Cataloguing, Classification and the University of Northampton Institutional Repository

Janet Milner, Metadata Specialist, Library and Learning Services, University of Northampton

Introduction to NECTAR

The project to set up the The University of Northampton's open access institutional repository was begun in 2007. Prompted by the RAE (Research Assessment Exercise), with its demands for comprehensive research output data and evidence of support for research activity, members of both the research community and the Library identified an opportunity to collaborate on a new digital repository. It remains a joint endeavour between Library and Learning Services, and the research community at the university – it is not just a 'library thing'.

Two groups were formed to further the project: a Steering Group, comprising members of both the research community and our department; and a Project Team, of which I was a part, and which was responsible for the day to day management and operation of the project.

The repository was soon christened NECTAR (Northampton Electronic Collection of Theses and Research). It was designed to capture, showcase and preserve the University's scholarly research output, including PhD and MPhil, but not undergraduate, theses. NECTAR's theses can now be harvested by the British Library's thesis service, EThOS.

Eprints software was selected, specifically the 'Professional package' offered by Eprints Services. This software is based on the Dublin Core metadata standard. It also provides for enhancing the records with Library of Congress classification.

Initially this was a mediated service – the School admin teams (usually the School Secretary) were trained to enter the metadata, the items being visible only to NECTAR Project Team staff. Since 2011, academics have been able to self-archive. They are now much more engaged with the process, and aware of the value-added services NECTAR provides for them.

The University Annual Research Report is now generated directly from NECTAR.

Although the start-up project is over, the Team still exists, as there is ongoing development work on NECTAR dealing with preservation issues, copyright concerns, the addition of full text, and of multimedia items, and further researcher engagement. NECTAR will shortly be enabled to assemble and submit the necessary publication information to the Research Excellence Framework (REF), for the University.

Current Repository Team

1. *Research Support Librarian*, who heads up all Research support for our Library and Learning Services

2. *Research Information Specialist*, whose special responsibility is to administer NECTAR, including liaison with content providers; checks and enhances metadata; verifies intellectual property rights; prepares full content for deposit; generates reports and statistics

3. *Information Systems Developer*, who deals with all the technical issues around NECTAR

4. *Metadata Specialist*, who adds some new entries, checks and enhances metadata as part of quality control, adds Library of Congress subject classifications, checks presentation, makes entries 'live'

5. *Cataloguing Assistant*, who has also been trained to add subject classifications, and verify metadata.

My Part

Metadata

My job title is Metadata Specialist, and I was fortunate in that it was felt useful that someone with metadata skills be involved as part of the Project Team from the beginning. Indeed, one of NECTAR's selling points to the University is the high quality of its metadata.

I was involved in the selection of the software, then the subsequent organisation of the metadata fields within that to enable the entries to be presented in the Harvard Referencing style format. I was also involved in the design of the fields within the template, so that the relevant metadata fields were available for each type of item. The fields available for a journal article, for example, won't necessarily be the same as those for a radio/ TV interview.

I attended a Repository Support Project event that was held about this time, on metadata for repositories. So I was able not only to brush up on my knowledge of the Dublin Core elements, but also to actually work with them for our repository project.

Currently I spend approximately one day each week on NECTAR work, although this increases at busy times.

Issues

- **Accuracy of metadata** (dates, location etc.). The potential for (albeit unintended) inaccurate metadata is great. Checking is paramount.

- **School of Arts entries** – pictures, artists' exhibitions, artefacts, artists' books, don't all fit into the metadata fields template very easily. This is resolved by discussion with the Head of the School of Arts, and with the rest of the Team, to achieve the best presentation of these more problematic entries.

- **Workload priorities**. Entries appear in the Review area in a sporadic manner, so sometimes I am swamped by a huge number waiting to be checked and made live, particularly when the Annual Research Report deadline is coming. This then impacts on my workload, causing time management issues.

Cataloguing

Cataloguing skills are a great asset for working with a repository. The same skills required for cataloguing are required for entering or checking metadata in a repository – attention to detail, accuracy, speed, consistency, the following of set rules and standards, verifying of data, authority control, punctuation. Whereas normally the cataloguer is entering data into fields in the MARC21 record format, to create an end record following AACR2 or RDA standards, in the repository we are using the elements in the Dublin Core format, to create a citation (or record) following the Harvard Referencing format.

Issues

- **Authority control**. Different people were entering different forms of name for creator, resulting in multiple entries in the name browse list for the same person. This was solved by our technical person in conjunction with Eprints – these are now auto-completed as soon as typing begins.   However, occasionally the auto-complete fails (either because for a spell it seems to get switched off or because folk type too fast!) then duplicate names still end up on the browse list. Periodically this is checked for duplicate names and fixed manually. I keep my beady eye on the creator names to pre-empt this. The same facility is there for item titles and conference titles. This not only picks up duplicate entries, but at the same time helps with

consistency in these fields.

- **Consistency**. I produced a set of guidelines covering, for example, the use of capital letters for different types of heading; what to do with a recurring honorary role (treated rather like a journal catalogue record); the same exhibition or paper given in different locations in the same year – this is entered on one record that includes them all. I have found my cataloguing background incredibly useful in sorting out these types of issues, or discussing them with and advising the Project Team. This crib sheet has proved invaluable whenever a new person comes on board to help with NECTAR, although it is now covered in scribbled annotations and needs a new edition!

I also contributed to the online field 'help' information. This is available to the user at the point of entry in each field and so useful to researchers as well as repository staff. It helps in the effort to achieve consistency in metadata presentation.
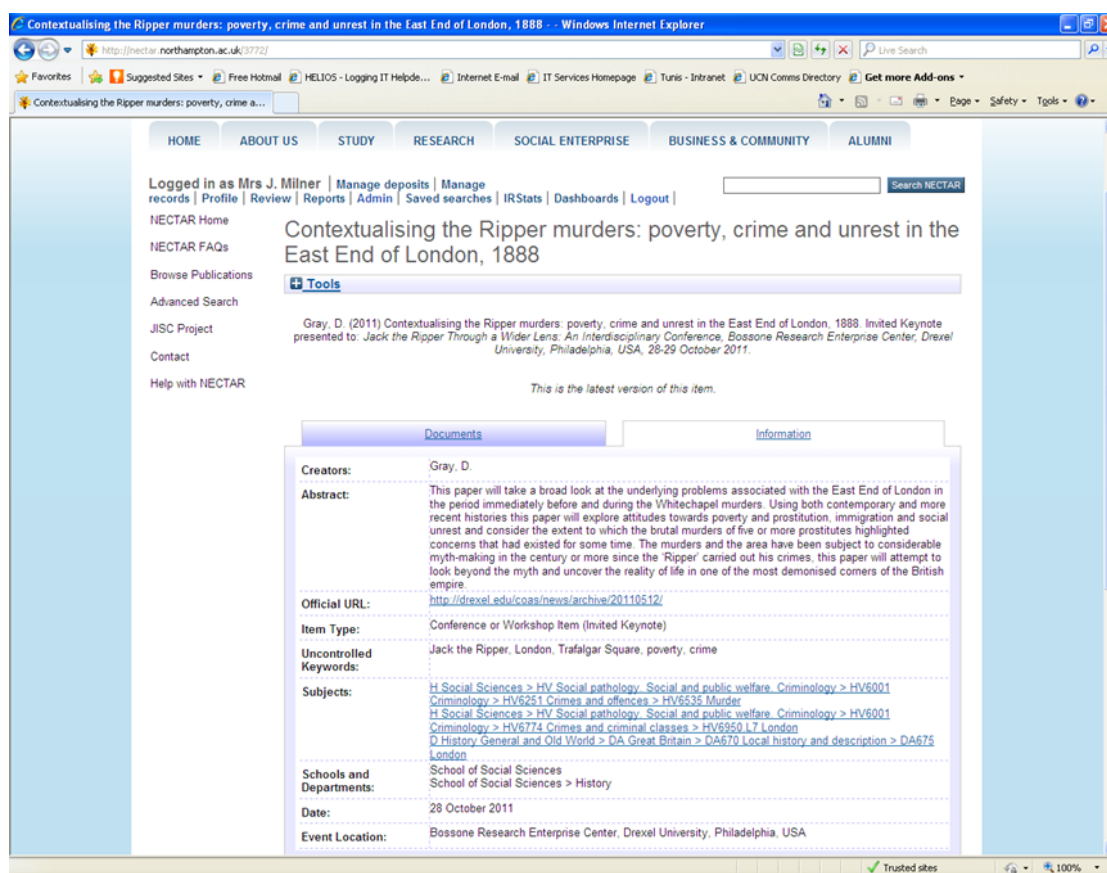
Classification

It was felt by the Research Support Librarian that adding subject classifications to the records would add value to them, in addition to the uncontrolled keywords that in some cases were not provided. Adding the Library of Congress classifications provides some vocabulary control and consistency over the subjects.

My knowledge and experience of the principles of subject classification (although not of Library of Congress!) were very useful. Also, having classified library stock at the University for some years, I had already gained some subject knowledge of the research areas being covered as the academics tend to research around the subjects they teach.

Issues

- **Classification scheme provided 'out of the box' too broad**. Eprints provides the basic Library of Congress classification outline, but we quickly found that these classes were too broad, and decided to add to the subject tree in NECTAR, the more relevant ones from the sub-classes, which are freely available on the LOC website. We have continued to add more as different subjects are covered. This took up extra time at the outset, less so now, as we already have most of those we use regularly. Up to three subject classes per NECTAR item entry are added.

Conclusions

Being part of the NECTAR team has given me the opportunity to use my cataloguing skills in a different area of work. I have enjoyed learning something different, and developing the skills I already had. I was forced into learning more about LOC classification, and this has proved useful in my normal cataloguing work – I now know what these are when I see them on catalogue records! What is exciting is that the institutional repository arena is continuing to move forward, and I am pleased to continue to be part of it.

References

Pickton, M. (2008) Gathering NECTAR at the University of Northampton. ALISS Quarterly. **3**(4), pp. 33-38. 1747-9258.

Pickton, M. (2012) Bringing a buzz to NECTAR. ALISS Quarterly. **7**(3), pp. 24-29. 1747-9258.