

## **A technology overview of a complex migration from Eprints 2 to Eprints 3**

E-LIS is an open access archive for scientific or technical documents, published or unpublished, on Librarianship, Information Science and Technology, and related areas. E-LIS relies on the voluntary work of individuals from a wide range of backgrounds and is non-commercial. It is not a funded project of an organization. It is community-owned and community-driven. We serve LIS researchers by facilitating their self-archiving, ensuring the long-term preservation of their documents and by providing world-wide easy access to their papers. It has been recently ranked third in the Directory of World Repositories: [http://www.webometrics.info/top200\\_rep.asp](http://www.webometrics.info/top200_rep.asp)

In this poster we list and briefly explain the challenges of a complex migration from Eprints 2 to Eprints 3. The Eprints 2 installation of E-LIS is quite complex and personalized. With this migration we want to transfer all specific features into the new version.

The migration is still ongoing and we have found those specific aspects:

### **-Purge inbox buffer**

E-LIS is an archive with many subscribers, around 3500, and many users have problems on learning how to insert a paper into the archive. This situation produces, as a result, that we have around 1400 papers inserted for test. Before the migration we need to purge the section of the archive that hosts those tests, the inbox buffer in Eprints software jargon.

We send an email to the users and we ask them if we can delete those papers that look as a test.

### **-Changes on metadata**

With the migration toolkit of Eprints 3 we transport the metadata configuration of version 2 to version 3. But we want also to improve metadata configuration; we add a new field (not mandatory, repeatable, with info about language) to handle an abstract in languages different from English. We work on setup and we modify the metadata exported from Eprints 2.

We find that we can modify in a perfect way metadata with a program, so we develop a program that modifies metadata but also logs the situations with ambiguities. E-LIS editors will recheck the situation logged by the program.

### **-UTF-8**

In the past E-LIS had problems on users input. Before the migration we check if the characters are UTF-8.

### **-Browsing of fields with long values and non-Latin-1 chars .**

One of the most important hacks of E-LIS is the modification done to better support the browse indexes.

In standard Eprints 2 and 3 fields with non-Latin-1 chars have severe limits on the numbers of possible chars if the field are used for a browsing index. With this hack this limit is improved.

### **-Specific modifications for editors**

E-LIS editors asked for modifications about checking and approving new papers. Into Eprints 3 we recreate those modifications using the new options about workflow and with new hacks into the cgi. In Eprints 2 the work was done only with modifications of cgi.

### **-Alerts improved**

In E-LIS we added abstracts into the alerts and we inserted an alert to the uploader if a paper is approved. We operate to have those functions also in our setup of Eprints 3.

#### -Statistics

E-LIS offers statistics about the usage of a paper with a software developed ad hoc, starting from the code of University of Tasmania. In Eprints 3 we think to use the new module developed by the University of Southampton. We need to transfer data registered in Eprints 2 into the new system. Statistics of the past years need to be maintained.

#### -Scripts developed for particular tasks

For many administrative tasks we developed perl scripts that work directly with MySQL tables, without using the Eprints 2 API. Now Eprints 3 API is better documented, so we rewrite the same scripts using it. But in some cases we still use direct calls to the database.

#### -Setup of static pages

Some static pages in Eprints 3 have the same appearance but also new features for the users, like RSS and the import options.

#### -Paracite

We transfer to Eprints 3 a version of Paracite improved by us, for better handling of URLs in the references.