Digital Repository Requirements

These requirements were used to evaluate Digital Repository Architecture Alternatives. Now that Fedora has been selected, this list is mostly of historical importance.

User requirements

1. DLP staff can easily create new collections.
2. DLP staff can easily manage existing collections.
3. Faculty members, librarians, and others outside the DLP can easily create collections using DLP-supported resources.
4. Other digital library professionals can easily index, search, and obtain metadata in DLP collections.

Requirements the final repository must accomodate:

1. [**Must**] The repository is able to store/search/deliver both text-centric and image-centric collections.
2. [**Want**] Audio and video objects can be stored/delivered.
3. [**Want**] Arbitrary document types (PowerPoint files, music notation files, etc.) can be stored/delivered.
4. [**Must**] All metadata stored in the repository conforms to (or is easily exportable to) a widely-recognized standard format.
5. [**Want**] A wide variety of standard metadata formats can be stored, including MARCXML, DC, METS, and VRA Core.
6. [**Want**] Metadata about a single object can be transformed into multiple formats.
7. [**May want**] "Exceptions" to metadata transformations can be stored, either as full records in an alternate format, or as single fields that override the transformed version.
8. [**Must**] It is possible to add custom fields to existing metadata schemas.
9. [**Must**] It is possible to use completely arbitrary metadata schemas.
10. [**Must**] Preservation is addressed, preferably in a way that is compatible with OAIS, preferably checking file integrity using checksums, and preferably using an automatic connection to HPSS.
11. [**High Want**] The repository system is based on open-source software, so we don't have to depend on someone else to add features we need.
12. [**Must**] The repository is able to manage data at many levels of granularity.
13. [**Must**] Access privileges to materials in the repository can be managed at many levels of granularity.
14. [**Must**] There is a facility for searching across all collections held by the repository.
15. [**Must**] External search systems can be easily connected to the repository, in order to deal with specialized search fields.
16. [**Must**] Metadata about multiple versions of files can be maintained, including master and delivery versions.
17. [**Want**] Repository materials (both media and metadata) can be easily accessed by course management systems, including Oncourse CL.
18. [**May Want**] "Certification" according to the standards described in the October 2005 RLG DigiNews

Open questions

1. Will we include collection-level information in the repository? This includes things like the web sites that host collections, and the timeline of Cushman's life.