2.0discussion

Meeting Notes

At the Open Repositories meeting, April 4, 2008 the committers met to talk about the roadmap and plan for the DSpace platform. In particular, each person there was asked what they were working on, and what work had been completed to date on 1.6. These are the notes from the meeting, please comment and edit as you see fit. For each participant I have outlined their input. No decisions were made, just started the process for getting everyones input and thoughts out to start working through the possible timeline.

- Richard Jones- Has a working module for abstracting the identifiers in 1.6. This work has been integrated into the JSP layer but needs to be incorporated into the manakin layer.
- Tim Donahue-Tim would like to make sure an updated statistics package is available in 1.6 or the next release. Th group felt the work done by Minho University and/or the GoSC interns from last year would be good candidates. He pointed out that currently in the 1.5 release the only statistics now available or in the JSP interface, not Manakin. Scott Yeadon also commented that ASPR was working on a general IR stats package that should be deployable on all IR software for evaluation the end of June 2008.
- Richard Rogers-Richard emphasized that moving towards "pluggable" design is important, and that we could support/deploy several IR stats packages with a pluggable design so the user could choose. That being said, he felt we need to at least have one available with the next release. He emphasized that whatever we choose to incorporate it should be pulled out as a module, to follow our current Maven practice of design and build. Richard has been working on the asset store abstraction layer that would allow users to decouple the asset store from the application and use 3rd party hardware/software platforms, such as Sun Honeycomb, Amazon, SRB and possible others. He also noted that rework of the bitstream format registry was going on(mainly by Larry Stone) and this code would be available so users could validate their formats against 3rd party sources such as GDFR, Pronom and the like.
- Jim Rutherford-Jim has been working hard on the core of 1.6, which has abstracted all of the database dependencies from the data model. This way users can choose not only Oracle and Postgres, but potentially other databases. In last years GSoC work was done to incorporate versioning in 1.6 and a full suite of unit tests has been developed by Jim for testing the new database abstraction.
- Scott Phillips- Scott would like to see a 1.5.1 release out very soon as there are 2 critical bug fixes required for 1.5- SWORD and Oracle bug fixes. this should be purely a bug fix release. Scott is working on developing an ETD management frontend called Verio. He felt this would not be ready until 2009 after 2.0 release. He is also looking at incorporating themes into a web administrator interface- so no coding is needed to change a limited number of themes in manakin. He emphasized he can only be available for community coding 25% of his time. Therefore- until others in the community or up to speed on Manakin coding and integration, this may be a limiting factor in what can be achieved if it needs to be integrated with Manakin.
- Scott Yeadon- Scott will be leaving us to work on another project, which does not involve DSpace this summer. He hopes that ANU will continue their involvement.
- Jim+Downing-???
- Rob Tansley- Rob theoretically, has 20% of his time to work on DSpace for the community. He commented we should publish a "best practices" guide for the community to let them know how they can minimize work required to migrate customizations as we release new code. He would also like to incorporate a Google web tool kit in the next release.
- Graham Triggs- Graham is working on a citation crosswalk. He felt we should figure out quickly the standard interfaces to the new data model and publish those so other developers could build on those interfaces.

In addition to these folks, myself (Michele Kimpton), Brad McLean (new foundation technical director) and Aaron Zeckoski (new developer for 2.0) were present. The meeting lasted 1 hour. It was agreed that to move forward we would have a, possibly, weekly conference call, and the Foundation would orchestrate this.

JISC funded work

JISC has graciously funded part of the work for the 2.0 project. Attached is the proposal http://wiki.dspace.org/index.php/Image:DSpacejiscproject.doc.pdf] submitted to JISC.

Post Meeting Notes and Comments

- Mark Diggory: (not in attendance) these are not strictly "my projects" but areas of work I hope to assist in driving forward through collaborative development in the community and GSoC projects.
  - Statistics Packages: I feel a Statistics Package for DSpace+1.6 is something that is a a side project and shouldn't require modification to DSpace directly to maintain. We need to start a coordinated project in the sandbox that may be used as a fulcrum point for consolidating statistics alternatives and coming up with a common solution, MIT as well is heading down a very short road to implementing a statistics solution for our reporting needs. It should be extensible and require little modification to the core codebase.
  - Improving Metadata associated with Containers and Items: I have a serious interest in seeing the Metadata management capabilities of Dspace enhanced for 1.6/2.0 and believe we can utilize certain projects this summer to prototype solutions that will open our metadata up to be more flexible, some of these may be classified under Semantic Web umbrellas, but all in all, actually do not alter the underlying infrastructure for managing DSpace metadata and simply redefine how the the tables are used and what objects actual metadata can be assigned to.
  - Crosswalk refactoring: The Manakin XML-UI work showed us that our current implementation of Crosswalks is not only inefficient, but overly complex. I hope to see likewise, with the project above, the possibility of taking all the work done in the Manakin Space involving DSpaceObject adapting and SAX pipelining, to address an alternative Crosswalk solution that will be more efficient and scalable.