only unifies the description of learning resources, but it also unlocks silos of semantically-rich descriptions of inter-related resources as described by domain experts (the instructors).

In use at University of Plymouth

The system was launched at the University of Plymouth, one of the existing focus group partners, in September 2008, initially with just 1000 students. Throughout the autumn semester, Plymouth have been increasing adoption of the system with the aim of giving access to all 22000 students in early 2009.

The future

During the first half of 2009, Talis will be rolling the system out to further focus group partners, and will continue to implement the system across its RLM customer base of 1/3rd of UK Universities (representing over 800,000 students and over 40,000 instructors) towards the end of the year and into 2010.

Having unified the descriptions of lists and added semantic richness to the data, future development can focus on the issues around context-aware recommendation systems. This will suggest related items either to instructors when creating lists (taking into account recognised topics or contexts) or suggesting alternatives to students when consuming items.

Key Benefits of Using Semantic Web Technology

- Re-using established ontologies to unify descriptions of resources and enable data interoperability with other systems and data sets without in-depth knowledge of the application. This also had the side effect of reducing the amount of proprietary modeling required and as such cut time to market
- Simplify the editing interface by allowing the user to interact with RDFa markup
- Using linked open data principles to allow linking within the linked open data cloud
- Using content negotiation principles for ease of integration to 3rd party systems
- Enabling semantic annotation of resources by instructors and students
- Exploring relationships between resources, topics, people and annotations, stored as RDF, for the purposes of building recommendation systems

References

1. <http://vocab.org/resourcelist/schema-20080519.html>
2. <http://vocab.org/aiiso/schema-20080514>
3. <http://ocoins.info/>
4. <http://purl.org/ontology/bibo/>
5. <http://www.zotero.org>
6. <http://libris.kb.se/?language=en>
7. <http://dynamicorange.com/2008/10/03/pages-screens-mvc-and-not-getting-it/>
8. <http://www4.wiwiw.fu-berlin.de/bizer/pub/LinkedDataTutorial/>
9. http://www.talis.com/platform/resources/assets/harnessing_sophisticated_mass.pdf

© Copyright 2009, [Talis Information Limited](#), [University of Plymouth](#)