Fedora UK and Ireland User Group (2010-12-13)

The Fedora UK and Ireland User Group met at the London School of Economics and Political Science (the LSE) on Monday 13th December with some 25 people in attendance.

**Short talks**

The first part of the day was given over to short talks and presentations on matters of potential interest to attendees.

**Update on Fedora (Steve Bayliss, Acuity Unlimited, Fedora 3.4 release manager)**

Presentation

**LSE's plans for Fedora (Simon McLeish, LSE)**

Presentation

**Hydra, BL@H and CLIF (Richard Green, University of Hull)**

Presentation

**Video storage at the Open University (James Alexander, OU)**

(Presentation requested)

**Modelling the repository at York (Julie Allinson, York University)**

Julie did not give a screen-based presentation but rather talked us through York’s approach to working with cModels which uses a four-part structure. She made references to a document which has now been made available on their Digital Library wiki.

**Datastore development on Fedora principles at Oxford (Anusha Ranganathan, Oxford University)**

Presentation

**The Kindura Project and DuraCloud (Simon Waddington, King's College London)**

Presentation

**Open discussions**

The later part of the meeting was given over to short, timed to 15 minutes, discussions on a range of subjects put forward by people at the meeting:

**Preservation**

- What is considered good practice in Fedora?
- Some have embedded PREMIS in METS. Hydra is using PREMIS, but avoiding METS as too complex. Oxford is using MODS and avoiding PREMIS.
- In relation to work on the AIMS project, there was a query about moving metadata from Archivematica into Fedora. It was noted that Archivematica omitted some core PREMIS fields.
- Preservation services - what is needed? It was noted that we are far more aware of what services can be delivered, but that implementing them hasn't moved on much in the last 10 years.
- From an institutional perspective, a repository tends to focus on enabling preservation, even if it isn't doing it. Not all institutions are capable of carrying out preservation.
- The management of disk images is a current area of interest.
- There is also the matter of the right to preserve.

**Community involvement**

- The main benefit of having our UK&I community was considered to be finding out who is working on similar issues, and being able to make contact to discuss them further.
- The idea of a separate fixed website point of access to information about the group was proposed and accepted. This is not intended to replace the information on the DuraSpace wiki, but simply provide an easy point of reference, drawing together information held elsewhere.
- Such a website could include: a blog roll, meeting information, a list of members.
- Meetings would always be open to Ireland and other EU members if they wished to attend.
- We need to ensure that the FEDORA-UKI mailing list maintains its focus on meetings, and keep discussion of issues to the main Fedora mailing lists.
Migration from Fedora 2.2.4 to 3.x

- National Library of Wales had recently carried out this migration.
- Those going through a migration were warned to watch what happens to checksums.
- The data migration tool provided by Fedora is very useful for migration up to 3.3, but not 3.4 (as DuraSpace thought all would have migrated by this stage).
- Evidence suggested that there are many who are still in the process of migrating: York used a tweaked version of the migration tool to get to 3.4.1. Don't expect the migration tools to be maintained, although it is recognised that there is a remaining issue with those wishing to switch to 3.4.1 because of the recent security fix.
- The migration tool expects certain files to be in certain places on the server.
- An alternative approach is being undertaken by Hull and Oxford, exporting, converting and then re-ingesting to the new version.
- Versions of associated software need attention on installing new versions of Fedora.
- It is hoped to move to a WAR file deployment in the future.

Fedora high level storage

- Low level storage, as currently implemented by Fedora, just stores blobs.
- A high level storage API can be cleverer about how objects are stored and enable proactive storage decisions.
- Aaron Birkland at Cornell is working on this, in part through the Data Conservancy project.
- The approach being explored by DuraSpace is still in flux, and input is welcome.
- High level storage would provide greater flexibility to deal with managed and external content, and also be better for preservation, as all information would be available for re-creation.
- Enhancements in Fedora 3.5 will lay the groundwork for taking this further.

Fedora cModels

- Spoken Word at Glasgow Caledonian have developed a video content model.
- Content models inheriting other content models would be useful.
- Content model aggregation? This is feasible, but only on a per object basis. York have done this, and will write up the use case.
- Could there be a content model pattern that objects can subscribe to? Can Enhanced Content Modelling (ECM) enable this? ECM features are additive.
- ECM is being used in Denmark (where it was developed) and at FIZ Karlsruhe in Germany (as part of eSciDoc).
- It was noted that within ECM datastream validation works on all XML datastreams.
- There is a fear of setting things in stone.
- It should be remembered that a repository is structural content modelling overall.

FESL and security

- The Fedora committers are keen to hear more about use cases for security and access control. The Open University offered itself as an example of multiple use cases.
- Many hide Fedora behind other systems: gated discovery. For example, Hydrangea uses security in the Solr index. Whether to adopt this approach or not is key to the repository architecture, and will be guided about what the repository is for.
- In Hydra all objects have a rights datastream, so all atomistic parts can be re-used if required.
- For such rights, the concept of human and machine-readable versions is useful.
- Another approach may be to associate licence objects to content. This can work well, but may have scaleability issues depending on the number of licences used.

Other contributions

Eliza Stefanova, who is currently based in London and attended the meeting offered us this presentation to share with the group describing the use of Fedora in the ShareTEC portal developed by a colleague of hers at Sofia University, Bulgaria.

Spoken Word Services, at Glasgow Caledonian University, were unable to attend but offered this summary of their recent work.