

Open Repositories 2010 Call for Papers, Posters and Workshops
“The Grand Integration Challenge”
DSpace User Session Proposal
Mark Diggory, Head of U.S. Operations, @mire NV

**DSpace Discovery:
Unifying DSpace Search and Browse with Solr**

One key innovation long awaited by the DSpace community is a more intuitive and unified search and browse experience. NESCent¹ and @mire NV² have collaborated to create a new Faceted Search and Browse experience for NESCent’s DSpace repository, Dryad³. DSpace Discovery is a modular Add-on for DSpace XMLUI that replaces DSpace search and browse with Solr. The implementation of Discovery’s Services utilize the DSpace Services API originally developed for DSpace 2.0 and back-ported for use within the recent release of DSpace 1.6.0. Thus, DSpace Discovery represents the next stage in @mire’s DSpace 2.0 development initiative.

DSpace search has always been Lucene based, much like Apache Solr. However, its browse capabilities pre-dated new faceted search technologies like Solr. The DSpace community historically has chosen to maintain its own custom solutions to browsing even in the face of such new technologies. The significant amount of time and resources devoted to the recent re-architecture of the Browse system in DSpace 1.5 has shown us that choosing local maintained custom solutions will have a tendency to increase the burden of maintenance. It can be further asserted that the cost of such custom solutions escalates as core developers move on to new projects and interests. These costs ultimately limit the communities ability to be innovative in the Institutional Repository niche’.

In such situations, the appropriate strategy to avoid the perceived future deadlock in development is to abandon local solutions and seek to utilize other popular Open Source tools that will more appropriately meet the needs of the community. This is most effective when such tools already have a diverse healthy community that is comprised of many different commercial and non-commercial interests. These “Best of Breed” solutions should always be considered over local homegrown tooling.

¹ <http://www.nescent.org>

² <http://atmire.com/>

³ <http://datadryad.org>

Apache Solr is a popular, scalable, distributed open source enterprise search web-application built upon Apache Lucene. Solr has many features including full-text search, hit highlighting, faceted search, dynamic clustering, database integration, and rich document indexing. Many of the world's largest Internet sites utilize Solr to provide rich faceted navigation. Likewise, Solr has become popular within the academic open source community with the advent of projects such as Blacklight and the Hatha Trust, both utilizing the tool for their search systems. Within the DuraSpace community, Fedora's GSearch tool supports Solr and recent projects such as Hydra have shown the power of its solution. As such, Solr clearly represents one point of commonality in integrating DSpace and Fedora.

While already mandated to be Open Source by Dryads funding obligations, @mire and NESCent are additionally at work to release DSpace Discovery under the DuraSpace license and maintained in parallel to DSpace by the community within the DSpace source-code management services⁴. @mire's goal in this approach is to exemplify how Add-on modules for DSpace community should be properly architected and maintained. NESCent and @mire are inviting participation in the DSpace Discovery project with the intent to get more community developers engaged in its evolution.

In summary; by leaving behind older in-house search and browse solutions and adopting a tool with a vibrant expert developer base, the DSpace community will not only be free from the maintenance burden of maintaining its own custom browse infrastructure, but will also gain all the benefits of many experienced professional contributors from the enterprise search sector. In doing so, DSpace community developers will be more able to focus more exclusively on integration with Fedora and enhancing DSpaces curatorial and scholarly publishing capabilities.

⁴ <http://scm.dspace.org/svn/repo/modules/dspace-discovery>