The OU Linked Open Data: Production and Consumption

Fouad Zablith, Miriam Fernandez, and Matthew Rowe

Knowledge Media Institute (KMi), The Open University
Walton Hall, Milton Keynes, MK7 6AA, United Kingdom
{f.zablith,m.fernandez,m.c.rowe}@open.ac.uk

Abstract. The aim of this paper is to introduce the current efforts toward the release and exploitation of The Open University's (OU) Linked Open Data (LOD). We introduce the work that has been done within the LUCERO project in order to select, extract and structure subsets of information contained within the OU data sources and migrate and expose this information as part of the LOD cloud. To show the potential of such exposure we also introduce three different prototypes that exploit this new educational resource: (1) the OU expert search system, a tool focused on finding the best experts for a certain topic within the OU staff; (2) the Social Study system, a tool that relies on Facebook information to identify common interest between a user's profile and recommend potential courses within the OU; and (3) Linked OpenLearn, an application that enables exploring linked courses, Podcasts and tags to OpenLearn units. Its aim is to enhance the browsing experience for students, by detecting relevant educational resources on the fly while studying an OpenLearn unit.

Keywords: Linked Data, education, expert search, social networks.

1 Introduction

The explosion of the Linked Open Data (LOD) movement in the last few years has produced a large number of interconnected datasets containing information about a large variety of topics, including geography, music and research publications among others. [2]

The movement is receiving worldwide support from public and private sectors like the UK[1] and US[2] governments, international media outlets, such as the BBC [5] or the New York Times [1], and companies with a social base like Facebook.[3] Such organisations are supporting the movement either by releasing large datasets of information or by generating applications that exploit it to connect data across different locations.

[1] http://data.gov.uk
[2] http://www.data.gov/semantic/index
[3] http://developers.facebook.com/docs/opengraph

R. García-Castro et al. (Eds.): ESWC 2011 Workshops, LNCS 7117, pp. 35–49, 2012.

Despite its relevance and the support received in the last few years, very few pieces of work have either released or exploited LOD in the context of education. One of these few examples is the DBLP Bibliography Server Berlin,[4] which provides bibliographic information about scientific papers. However, education is principally one of the main sectors where the application of the LOD technologies can provoke a higher impact.

When performing learning and investigation tasks, students and academics have to go through the tedious and laborious task of browsing different information resources, analysing them, extracting their key concepts and mentally linking data across resources to generate their own conceptual schema about the topic. Educational resources are generally duplicated and dispersed among different systems and databases, and the key concepts within these resources as well as their inter and intra connections are not explicitly shown to users. We believe that the application of LOD technologies within and across educational institutions can explicitly generate the necessary structure and connections among educational resources, providing better support to users in their learning and investigation tasks.

In this context, the paper presents the work that has been done within The Open University (OU) towards the release and exploitation of several educational and institutional resources as part of the LOD cloud. First, we introduce the work that has been done within the LUCERO project to select, extract and structure subsets of OU information as LOD. Second, we present the potential of this data exposure and interlinking by presenting three different prototypes: (1) the OU expert search system, a tool focused on finding the best experts for a certain topic within the OU staff; (2) the Social Study system, a tool focused on exploiting Facebook information to identify common interests between a user and recommend potential courses within the OU, and; (3) Linked Open Learn, an application that enables exploring linked courses, Podcasts and tags to OpenLearn units.

The rest of the paper is organised as follows: Section 2 presents the state of the art in the areas of LOD within the education context. Section 3 presents the work that has been done within the LUCERO project to expose OU data as part of the LOD cloud. Sections 4, 5 and 6 present example prototype applications that consume the OU's LOD for Expert Search, Social Study and Linked OpenLearn respectively. Section 7 describes the conclusions that we have drawn from this work, and section 8 presents our plans for future work.

2 Related Work

While LOD is being embraced in various sectors as mentioned in the previous section, we are currently witnessing a substantial increase in universities adopting the Linked Data initiative. For example, the University of Sheffield's Department of Computer Science[5] provides a Linked Data service describing research

[4] `http://www4.wiwiss.fu-berlin.de/dblp/`
[5] `http://data.dcs.shef.ac.uk`

groups, staff and publications, all semantically linked together [6]. Similarly the University of Southampton has recently announced the release of their LOD portal (http://data.southampton.ac.uk), where more data will become available in the near future. Furthermore, the University of Manchester's library catalogue records can now be accessed in RDF format[6]. In addition, other universities are currently working on transforming and linking their data: University of Bristol,[7] Edinburgh (e.g., the university's buildings information is now generated in LOD[8], and Oxford[9]. Furthermore the University of Muenster announced a funded project, LODUM, the aim of which is to release the university's research information as Linked Data. This includes information related to people, projects, publications, prizes and patents.[10]

With the increase of the adoption of LOD publishing standards, the exchange of data will be much easier, not only within one university, but also across the LOD ready ones. This enables, for example, the comparison of specific qualifications offered by different universities in terms of courses required, pricing and availability.

3 The Open University Linked Open Data

The Open University is the first UK University to expose and publish its organisational information in LOD.[11] This is accomplished as part of the LUCERO project (Linking University Content for Education and Research Online)[12], where the data extraction, transformation and maintenance are performed. This enables having multiple hybrid datasets accessible in an open way through the online access point: http://data.open.ac.uk.

The main purpose of releasing all this data as part of the LOD cloud is that members of the public, students, researchers and organisations will be able to easily search, extract and, more importantly, reuse the OU's information and data.

3.1 Creating the OU LOD

Detailed information about the process of LOD generation within the OU is available at the LUCERO project website.[12] We briefly discuss in this section the steps involved in the creation of Linked Data. To achieve that, the main requirement is to have a set of tools that generate RDF data from existing data sources, load such RDF into a triple store, and make it accessible through a web access point.

[6] http://prism.talis.com/manchester-ac
[7] https://mmb.ilrt.bris.ac.uk/display/ldw2011/University+of+Bristol+data
[8] http://ldfocus.blogs.edina.ac.uk/2011/03/03/university-buildings-as-linked-data-with-scraperwiki
[9] http://data.ox.ac.uk
[10] http://www.lodum.de
[11] http://www3.open.ac.uk/media/fullstory.aspx?id=20073
[12] http://lucero-project.info

Given the fact that the OU's data repositories are scattered across many departments, using different platforms, and subject to constant update, a well-defined overflow needs to be put in place. The initial workflow is depicted in Figure 1, and is designed to be efficient in terms of time, flexibility and reusability. The workflow is component based, and the datasets characteristics played a major role in the implementation and setup of the components. For example, when the data sources are available in XML format, the XML updater will handle the process of identifying new XML entities and pass them to the RDF extractor, where the RDF data is generated, and ready to be added to (or removed from) the triple store. Finally the data is exposed to the web, and can be queried through a SPARQL endpoint.[13]

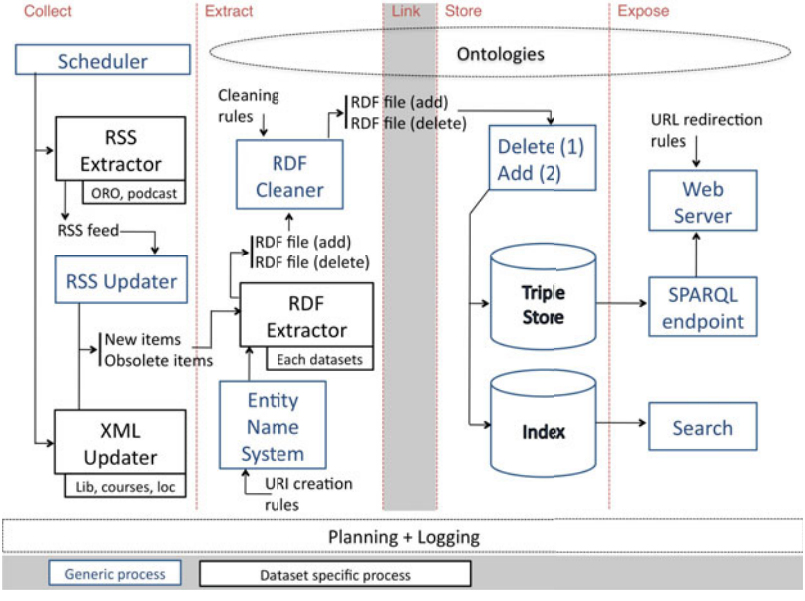

Fig. 1. The LUCERO workflow

The scheduler component takes care of initiating the extraction/update process at specific time intervals. This update process is responsible for checking what was added, modified, or removed from the dataset, and accordingly applies to the triple store the appropriate action. Having such a process in place is important in the OU scenario where the data sources are continuously changing. Another point worth mentioning is the linking process that links entities coming from different OU datasets (e.g., courses mentioned in Podcast data and library records), in addition to linking external entities (e.g., course offerings in

[13] `http://data.open.ac.uk/query`

a GeoNames defined location[14]). To achieve interlinking OU entities, independently from which dataset the extraction is done, we rely on an Entity Named System, which generates a unique URI (e.g., based on a course code) depending on the specified entity (this idea was inspired from the Okkam project[15]) . Such unique URIs enable a seamless integration and extraction of linked entities within common objects that exist in the triple store and beyond, one of the core Linked Data requirements [3].

3.2 The Data

Data about the OU courses, Podcasts and academic publications are already available to be queried and explored, and the team is now working to bring together educational and research content from the university's campus information, OpenLearn (already available for testing purposes) and library material. More concretely, data.open.ac.uk offers a simple browsing mechanism, and a SPARQL endpoint to access the following data:

- The Open Research Online (ORO) system[16], which contains information about academic publications of OU research. For that, the Bibliographic Ontology (bibo)[17] is mainly used to model the data.
- The OU Podcasts,[18] which contain Podcast material related to courses and research interests. A variety of ontologies are used to model this data, including the W3C Media Ontology,[19] in addition to a specialised SKOS[20] representation of the iTunesU topic categories.
- A subset of the courses from the Study at the OU website,[21] which provides courses information and registration details for students. We model this data by relying on the Courseware,[22] AIISO,[23] XCRI,[24] MLO[25] and GoodRelations ontologies [4], in addition to extensions that reflect OU specific information (e.g., course assessment types).
  Furthermore, there are other sources of data that are currently being processed. This includes for example the OU YouTube channel,[26] the library catalogue, and public information about locations on the OU campus (e.g., buildings) and university staff.

[14] http://www.geonames.org
[15] http://www.okkam.org
[16] http://oro.open.ac.uk
[17] http://bibliontology.com/specification
[18] http://podcast.open.ac.uk
[19] http://www.w3.org/TR/mediaont-10
[20] http://www.w3.org/2004/02/skos
[21] http://www3.open.ac.uk/study
[22] http://courseware.rkbexplorer.com/ontologies/courseware
[23] http://vocab.org/aiiso/schema
[24] http://svn.cetis.ac.uk/xcri/trunk/bindings/rdf/xcri_rdfs.xml
[25] http://svn.cetis.ac.uk/xcri/trunk/bindings/rdf/mlo_rdfs.xml
[26] http://www.youtube.com/user/TheOpenUniversity

4 The OU Expert Search

Expert search can be defined as the task of identifying people who have relevant expertise in a topic of interest. This task is key for every enterprise, but especially for universities, where interdisciplinary collaborations among research areas is considered a high success factor. Typical user scenarios in which expert search is needed within the university context include: a) finding colleagues from whom to learn, or with whom to discuss ideas about a particular subject; b) assembling a consortium with the necessary range of skills for a project proposal, and; c) finding the most adequate reviewers to establish a program committee.

As discussed by Yimam-Seid and Kobsa [7], developing and manually updating an expert system database is time consuming and hard to maintain. However, valuable information can be identified from documents generated within an organisation [8]. Automating expert finding from such documents provides an efficient and sustainable approach to expertise discovery.

OU researchers, students and lecturers constantly produce a plethora of documents, including for example conference articles, journal papers, thesis, books, reports and project proposals. As part of the LUCERO project, these documents have been pre-processed and made accessible as LOD. The purpose of this application is therefore to exploit such information so that OU students and researchers can find the most appropriate experts starting from a topic of interest.[27]

4.1 Consumed Data

This application is based on two main sources of information: (a) LOD from the Open Research Online system, and (b) additional information extracted from the OU staff directory. The first information source is exploited in order to extract the most suitable experts about a certain topic. The second information source complements the previous recommended set of experts by providing their corresponding contact information within the OU. Note that sometimes, ex-OU members and external collaborators or OU researchers may appear in the ranking of recommended experts. However, for those individuals, no contact information is provided, indicating that those experts are not part of the OU staff.

As previously mentioned, the information provided by Open Research Online contains data that describe publications originating from OU researchers. In particular, among the properties provided for each publication, this system exploits the following ones: a) the title, b) the abstract, c) the date, d) the authors and, e) the type of publication, i.e., conference paper, book, thesis, journal paper, etc.

To exploit this information the system performs two main steps. Firstly when the system receives the user's query, i.e., the area of expertise where a set of experts need to be found (e.g., "semantic search"), the system uses the title and

[27] The OU Expert Search is accessible to OU staff at:
`http://kmi-web15.open.ac.uk:8080/ExpertSearchClient`

abstract of the publications to find the top-n documents related to that area of expertise. At the moment n has been empirically set to 10.

Secondly, once the top-n documents have been selected, the authors of these documents are extracted and ranked according to five different criteria: (a) original score of their publications, (b) number of publications, (c) type of publications, (d) date of the publications and, (e) other authors of the publication.

The initial score of the publications is obtained by matching the user's keyword query against the title and the abstract of the OU publications. Publications that provide a better match within their title and abstract against the keywords of the query are ranked higher. This matching is performed and computed using the Lucene[28] text search engine. Regarding the number of publications, authors with a higher number of publications (among the top-n previously retrieved) are ranked higher. Regarding the type of publication, theses are ranked first, then books, then journal papers, and finally conference articles. The rationality behind this is that an author writing a thesis or a book holds a higher level of expertise than an author who has only written conference papers. Regarding the date of the publication, we consider the 'freshness' of the publications and continuity of an author's publications within the same area. More recent publications are ranked higher than older ones, and authors publishing in consecutive years about a certain topic are also ranked higher than authors that have sporadic publications about the topic. Regarding other authors, experts sharing a publication with fewer colleagues are ranked higher. The rationality behind this is that the total knowledge of a publication should be divided among the expertise brought into it, i.e., the number of authors. Additionally we also consider the order of authors in the publication. Main authors are considered to have a higher level of expertise and are therefore ranked higher.

To perform the first step (i.e., retrieving the top-n documents related to the user's query) we could have used the SPARQL endpoint and, at run-time, searched for those keywords within the title and abstract properties of the publications. However, to speed the search process up, and to enhance the query-document matching process, we have decided to pre-process and index the title and abstract information of the publications using the popular Lucene search engine. In this way, the fuzzy and spelling check query processing and ranking capabilities of the Lucene search engine are exploited to optimise the initial document search process.

To perform the second step, once the top-n documents have been selected, the rest of the properties of the document (authors, type, and date) are obtained at run-time using the SPARQL endpoint.

Finally, once the set of authors have been ranked, we look for them in the OU staff directory (using the information about their first name and last name). If the author is included in the directory, the system provides related information about the job title, department within the OU, e-mail address and phone number. By exploiting the OU staff directory we are able to identify which experts are

[28] `http://lucene.apache.org/java/docs/index.html`

members of the OU and which of them are external collaborators, or old members not further working for the institution.

Without the structure and conceptual information provided by the OU LOD, the implementation of the previously described ranking criteria, as well as the interlinking of data with the OU staff directory, would have required a huge data pre-processing effort. The OU LOD provides the information with a fine-grained structure that facilitates the design of ranking criteria based on multiple concepts, as well as the interlinking of information with other repositories.

4.2 System Implementation

The system is based on lightweight client server architecture. The back end (or server side) is implemented as a Java Servlet, and accesses the OU LOD information by means of HTTP requests to the SPARQL endpoint. Some of the properties provided by the LOD information (more particularity the title and the abstract of the publications) are periodically indexed using Lucene to speed-up and enhance the search process by means of the exploitation of its fuzzy and spell checker query processing, and ranking capabilities. The rest of the properties (authors, date, and type of publications) are accessed at run time, once the top-n publications have been selected.

The front end is a thin client implemented as a web application using HTML, CSS and Javascript (jQuery).[29] The client doesn't handle any processing of the data, it only takes care of the visualisation of the search results and the search input. It communicates with the back-end by means of an HTTP request that passes as a parameter the user's query and retrieves the ranking of authors and their corresponding associated information by means of a JSON object.

4.3 Example and Screenshot

In this section, we provide an example of how to use the OU expert search system. As shown in Figure 2, the system receives as a keyword query input "*semantic search*", with the topic for which the user aims to find an expert. As a result, the system provides a list of authors ("*Enrico Motta*", "*Vanessa Lopez*", etc), who are considered to be the top OU experts in the topic. For each expert, if available, the system provides the contact details (department, e-mail, phone extension) and the top publications about the topic. For each publication, the system shows its title, the type of document, and its date. If the user passes the cursor on the top of the title of the publication, the summary is also visualised (see the example in Figure 2 for the publication "*Reflections of five years of evaluating semantic search systems*"). In addition the title of the publication also constitutes a link to its source page in data.open.ac.uk.

[29] `http://www.jquery.com`

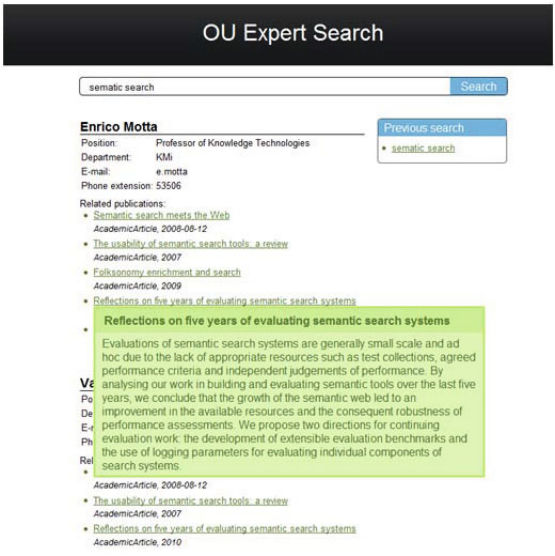

Fig. 2. The OU Expert Search system

5 Social Study

The Open University is a well-established institution in the United Kingdom, offering distance-learning courses covering a plethora of subject areas. A common problem when deciding on which course to study is choosing a course that is relevant and close to an individual's interests. One solution to this problem is to take advantage of existing profile information to bootstrap the decision process, in doing so leveraging information describing a person's interests upon which possible course for studying could be pursued.

Based on this thesis, Social Study[30] combines the popular social networking platform Facebook with the OU Linked Data service, the goal being to suggest Open University courses that share common themes with a user's interests.

5.1 Consumed Data

Social Study combines information extracted from Facebook with Linked Data offered by The Open University, where the former contains the profile information describing a given user - i.e. his/her interests, activities and '*likes*' - while the latter contains structured, machine-readable information describing courses offered by The Open University.

Combining the two information sources, in the form of a 'mashup', is performed using the following approach. First the user logs into the application – using Facebook Connect – and grants access to their information. The application then extracts the concepts that the user has expressed an interest in on his/her profile.

[30] `http://www.matthew-rowe.com/SocialStudy`

In Facebook such interests can be expressed through one of three means: *interests* - where the user explicitly states that they are interested in a given subject or topic; *activities* - where the user describes his/her hobbies and pastimes, and; *likes* - where the user identifies web pages that he/she is interested in that are then shared with the individual's social network.

This collection of concepts extracted from each of these interest facets provides the profile of the given user. To suggest courses from this collection, the OU SPARQL endpoint is queried for all courses on offer, returning the title and description of each course. This information is then compared with the profile of the user as follows: each of the concepts that make up the user's interest profile - in the form of ngrams - are compared against the description and title of each course, and the frequency of concepts matches is recorded for each course.

The goal of Social Study is to recommend relevant courses to the user given their interests, therefore the greater the number of concept matches, the greater the likelihood that the course is suited to the user. The courses are then ranked based on the number of overlapping concepts, allowing the user to see the most relevant courses at the top of the list, together with the list of concepts that informed the decision for the rank position of the course.

If for a moment we assume a scenario where Linked Data is not provided by the OU, then the function of Social Study could, in theory continue, by consuming information provided in an alternative form - given that the query component for the course information could be replaced by another process. However, the presence of Linked Data made the effort required to access and process courses information minimal. This work was an evolution of previous work that attempted to utilise the terms found in wall posts on Facebook in order to inform study partners and relevant courses. In evolving such work our intuition is that the user interest profile that is presented on such a platform can be bootstrapped to suggest courses, thereby avoiding the time-consuming task of manual profile population - from which course suggestions would then be derived.

5.2 System Implementation

The application is live and available online at the previously cited URL. It is built using PHP, and uses the Facebook PHP Software Development Kit (SDK)[31]. Authentication is provided via Facebook Connect,[32] enabling access to Facebook information via the Graph API. The ARC2 framework[33] is implemented to query the remote SPARQL endpoint containing The Open University's Linked Data, and parse the returned information accordingly.

5.3 Example and Screenshot

To ground the use of Social Study, Figure 3 shows an example screenshot from the application when recommending courses for Matthew Rowe – one of the

[31] `https://github.com/facebook/php-sdk`
[32] `http://developers.facebook.com/docs/authentication`
[33] `http://arc.semsol.org`

authors of this paper. The screenshot displays to the end user the order of courses together with the common interest concepts that their interest profile shares with those courses. The top-ranked course "*The technology of music*" matches the interest concepts *music* and *techno* that the user has specified in their profile. The greater the number of shared concepts with the course is, the greater the likelihood that the user will be interested in the course.

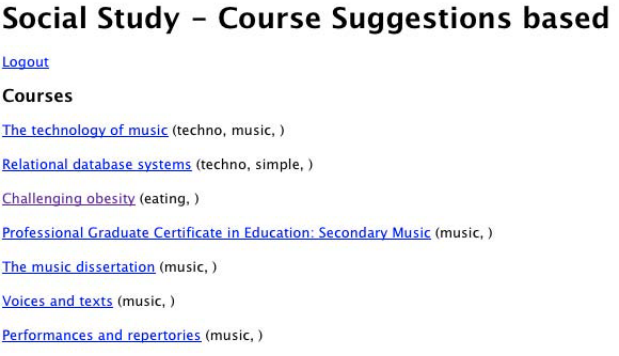

Fig. 3. Social Study showing the top ranked courses together with the interest concepts

6 Linked OpenLearn

The Open University offers a set of free learning material through the OpenLearn website.[34] Such material cover various topics ranging from Arts[35], to Sciences and Engineering.[36] In addition to that, the OU has other learning resources published in the form of Podcasts, along with courses offered at specific presentations during the year. While all these resources are accessible online, connections are not always explicitly available, making it hard for students to easily exploit all the available resources. For example, while there exists a link between specific Podcasts and related courses, such links do not exist between OpenLearn units and Podcasts. This leaves it to the user to infer and find the appropriate and relevant material to the topic of interest.

Linked OpenLearn[37] is an application that enables exploring linked courses, Podcasts and tags to OpenLearn units. It aims to facilitate the browsing experience for students, who can identify on the spot relevant material without leaving the OpenLearn page. With this in place, students are able, for example, to easily find a linked Podcast, and play it directly without having to go through the Podcast website.

[34] `http://openlearn.open.ac.uk`
[35] OpenLearn unit example in Arts:
`http://data.open.ac.uk/page/openlearn/a216_1`
[36] A list of units and topics is available at: `http://openlearn.open.ac.uk/course`
[37] `http://fouad.zablith.org/apps/linkedopenlearn`

6.1 Consumed Data

Linked OpenLearn relies on The Open University's Linked Data to achieve what was previously considered very costly to do. Within large organisations, it's very common to have systems developed by different departments, creating a set of disconnected data silos. This was the case of Podcasts and OpenLearn units at the OU. While courses were initially linked to both Podcasts and OpenLearn in their original repositories, it was practically hard to generate the links between Podcasts and OpenLearn material. However, with the deployment of Linked Data, such links are made possible through the use of coherent and common URIs of represented entities.

To achieve our goals of generating relevant learning material, we make use of the courses, Podcasts, and OpenLearn datasets in data.open.ac.uk. As a first step, while the user is browsing an OpenLearn unit, the system identifies the unique reference number of the unit from the URL. Then this unique number is used in the query passed to the OU Linked Data SPARQL endpoint (http://data.open.ac.uk/query), to generate the list of related courses including their titles and links to the study at the OU pages.

In the second step, another query is sent to retrieve the list of Podcasts related to the courses fetched above. At this level we get the Podcasts' titles, as well as their corresponding downloadable media material (e.g., video or audio files), which enable users to play the content directly within the OpenLearn unit page. Finally the list of related tags are fetched, along with an embedded query that generates the set of related OpenLearn units, displayed in a separate window. The user at this level has the option to explore a new unit, and the corresponding related entities will be updated accordingly. The application is still a prototype, and there is surely room for further data to extract. For example, once the library catalogue is made available, a much richer interface can be explored by students with related books, recordings, computer files, etc.

6.2 System Implementation

We implemented the Linked OpenLearn application in PHP, and used the ARC2 library to query the OU Linked Data endpoint. To visualise the data on top of the web page, we relied on the jQuery User Interface library,[38] and used the dialog windows for displaying the parsed SPARQL results. The application is operational at present, and is launched through a Javascript bookmarklet, which detects the OpenLearn unit that the user is currently browsing, and opens it in a new iFrame, along with the linked entities visualised in the jQuery boxes.

6.3 Example and Screenshot

To install the application, the user has to drag the applications' bookmarklet[39] to the browser's toolbar. Then, whenever viewing an OpenLearn unit, the user

[38] `http://www.jqueryui.com`

[39] The bookmarklet is available at: `http://fouad.zablith.org/apps/linkedopenlearn,` and has been tested in Firefox, Safari and Google Chrome.

clicks on the bookmarklet to have the related entities displayed on top of the unit page. Figure 4 illustrates one arts related OpenLearn unit (referenced earlier), with the connected entities displayed on the right, and a running Podcast selected from the "Linked Podcasts" window. The user has the option to click on the related course to go directly to the course described in the Study at the OU webpage, or click on linked tags to see the list of other related OpenLearn units, which can be browsed within the same window.

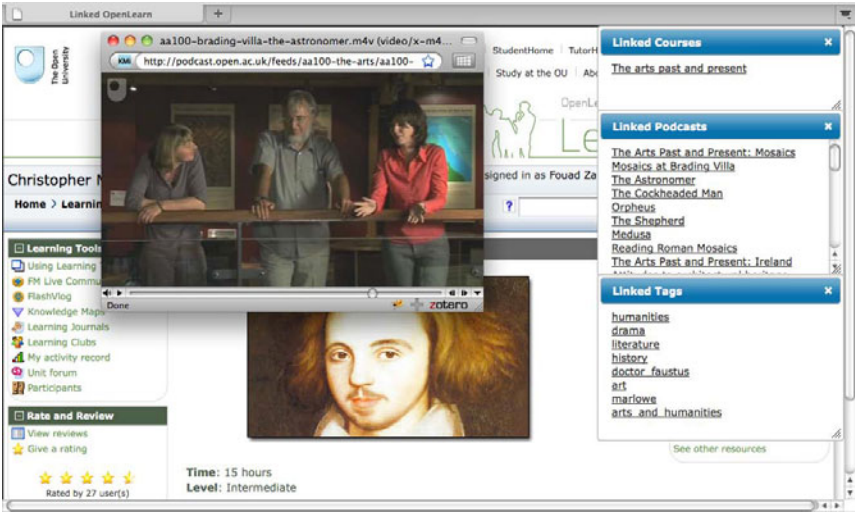

Fig. 4. Linked OpenLearn screenshot

7 Conclusions

In this section we report on our experiences when generating and exploiting LOD within the context of an educational institution. Regarding our experience on transforming information distributed in several OU repositories and exposing it as LOD, the process complexity was mainly dependent on the datasets in terms of type, structure and cleanliness. Initially, before any data transformation can be done, it was required to decide on the vocabulary to use. This is where the type of data to model plays a major role. With the goal to reuse, as much as possible, already existing ontologies, it was challenging to find the adequate ones for all our data. While some vocabularies are already available, for example to represent courses, it required more effort to model OU specific terminologies (e.g., at the qualifications level). To assure maximum interoperability, we chose to use multiple terminologies (when available) to represent the same entities. For example, courses are represented as modules from the AIISO ontology, and at the same time as courses from the Courseware ontology. Other factors that affected the transformation of the data are the structure and cleanliness of the

data sources. During the transformation process, we faced many cases where duplication, and information not abiding to the imposed data structure, hampered the transformation stage. However, this initiated the need to generate the data following well-defined patterns and standards, in order to get easily processable data to add to the LOD.

Regarding our experience exploiting the data, we have identified three main advantages of relying on the LOD platform within the context of education. Firstly the exposure of all these material as free Web resources provides open opportunities for the development of novel and interesting applications like the three presented in this paper. The second main advantage is the structure provided by the data. This is apparent in the OU Expert Search system, where the different properties of articles are exploited to generate different ranking criteria, which when combined, provide much stronger support when finding the appropriate expertise. Finally, the links generated across the different educational resources have provided a new dimension to the way users can access, browse and use the provided educational resources. A clear example of this is the exploitation of LOD technology within the OpenLearn system, where OpenLearn units are now linked to courses and Podcasts, allowing students to easily find through a single Web page relevant material that could be of interest.

We believe that universities need to evolve the way they expose knowledge, share content and engage with learners. We see LOD as an exciting opportunity that can be exploited within the education community, especially by interlinking people and educational resources within and across institutions. This interlinking of information will facilitate the learning and investigation process of students and research staff, enhancing the global productivity and satisfaction of the academic community. We hope that, in the near future, more researchers and developers will embrace LOD approach, by creating new applications and learning from previous experiences to expose more and more educational data in a way that is directly linkable and reusable.

8 Future Work

The application of Linked Data within the OU has opened multiple research paths. Regarding the production of Linked Data, in addition to transforming the library records to LOD, the LUCERO team is currently working on connecting the OU's Reading Experience Database (RED)[40] to the Web of data. Such database aims to provide access and information about reading experiences around the world. It helps the readership for books issued in new editions for new audiences in different countries to be tracked. Its publication as LOD is an interesting example about how the integration of Linked Data technology can open new investigation paths to different research areas, in this case humanities.

Regarding the consumption of LOD, we envision, on the one hand, to enhance the three previously mentioned applications and, on the other hand to generate new applications as soon as more information is available and interconnected.

[40] `http://www.open.ac.uk/Arts/reading`

As example of the former, for the Social Study application we plan to extend the current approach for identifying common terms between users' interests and courses information, to instead utilise common more generic concepts. At present the use of merely interest concepts from within Facebook may be too specific to suggest relevant course to users - in many cases users expressed interest in a particular niche band that could instead have been replaced with a concept describing the genre of music. By instead using concepts, we believe that the suggested courses would be more accurate and suitable for studying. As an example of the latter, we aim to generate a search application over the RED database, able to display search results on an interactive map and link them not just to relevant records within the RED database, but also with relevant objects of the LOD cloud.

References

1. Bizer, C.: The emerging web of linked data. IEEE Int. Systems, 87–92 (2009)
2. Bizer, C., Heath, T., Berners-Lee, T.: Linked data-the story so far. Int. J. Semantic Web Inf. Syst. 5(3), 1–22 (2009)
3. Heath, T., Bizer, C.: Linked Data: Evolving the Web into a Global Data Space (2011)
4. Hepp, M.: GoodRelations: an ontology for describing products and services offers on the web. Knowledge Engineering: Practice and Patterns, 329–346 (2008)
5. Kobilarov, G., Scott, T., Raimond, Y., Oliver, S., Sizemore, C., Smethurst, M., Bizer, C., Lee, R.: Media meets semantic web how the bbc uses dbpedia and linked data to make connections, pp. 723–737 (2009)
6. Rowe, M.: Data.dcs: Converting legacy data into linked data. In: Linked Data on the Web Workshop, WWW 2010 (2010)
7. Yimam-Seid, D., Kobsa, A.: Expert-finding systems for organizations: Problem and domain analysis and the DEMOIR approach. Journal of Organizational Computing and Electronic Commerce 13(1), 1–24 (2003)
8. Zhu, J., Huang, X., Song, D., Rüger, S.: Integrating multiple document features in language models for expert finding. Knowledge and Information Systems 23(1), 29–54 (2010)