

# Using DSpace as a closed research repository

Hind, Ianthe; Taylor, Robin

## Introduction

This year the University of Edinburgh introduced a dual DSpace<sup>1</sup> repository architecture: A closed repository to hold all research outputs (no full-text required and password protected) and an open access repository (full-text only). This presentation will focus on the closed repository.

In our repository we have our university hierarchy: colleges and schools as the communities and the academic staff at the collection level. This is not the conventional way of setting up a DSpace repository and immediately allows an author to be associated with a collection.

We used the item based submission from the 2009 Google Summer of Code Submission Enhancements<sup>2</sup>. This has allowed us to only display relevant metadata based on publication type during submission.

We have added functionality for an academic (on the collection level) to export a list of their publications to their own personal webpage by inserting some JavaScript (or as XML). This dynamically fetches the list of their publications from the repository at each load, with the most recent publications and orderings as defined by the academic reflected in the list. The same can be done by a research administrator (on the community level) to export a list of publications for their school or research group.

We have developed the ability to export items across to our open repository using SWORD – open access items can be copied across on deposit and those under embargo can be copied across once the embargo period has passed.

## Academics as Collections, Departments as Communities

Our repository follows our official university hierarchy. The colleges and schools are our communities and the academics are our collections. Each community and collection has an extra field in the database which is a university based identifier. Each collection has an associated eperson which is linked via the identifier (using netid for the eperson).

When a user logs into our repository, via our university's single-sign on, they are taken directly to their collection. This set up allows us to have our publication's ordered by academic easily without the need for a naming authority to match on author names.

This unconventional setup has been done with the Research Excellence Framework<sup>3</sup> in mind as it allows the data to be queried easily from a management perspective:

## Item-Type Based Submission

Using the code developed during the 2009 Google Summer of Code Submissions Enhancements, we have item-type based submission within our repository. The first screen shown during submission asks for a publication type to be selected. Based on the choice made, the relevant metadata for that publication type is displayed on the next submission page.

The configuration required to setup up the publication types and their associated metadata can be done via the UI without `input-forms.xml` and `item-submission.xml`

---

<sup>1</sup> <http://www.dspace.org/>

<sup>2</sup> <http://www.fedora-commons.org/confluence/display/DSPACE/Google+Summer+of+Code+2009+Submission+Enhancements>

<sup>3</sup> <http://www.hefce.ac.uk/research/ref/>

having to be edited directly (as was the case previously). The forms are displayed according to the item-type rather than the collection handle (traditional way).

### **Publications List Aspect**

We are using the DSpace XML user interface (Manakin<sup>4</sup>) for our repository and created a publications list aspect for this functionality. There are three main parts to this, the ranking, which allows a ranking to be placed on an items in a community or collection; the ordering, which allows the user to chose the order in which they would like the items displayed in the publications list and then the publications list itself.

### **Ranking**

We have added the ability for a user (academic or research administrator) to assign rankings to their publications. The publications can then be ordered in their publications list according to this ranking. When ordering by ranking, items are ordered in decreasing ranking order so those with a higher ranking score are shown first followed by those with a lower ranking score. Publications assigned a ranking of 0 are not shown in the publications list which allows for confidential items to be excluded from the publications list.

It is worth noting, that ranking is per community or collection per item. So if an academic assigns a ranking to a publication, the ranking assigned to that publication by a research admin may be different (where the collection is a descendant of the community).

### **Ordering**

When the option to export a publication list is selected, the user is presented with the following ordering options:

- By ranking – items are listed in descending ranking order
- By publication type and then ranking – items are listed under the headings of the publication types included and then ordered in descending ranking order
- By publication type and then date descending – items are listed under the headings of the publication types included and then ordered in descending date order i.e. with the most recent publications first.
- By publication type and then date ascending – items are listed under the headings of the publications types included and then ordered in ascending date order i.e. with the oldest publications first.
- By date descending – items are listed with the most recent publications first.
- By date ascending – items are listed with the oldest publications first.

There is an option to choose how many publications to display – the default is for all publications to be included but allows for the functionality of displaying the top three publications for an academic for example.

### **Publications List**

Once the ordering has been selected, there is an option to choose display format, either embeddable HTML or XML. For the embeddable HTML, the code to be pasted into the web page where the publications list is to appear is displayed. The code is a snippet of JavaScript that uses JQuery to query the repository and format the results in a suitable format. How the publications list will appear on their web page is shown below the code. For the XML option, we have created a basic XML structure based loosely on the BibTeX<sup>5</sup> terminology. This option will be used by those wanting to manipulate the data and display it how they choose. Both of these options will update dynamically based on the repository content and rankings set for the publications.

---

<sup>4</sup> <http://di.tamu.edu/projects/xmlui/manakin/>

<sup>5</sup> [http:// www.bibtex.org/](http://www.bibtex.org/)

As it stands, the publications in the publications list are listed in the Harvard citation style. It is intended that the user will be able to select which citation style they would like their publications list displayed. We are currently investigating using Citeproc<sup>6</sup> to implement this solution.

We have provided a way for data to be exported from our dark publications repository in a form that is useful for academics and schools. The more ways in which the academic and research administrators can engage with the publications repository the better, and the publications list provides a way for the data that they have spent time depositing to be exported in a way that is useful to them.

### **SWORD between two DSpace Repositories**

Direct deposits will be made into our closed repository. The user is presented with access options during submission where open access items can be marked for transfer to our open repository. This can also be done after an embargo period has expired. We chose to use the SWORD capabilities of DSpace for this transfer from our closed repository to our open repository. The open repository already had SWORD ingestion enabled (being DSpace) so the work was in creating the ability to create the SWORD ingestion package for the open repository from our closed repository, selecting the appropriate collection in the open repository to deposit into and passing the open repository handle back after deposit to be stored in the closed repository.

### **Presentation Summary**

In the presentation, we will set the scene for our closed publications repository and explain the motivations for the extensions we made to the DSpace repository framework. We will then give a demonstration of our closed repository showing each of the extensions in action.

### **Conclusion**

We have used DSpace for a closed repository successfully with extensions to suit our needs:

- The communities and collections are based on our university structure with the colleges, schools and departments being our communities and the academics being our collections.
- Items are submitted based item-type so only relevant metadata is displayed on the submissions forms which can be configured directly by an administrator through the UI.
- A publications list can be created on either a community or collection level displaying all the items in that group either as embeddable HTML or as an XML file.
- Items can be transferred from our closed repository to our open repository using SWORD with the ability for this to happen after an embargo period.

---

<sup>6</sup> <http://xbiblio.sourceforge.net/citeproc/>