

Open Repositories 2013

Proposal For: Main Conference Panel

Title: Migrating Repository Islands

**Presenters: Mark Leggott, University of Prince Edward Island
Richard Shrake, American Philosophical Society
Mitch MacKenzie, discoverygarden inc.**

Software, like any technology application, seems to become obsolete as soon as it is implemented. The need to change can come from simple bug fixes and enhancements to significant rewrites that may require changes to data and lengthy migrations. When combined with the need for staff to accommodate changes in their local business practices, changing software can lead to major disruption in an organization. While not as complex as an ERP system, repositories can present challenges when migrations are desirable or necessary. This is especially true when it comes to metadata, which is the “bottom of the iceberg” in many migrations, the unseen challenges that come from the need to effect data normalization and data mappings from one schema to another and more. This panel will provide details of recent migrations between versions of the same repository system, and between different systems. The opportunities and challenges of data migration will be highlighted.

Repository systems by definition need to ensure that system changes and migrations do not impact the integrity of the stored assets and metadata. The Islandora system provides an exemplary approach to repository development, maintaining a clear separation between interface functionality and the repository of data and metadata. By storing metadata and assets in the Fedora filesystem, unencumbered by additional layers of technology, the ability to migrate core repository content is optimized regardless of future requirements. Metadata is packaged in a well-defined, METS-like XML structure (FoXML), including descriptive, technical, administrative and preservation metadata. The panelists will consider 4 approaches to repository migration.

Migrating Between Software Versions - The latest Islandora release provides an excellent example of a system that insulates the user from lengthy and complex migrations. Islandora is able to run both Drupal 6 and 7 front-ends on the same Fedora repository, providing a powerful example of the ability to leverage new functionality in the Islandora 7 release, while maintaining previous Islandora 6 collections. This in itself is a unique aspect of the Islandora architecture, allowing institutions to maintain previous efforts while building on the new. The system also facilitates the complete migration from earlier Drupal releases to the latest stable release, a migration that is typically much more complex in the absence of the Fedora/Islandora data layer.

Migration Between Repository Systems: In-House Model - The American Philosophical Society migrated from a ContentDM system to Islandora 6, developing its own tools and processes. This

migration involved approximately 10,000 digital objects (images and audio) and several iterations of customized, non-standard metadata. In seeking to not only migrate the objects and metadata from one system to another but to also correct errors in description and initial design, the APS model consisted of a combination of batch uploading strategies and manual data entry. A corollary and requisite task involved the updating of every reference to the objects from access tools such as the Archivists' Toolkit and the primary APS website.

Migration Between Systems: Service Model - discoverygarden Inc. has provided migration services from DigiTools, ContentDM, DSpace, Philologic and a number of proprietary systems such as SharePoint. Even when systems support standards such as Dublin Core, the challenge of migrating metadata to another system can be significant, especially when local customizations have been made. The migration process requires extensive stakeholder coordination in major phases, including: content cataloguing and inventory discovery, content mapping, transformation, and enhancement, migration execution, and quality assurance. The management of these process steps and the tools used will be discussed. Steps of the migration process which are amenable to automation will also be highlighted.

System Integration/Synchronization - Synchronizing data between systems, rather than effecting a complete migration is another option, especially when data in a repository can provide information to multiple interfaces. The new Islandora Bridge module will provide a synchronization between Drupal and Fedora, allowing Drupal developers to leverage the complete suite of standard Drupal modules, while ensuring effective data preservation and future migration with the Fedora layer. The University of PEI has developed a flexible approach to their new institutional repository, including harvesting data from the enterprise Research Management system and the option of synchronizing data from the Fedora repository to other campus Drupal installations.