# 07 Digital repositories

## Introduction

Repository software aims to provide a managed environment for digital objects, such as documents and images, and their metadata. Repository software will generally include tools which allow curators and users to exploit the stored objects and their metadata. Paradigm evaluated DSpace and Fedora, two of many repository systems, with the objective of assessing their potential as preservation repositories for personal digital archives and to select software that the project would use in its prototype system. The online version of the Workbook contains some how-to guides for installing DSpace and Fedora and on ingesting file directories into Fedora.[1]

Paradigm distinguished between access and preservation repositories very early in the life of the project. Repositories catering for less sensitive materials with shorter embargoes, or materials where rightsholders permissions to publish can feasibly be sought, often combine preservation and access functions in a single repository. Paradigm's work suggests that this is not the optimal solution for personal digital archives because:

- Preservation repositories have different functional requirements to access/presentation repositories.
- Preservation and access have different metadata requirements.
- Preservation repositories do not have the same infrastructure/performance requirements as presentation repositories.
- Access repositories must be networked, but closed materials are afforded better security when placed in a preservation repository that is isolated from the network.
- Preservation repositories which are networked will require much thought to be given to security of materials server-side and client-side, and in transit.
- Preservation repositories may have different backup and disaster recovery requirements.
- The management of preservation repositories calls for different skills and experience.
- Preservation repositories have different users (archivists) than access repositories (researchers, general public).
- Some items held in preservation repositories may never be published to online access repositories (e.g. software maintained by the repository for data extraction purposes; an image of a depositor's hard disk; an old version of a file that is no longer accessible using current computing environments).

Paradigm therefore proposes that:

- Born-digital archives be held in preservation-only repositories while they are closed to researchers.
- That institutions secure online preservation repositories appropriately, or opt for offline repositories that are simpler to secure (but will require a local back up routine).
- When an archive, or parts of it, is opened to research, readers may order a dissemination copy of a born-digital archive to read in a controlled environment.
- When all restrictions relating to privacy, rights and other content liability expire, 'access copies' of born-digital archives may be published to an online 'access repository', but master copies should remain in the preservation repository. The preserva-

1 Online Paradigm Workbook. URL: <http://www.paradigm.ac.uk/workbook/>

tion service should be responsible for supplying new versions of 'access copies' in accessible formats as appropriate.

This solution means that those functions of preservation repositories that would be redundant in an access repository need not be imposed on them and vice versa.

# Comparing repository software for preserving personal digital archives

## Methodology

### Defining requirements

The needs of managers, archivists, system administrators and developers should be considered when evaluating repository software. These needs are several, sometimes conflicting and overlapping between groups in such a way that it is challenging to produce a coherent model for comparison. Many needs can be characterised as functional requirements, but others which relate to extensibility of the software, sustainability of the user/developer community and usability are equally important. Paradigm drew on many sources to draw up evaluation criteria, including the archivists' experience of traditional archival principles, functions and activities; the OAIS model; other repository evaluations and the RLG/NARA Audit Checklist for Certifying Digital Repositories.[1]

### Evaluation methodology

The project did not adopt a formal methodology to evaluate the repositories. However the following were among the methods used:

- Examining the release documentation of both repositories.
- Installing and configuring the repositories (and some related tools) in a test-bed environment.
- Examining documentation from members of the user community, both on the project wikis and at institutional and project websites.
- Examining earlier comparisons of repository software.
- Monitoring the mailing lists for both repository communities.
- Talking to other users of the software.
- Testing the repositories with exemplar collections of born-digital archives – this included archivists preparing and submitting materials for ingest and researchers testing access mechanisms.

## General findings

A number of general points can be made about Paradigm's experiences with DSpace and Fedora.

### DSpace and Fedora are not equal

A more detailed comparison of DSpace and Fedora appears later in this chapter, but it is worth noting here that the two have some fundamental differences that may help potential users decide which is most suitable for their needs. While the two softwares are different, both user communities participate in a wider 'repository community', which brings together users of many repository softwares such as Eprints and Greenstone, as well as DSpace and Fedora. This inter-community dialogue encourages basic interoperability and an exchange of ideas and practice.

1 RLG and NARA, *An Audit Checklist for the Certification of Trusted Digital Repositories*, Draft for Public Comment (August 2005). URL: <http://www.rlg.org/en/page.php?Page_ID=20769>

### DSpace

**Background:** developed jointly by MIT Libraries and Hewlett-Packard Labs to act as a repository for the intellectual output of research organisations between March 2000 and November 2002.

**Licence:** BSD

**Current version:** 1.4.1 (since 7 December 2006)

**Technology:** DSpace is written in Java, and provides a Java Application Programming Interface (API), a web application that runs in Apache Tomcat and command line tools. An architecture diagram is available.[1]

**Data model:** DSpace repositories create:

- Communities (e.g. a university department), which can have Collections and Sub-Collections.
- Collections and Sub-Collections are containers for grouping related Items; Collections can be part of multiple Communities.
- Items can be part of multiple Collections and contain Bundles.
- Bundles contain one or more Bitsreams (DSpace refers to digital files as Bitsreams).

**Storage:** Metadata is stored in a relational database management system (RDBMS); data is stored in a file system.

**Version(s) tested:** 1.3.2

**Test-bed platform:** SUSE Linux 10, Apache Tomcat, Java, PostgreSQL (actually an object relational DBMS).

**Development priorities:** unclear at present; a technical architecture group has been formed as part of the DSpace governance.

DSpace has been implemented by several institutions looking to develop repositories of simple objects, such as academic papers or e-theses, to enhance their accessibility to the research community. Membership of the DSpace community also includes institutions developing an interest in preservation and curating other kinds of materials, such as images and datasets.

DSpace is self-contained and straightforward; a usable system can be deployed out-of-the-box with relative ease. This simplicity is largely due to the fact that DSpace comes pre-configured with a standard user model, data model and workflows. This ready-made simplicity undermines the utility of the software in an archival context; the software's models and workflows are strongly biased towards an open access repository for academic output, which was the original purpose of the design, and are not well suited to the highly structured collections or complex objects that are commonly associated with personal digital archives.

The recommendations of the DSpace Architecture Review Group suggest that Version 2 of DSpace (no exact release date, but 2009 is probable) will bear more resemblance to Fedora, by providing better functionality for a range of contexts and activities, an extension framework for third party developers, and the ability to operate on a large scale.

### Fedora

**Background:** developed at the Universities of Cornell and Virginia with funding from the Andrew W. Mellon foundation. Now established as Fedora Commons, a non-profit organisation. Recently awarded a $4.9M grant for further development from the Gordon and Betty Moore Foundation.

**Licence:** Fedora is available under the Educational Community License 1.0 (ECL); third party packages associated with its use are distributed under a variety of other licences.

**Current version:** 2.2 (since 19 January 2007)

1 DSpace, 'DSpaceSystem Documentation: Architecture', *DSpace website*. URL: <http://www.dspace.org/index.php?option=com_content&task=view&id=145>

**Technology:** Fedora is written in Java and runs as a web application in Apache Tomcat. It provides a number of open APIs that are exposed as SOAP and REST web services: Management API (API-M), Access API (API-A), Access-Lite API (API-A-Lite, also includes Search API), Management-Lite API (API-M-Lite) and Resource Index Search API. Fedora also supplies a client with a Graphical User Interface (GUI) and command line tools.

*Fedora provides three local web services:* Saxon XSLT Processor Local Service, FOP Local Service (for PDF Transformation) and the Image Manipulation Local Service.

*The Fedora framework currently provides three services that interface with the Fedora repository service:* Generic Search Service (GSearch), Directory Ingest Service (DirIngest) with a GUI tool called SIP Creator for preparing submissions to the DirIngest service and OAI Provider Service (PROAI).

**Data model:** Fedora has three kinds of object: data, behaviour definition and behaviour mechanism.

- Data Objects must contain an ID and Dublin Core metadata; they may also contain one or more datastreams (digital file), XML metadata of any kind and RDF XML metadata to describe relationships with other objects.
- Behaviour definition and behaviour mechanism objects provide disseminators for one or more datastreams in data objects.

**Storage:** Metadata is stored in a relational database management system (RDBMS); data is stored in a file system within the repository or externally.

**Test-bed platform:** SUSE Linux 10, Apache Tomcat, Java, PostgreSQL.

**Version(s) tested:** 2.1, 2.1.1, 2.2.

**Development priorities:** the community has established a series of working groups: Preservation Services, Search Services, Workflow Services and Content Models for Datastreams and Disseminators. The envisaged framework includes services for preservation monitoring, event notification, etc., but it is unclear when these will be implemented.

Fedora was designed to be a repository for all materials and all purposes from the beginning, although it is fair to say that many early users were developing repositories for access purposes. The Fedora community has evolved alongside the DSpace community; several members are now using Fedora for preservation and for complex objects and highly structured collections. The Fedora community has demonstrated an interest in preservation functions with the formation of a preservation working group (established 2005) and some useful preservation-related features in the 2.2 release of the software (January 2007).

Fedora is more complex than DSpace because it is a repository architecture as much as a repository. It was designed to be flexible, so that users could employ any kind of data model; and to be extensible, so that users could add whatever clients or services they needed to the Fedora framework. As a Fedora repository matures, it is likely to use several web services – this creates a distributed service-oriented architecture system, as opposed to the self-contained system presented by DSpace. This flexibility has immense potential, but comes at a cost. The implementing institution must do more than basic installation, configuration and customisation of the software; it must be prepared to design its own user models, data models, workflows and tools, or to adopt them from comparable implementations within the Fedora community. This means that Fedora-based repositories can be very different to one another and strong analytical and programming skills may be needed by the repository team.

Fedora does provide a client and web interface out-of-the-box, This basic install feels unpolished and incomplete; the system designers anticipated that adopters would design content models, services and interfaces particular to their varied implementation needs. The community has produced and published tools for some tasks, but the ‘open access’ origins of the repository movement mean that many relate to access rather than preservation. Some tools developed by the user community are not well documented for use by newcomers; others are not shared with the wider community at all. The availability of metadata and preservation tools is likely to improve as the combined experience of the Fedora repository community grows.

## Scope and content of documentation

Comparing repository softwares is no small task for those new to repositories as much of the supporting documentation is aimed at technical audiences (developers and systems administrators), rather than managerial/professional users (archivists and librarians) or end-users (researchers). The nature of the documentation makes the process of learning how to install, configure and use the repositories, much less evaluate them, time-consuming. Some of the problems Paradigm encountered include:

- It is difficult for newcomers to know what to read; some guidance of what to read, who should read what, where to find it and what order to read it in would be useful.
- Up-to-date management documentation, which translates the technical attributes of the system into features and benefits which might inform a selection process, or implementation decisions, is lacking.
- User-level documentation that would assist non-technical users in utilising the systems needs improvement.
- Technical documentation for system administrators and software developers getting to grips with the system is very distributed and could be organised better.
- The user-community needs access to documentation for older versions of the software, but it can be difficult to identify which release a document relates to, or whether an existing a document has been updated to describe features in a new release.
- In documentation published by institutions or projects, it is not always clear which version of the software is being referred to.
- The explanations for design choices are not always available; these would help newcomers understand why some features work the way they do, what they might be used for, or which configuration might suit their needs best.

Documentation provides instructions for configuring features, such as ‘authorisation’ or ‘search’, but does not detail the benefits of such features in language that is readily accessible to those whose needs should determine the selection and customisation of repository systems. Much of the documentation aims to answer ‘how’ questions; some more attention to ‘why’ questions and relevant use cases would be beneficial. There is very little in the way of supporting information that could assist implementation decisions; information which describes the advantages and disadvantages of a given implementation, and its intended purpose, would be a helpful addition to the documentation of repository software.

## The boundaries of repository software

There is no end-to-end repository system available and no clear-cut definition stipulating which functions and activities should be the responsibility of the repository software and which should be the responsibility of another entity. Some aspects of policy and procedure may form part of a manual process and need not be automated at all, while others may be automated but devolved to a service outside the repository software’s remit. Some important, non-trivial, parts of the preservation process take place outside of repository software. One example is the assembly of metadata required to support the digital archives into METS Archival Information Packages for submission to the repository system. Some functions and activities which would be desirable in preservation repositories, such as obsolescence monitoring and interoperability with external file format registries, are partly dependent on services external to the repository’s organisation.

The criteria presented below must be met by the repository system used by the repository service. The repository software could form all of the system, or a central part of that system. This means that some of these criteria need not be the responsibility of the repository software, but of a service that may be used in conjunction with that software.

## Coming soon

Some of the functionality that might interest those implementing preservation repositories has

yet to be implemented in any repository software, but is on the development roadmap of several. It can be difficult to keep track of what functionality is in the pipeline and when it is due to arrive. The availability of more detailed information about current developments in the community would improve co-ordination among adopting institutions and be useful in planning local development priorities.

## Archival concerns

The care of material in a repository catering for born-digital personal archives is the duty of professional archival staff. Archivists must be able to understand and have faith in the system's security and its processes, and be able to interact with the system confidently. Archivists have a duty of care to ensure the authenticity, continuing availability and robustness of archival material, both to the creators of archival material and the researchers who will use it, and to support the eventual use of archival material by researchers and to satisfy freedom of information requests.

| **Name** | ***Supports audit trails*** |
|---|---|
| Detail | *Born-digital personal archives will be retained by their collecting institutions indefinitely, and it can be assumed that during this time items will be moved to new storage media and new formats on numerous occasions. The repository must provide mechanisms to demonstrate that an item is as it was when submitted to the repository – that it is authentic.* |
| DSpace | Partial support |
| Detail | Some activities, such as submitting and approving a bitstream, are recorded as qualified Dublin Core metadata using description provenance – this records the name, the date and time, filenames, size in bytes and an MD5 checksum. |
| Fedora | Partial support |
| Detail | Modifications to files or metadata are logged in Fedora's audit metadata, which records information about who did what and when, and associates it with the object. Metadata associated with migration or preservation events are not created by Fedora, though Fedora could support the addition of such metadata. |

| **Name** | ***Supports unique identification of metadata, digital files and conceptual objects.*** |
|---|---|
| Detail | *In order to maintain intellectual control of items in the repository it must be possible to apply unique identifiers to objects and metadata managed by the repository.* |
| DSpace | Supported |
| Detail | Each Community, Collection and Item is allocated a Handle in the current version of DSpace. In version 2 DSpace will support other persistent identifiers in addition to Handles, and it will be possible to apply identifiers at more granular levels. |
| Fedora | Supported |
| Detail | Each object, file and metadata (and version thereof) is given a unique identifier by the Fedora repository. Repositories may opt to use the Handle system with Fedora. For example, see the VTLS OSC suite of tools, which includes a service for integrating the Handle System with Fedora.[1] |

| **Name** | ***Supports reliable binding of metadata and digital object*** |
|---|---|
| Detail | *In order to maintain intellectual control of items in the repository it must be possible to permanently associate an archival item with its metadata, both within the repository and in any export functionality.* |
| DSpace | Partial support |
| Detail | Each Community, Collection and Item can have its own metadata. Individual files which make up an item are allocated basic metadata, which is displayed by the containing Item's metadata; it is unclear which metadata belongs to which file. Recommendations for version 2 of DSpace include allowing metadata at more granular levels. |
| Fedora | Partial support |

1 Fedora, 'Tools', *Fedora website*. URL: <http://www.fedora.info/tools/>

| | |
|---|---|
| Detail | Dependent on implementation. If multiple files and their metadata are stored within a single object wrapper, then the repository must itself implement conventions which specify which metadata belongs to which files. If repositories use an atomistic model, with one file and its metadata to an object, metadata and object are unambiguously connected. |

| **Name** | ***Supports referenced metadata*** |
|---|---|
| Detail | *Some metadata is applicable to several objects and is best held once, such as an EAD collection level archival description, or rights metadata. Other metadata, such as file format registries, may be curated in repositories external to the organisation.* |
| DSpace | Not supported |
| Detail | |
| Fedora | Supported |
| Detail | Files and metadata may be held outside of the repository and referred to; relationships between objects in the repository can also be formed. |

| **Name** | ***Supports complex inter-object relationships*** |
|---|---|
| Detail | *Meaning in archival materials relies heavily on context; it is necessary that the repository supports complex hierarchical relationships found in archives.* |
| DSpace | Not supported |
| Detail | The DSpace data model is designed for flatter collections and is not well-suited to complex structures. |
| Fedora | Supported |
| Detail | Fedora can support complex multi-level relationships through its RDF metadata. It is also possible to ingest METS structural maps to reflect the original order of an archival accession, to ensure that this is preserved for the archivist who will catalogue the archive. |

| **Name** | ***Supports appropriate metadata standards*** |
|---|---|
| Detail | *Support for open and widely adopted metadata standards increases object portability, tool availability and the likelihood of recruiting staff familiar with metadata employed by the repository. Support for PREMIS preservation metadata has not been incorporated into any repository yet, and support for technical metadata is very limited.* |
| DSpace | Partial support |
| Detail | METS, OAI-PMH and Dublin Core are supported. Additional metadata may be added as 'serialized datastreams'. |
| Fedora | Partial support |
| Detail | Fedora stores metadata in its native FOXML, which can be exported to METS (a Fedora extension of METS); it also supports OAI-PMH and Dublin Core. Fedora can store any kind of valid XML metadata and can be configured to index this metadata using the Fedora Generic Search Service.<br><br>Metadata extraction support is limited to a web service for the Jhove validation and technical metadata extraction. There are currently no tools to generate or act on PREMIS preservation metadata.<br><br>Relationships between objects can be recorded using METS structural maps or via RDF metadata, but Fedora provides no interfacing with those relationships (e.g. would not display a complex object). |

| **Name** | ***Supports simple and complex objects*** |
|---|---|
| Detail | *Personal digital archives contain a range of simple objects, consisting of a single file, and complex objects, which are composed of multiple files that must be reassembled to recreate the object. The repository should be capable of supporting both kinds of object.* |
| DSpace | Partial support |
| Detail | Allows multiple files to be bundled together in an Item, but this limits the metadata that can be applied. |

| | |
|---|---|
| Fedora | Supported |
| Detail | The Fedora data model allows users to bundle files together in an object, or to store files in their own objects and create relationships between them. |

| **Name** | ***Supports multiple types and formats*** |
|---|---|
| Detail | *Personal digital archives can contain a wide variety of material, from email to simple image files, from spreadsheets to word-processed documents, from websites to audio files. The repository should be capable of supporting a wide range of object types and formats.* |
| DSpace | Supported |
| Detail | DSpace has a bitstream registry which details the formats that the repository accepts, and the level of support the repository provides for them. Additional formats may be added to the registry. |
| Fedora | Supported |
| Detail | Supports any mime-type. |

| **Name** | ***Supports automatic metadata creation*** |
|---|---|
| Detail | *The preservation of born-digital archives requires a great deal of metadata. The automation of this metadata is extremely advantageous.* |
| DSpace | Partial support |
| Detail | Some audit metadata, etc., is created automatically. Much of the metadata must be input through the web user interface. |
| Fedora | Partial support |
| Detail | Audit metadata is created automatically, and checksum metadata may be created automatically. A Jhove Metadata Extraction Service[1] is available to add some technical metadata. The SIP Creator/Dir Ingest service[2] can automate the creation of relationship metadata. Much metadata, including descriptive and preservation metadata, must be compiled manually. |

| **Name** | ***Supports bulk ingest*** |
|---|---|
| Detail | *Digital materials must be properly ingested into a managed environment as soon as possible, bulk ingest is therefore highly desirable.* |
| DSpace | Supported |
| Detail | Provides a command-line bulk ingest tool; files must be arranged according to a specified hierarchy to map to the DSpace data model. |
| Fedora | Supported |
| Detail | The Fedora Management web service has SOAP-based operations to ingest digital objects in different XML wrapper formats (METS and FOXML). This same web service has other SOAP-based operations to add datastream content to an object that is already in the Fedora repository. Fedora also has a separate “Directory Ingest” service that runs as a web application; this service accepts a zip file that contains a hierarchical directory of files along with a METS manifest file, opens the zip file and calls the Fedora Management web service to ingest each file as a digital object, preserving the hierarchical directory relationships. |

| **Name** | ***Supports bulk export*** |
|---|---|
| Detail | *Bulk export will be necessary for an institution moving to another repository technology, or one returning deposited materials to a creator. Archival materials and their metadata are likely to be moved to the next version of the repository software, and beyond that will one day be migrated to an entirely new system. It should be possible to easily migrate objects and metadata, and preference should therefore be given to implementations of metadata standards which are open and widely adopted.* |
| DSpace | Supported |
| Detail | Provides a command-line tool (dspace-export) that outputs a METS file per collection with references to the digital files (called bitstreams by DSpace) in the collection. DSpace can also export in the DSpace ingest format. |

1 Visionary Technology in Library Solutions, ‘VTLS Open Source Components’, *Visionary Technology in Library Solutions website*. URL: <http://www.vtls.com/Products/osc.shtml>

2 Fedora, ‘Fedora Directory Ingest Service’, *Fedora website*. URL: <http://www.fedora.info/download/2.2/services/diringest/doc/index.html>

| | |
|---|---|
| Fedora | Supported |
| Detail | From the GUI client, command-line or through a homegrown SOAP client. |

| **Name** | ***Supports appropriate content models*** |
|---|---|
| Detail | *Content models allow repositories to specify how particular classes of object should be treated. This increases efficiency and quality.* |
| DSpace | Not supported |
| Detail | The DSpace content model is rigid, and characterised by the Community and collection concepts of a repository for academic output. |
| Fedora | Partial support |
| Detail | Fedora allows the user to define their own content models. Work on formalising content models,[1] including defining a content model definition language, is underway. |

| **Name** | ***Supports format identification*** |
|---|---|
| Detail | *Reliably identifies an object as being of a particular format and assigns this metadata.* |
| DSpace | Not supported |
| Detail | Objects are associated with a format manually. The permitted bitstream formats recognised by the system are stored in the bitstream format registry. The contents of the bitstream format registry are entirely user-defined, though the system requires that the two default formats are present (Unknown and License). |
| Fedora | Not supported |
| Detail | Datastreams are manually associated with a mime type and optionally a format URI (this is a user-assigned URI which supports identification of the media type of an object in a more specific way than using a MIME type). |

| **Name** | ***Supports file validation*** |
|---|---|
| Detail | *Validates an object against a specification to evaluate its correctness and completeness.* |
| DSpace | Not supported |
| Detail | A command-line tool to run Jhove over the DSpace asset store has been developed by the DSpace community.[2] |
| Fedora | Not supported |
| Detail | Use of the Jhove tool in conjunction with Fedora provides validation for some formats. |

| **Name** | ***Supports versioning*** |
|---|---|
| Detail | *Allows the repository to keep older versions of metadata and files.* |
| DSpace | Not supported |
| Detail | The proposed changes to come in version 2 of DSpace will introduce versioning and the concept of Manifestations for Items, which may have their own metadata records. |
| Fedora | Supported |
| Detail | As of version 2.2, Fedora allows users to decide whether each metadata or digital file is versionable, or whether older versions should be overwritten by newer versions. For datastreams or metadata that are versionable, changes result in a new timestamped version being created. Older versions remain accessible. |

| **Name** | ***Easy to use workflows*** |
|---|---|
| Detail | *Archivists must work with the repository in order to apply professional treatment to the processing of these assets. It is important that repository interfaces support use by less technical users.* |
| DSpace | Partial support |

1 Fedora, 'Content Models Overview', *Fedora website.* URL: <http://www.fedora.info/wiki/index.php/Content_Models_Overview>

2 DSpace, 'TechMDExtractor', *DSpace website.* URL: <http://wiki.dspace.org/index.php/TechMDExtractor>

| | |
|---|---|
| Detail | Provides ingest workflow via a web user interface for non-technical users. The architecture group has proposed that version 2 of DSpace support a wider variety of workflows, which go beyond initial ingest and include migration, versioning and export and that these should be configured by users through interfaces provided by DSpace. The DSpace community are also evaluating workflow engines. |
| Fedora | Not supported |
| Detail | Fedora's design anticipates the creation of a workflow outside of the repository. It provides a basic client which is usable (with training) for working on single items, but the open source workflow interfaces designed by other Fedora users (such as Fez and Elated) do not meet the processing requirements for archival materials. |

| **Name** | ***Supports appropriate security mechanisms*** |
|---|---|
| Detail | *Born-digital archives will often be subject to embargo for a number of years owing to privacy and other concerns. Once privacy concerns cease, copyright still influences the manner in which the archives may be used. Security is of the utmost importance in building the confidence of potential donors; a security breach could be disastrous for the reputation of an archival repository and could have serious implications for collection development.* |
| DSpace | Supported |
| Detail | Provides data transfer encryption (SSL).<br>Authenticates users via a web user interface or LDAP.<br>Supports different user accounts and roles, and has a web interface for editing permission policies.<br>From version 2 Epeople (DSpace terminology for users) will have persistent identifiers in the form of URIs.<br>Direct access to Java API, database and filesystem requires user privileges on the machine hosting the DSpace repository. |
| Fedora | Supported |
| Detail | See Fedora's security documentation.[1]<br>Can restrict access to Management and Access APIs based on IP address.<br>Management API is protected by basic HTTP authentication.<br>Can provide data transfer encryption (SSL).<br>Can create multiple users (with roles and permissions that can be used in XACML access policies) in fedora-users.xml file; by default supports a single known user (fedoraAdmin) and other users are anonymous. Multiple users are needed for audit trail purposes.<br>Can defer authentication to application; Fedora therefore authenticates the application and expects the application to undertake user authentication.<br>XACML can be used to define repository level policies and item-level policies. Policies can be very granular, e.g. restricting access to a file but allowing metadata access.<br>Repository administrators are expected to provide the storage locations of metadata and content objects with adequate security.<br>As of v 2.2 Fedora can authenticate users against an LDAP server.[2] |

| **Name** | ***Supports technology watch*** |
|---|---|
| Detail | *A digital repository of personal digital archives will contain multiple material types which are submitted in a variety of different formats. It will be necessary to automate some technology watch functions to monitor the status of the materials in the archive so that preservation actions can be planned, prioritised and implemented as necessary. The repository should alert administrators to file formats which are at risk of obsolescence.* |
| DSpace | Not supported |
| Detail | An event mechanism has been proposed for version 2 of DSpace and the current Event-Mechanism prototype being worked on for version 1.5 might provide a basis to meet this requirement. |
| Fedora | Not supported |
| Detail | A preservation monitoring service (based on event notification) is planned. |

1 Fedora, 'Securing Your Fedora Repository', *Fedora website.* URL: <http://www.fedora.info/download/2.2/userdocs/server/security/securingrepo.html>

2 Fedora, 'Authenticating Fedora Against LDAP', *Fedora website*. URL: < http://www.fedora.info/wiki/index.php/Authenticating_Fedora_2.2_against_LDAP>

| Name | ***Supports notification of objects due for review, or opening for research*** |
|---|---|
| Detail | *The repository should notify the administrator when objects can be made accessible to researchers, or when their status should be reviewed.* |
| DSpace | Not supported |
| Detail | An event mechanism has been proposed for version 2 of DSpace and the current Event-mechanism prototype being worked on for version 1.5 might provide a basis to meet this requirement. |
| Fedora | Not supported |
| Detail | If the planned event notification service materialises this might satisfy this requirement. |

| Name | ***Provides reporting features*** |
|---|---|
| Detail | *The repository should be able to generate statistics that would be useful for planning and prioritising preservation strategies. One such report might be on the file formats represented in the repository. It should also be able to provide useful statistical information, such as the quantity and quality of material ingested into the repository in a given period.* |
| DSpace | Partial support |
| Detail | Some statistical reports can be generated by analysing DSpace's log files. |
| Fedora | Not supported |
| Detail | The features documentation alludes to a reporting utility which does not appear to exist? |

| Name | ***Supports digital provenance metadata*** |
|---|---|
| Detail | *The repository should allow users to trace migrated objects back to the original submission, with an account of the object's migration history.* |
| DSpace | Partial support (experimental) |
| Detail | The History subsystem (referred to at the DSpace Sourceforge website[1]) is explicitly invoked when significant events occur (e.g., accepting an item into the archive). The functionality of this part of DSpace is documented as a largely untested experiment. A replacement for inclusion in version 1.5 is being worked on.[2] |
| Fedora | Supported |
| Detail | As of version 2.2, Fedora supports journaling alongside the existing auditing and versioning functionality. There is no explicit functionality though to provide an account history at present and how the digital provenance metadata could be used would be dependent on the content model used. |

| Name | ***Supports integrity monitoring for metadata and objects*** |
|---|---|
| Detail | *The repository should monitor digital objects and metadata to ensure that they have not been damaged accidentally, through media failure or maliciously. The OAIS model refers to this as fixity information.* |
| DSpace | Supported |
| Detail | Since version 1.4 DSpace has supported checksum checking via a command line tool.[3] Digital signatures are not supported. |
| Fedora | Supported |
| Detail | As of version 2.2, Fedora supports the addition of a checksum to all digital files and metadata that can be checked by the repository.[4] Digital signatures are not supported. |

| Name | ***Supports backup and restore*** |
|---|---|
| Detail | *The repository should be easily restored from backup in the event of a disaster.* |
| DSpace | Supported |

1 DSpace 'Dspace History System', *DSpace website*. URL: <http://www.dspace.org/index.php?option=com_content&task=view&id=149#history>
2 DSpace, ' History System Protoype for DSpace 1.5', *DSpace website.* URL: <http://wiki.dspace.org/index.php/HistorySystemPrototype>
3 DSpace, 'Configure checksum checker', *DSpace website.* URL: <http://wiki.dspace.org/index.php/Configure_checksum_checker>
4 Fedora, 'Checksums on Datastreams in Fedora', *Fedora website.* URL: <http://www.fedora.info/download/2.2/userdocs/server/features/checksumming.html>

| | |
|---|---|
| Detail | Information on how to organise backup for a DSpace repository is available.[1] |
| Fedora | Supported |
| Detail | Fedora 2.2 provides a journaling module[2] that allows a repository to be mirrored, or to restore a Fedora repository to the exact state before failure, rather than the state at last backup. |

| **Name** | ***Is extensible*** |
|---|---|
| Detail | *The longer-term sustainability of the system will be reliant on its modularity. Monolithic systems are not easily updated to accommodate new needs, while modular systems can be enhanced piecemeal.* |
| DSpace | Supported |
| Detail | Supports add-ons; DSpace has rules for 'well-behaved add-ons',[3] but the community has acknowledged that this design should be changed; the architecture group is therefore recommending the adoption of an open source extension framework in version 2 of DSpace. |
| Fedora | Partial support |
| Detail | The repository software and related services can be distributed over different hardware. Additional homegrown or externally sourced services may be added to the Fedora framework. |

| **Name** | ***Is scalable*** |
|---|---|
| Detail | *At present, the volume of born-digital archives relative to their paper counterparts is small. This balance will change over time and archives can expect to receive greater quantities of digital materials in future. The volume of metadata will also increase over time, and migrated versions of objects and emulators with their own metadata may be added to the repository. The repository system should scale to manage millions of digital materials; this requires the repository to have the capacity to manage large quantities of material, to support mass throughput of material when ingesting and exporting, and to support several concurrent processes while maintaining acceptable performance.* |
| DSpace | Not supported |
| Detail | DSpace is known to have scalability problems; as is, it may be suitable as a short-term repository. The architecture group working on version 2 of DSpace are aiming to make the software scale to 10 million items and have made recommendations that may improve the architecture of the repository. |
| Fedora | Supported |
| Detail | NSDL have tested Fedora with 1 million objects, and the community is looking to test up to 30 million objects. |

| **Name** | ***Supports basic searching*** |
|---|---|
| Detail | *Searching across key metadata fields and ideally full text searching for textual objects will facilitate archivist- and researcher-generated queries.* |
| DSpace | Supported |
| Detail | DSpace supports searching for one or more keywords in metadata or extracted full-text and browsing though title, author, date or subject indexes. DSpace uses the Lucene search engine and the search indexes are configurable, enabling customisation of which DSpace metadata fields are indexed. |
| Fedora | Supported |

1 DSpace, ' Backing Up and Restoring a DSpace Instance', *DSpace website.* URL: <http://wiki.dspace.org/index.php/BackupRestore>

2 Fedora, 'Journalling Guide', *Fedora website.* URL: <http://www.fedora.info/download/2.2/userdocs/server/journal/index.html>

3 DSpace, 'AddOnMechanism', *DSpace website.* URL: <http://wiki.dspace.org/index.php/AddOnMechanism>

| | |
|---|---|
| Detail | Fedora indexes select system metadata fields and the primary Dublin Core record for each object. The Fedora repository system provides a search interface for both full text and field-specific queries across these metadata fields.[1]<br>The Gsearch service introduced with version 2.2 augments this with indexing of Fedora FOXML records, including the text contents of datastreams and the results of disseminator calls, searching the index, and the ability to plugin selected search engines, so far Lucene and Zebra.[2] |

## Managerial concerns

Managers must be content that commitment to a given open source repository and its community is a sound managerial decision. The software may be free, but there are costs and risks that should be assessed. Factors relevant here relate mainly to the depth and breadth of the user community and the organisation's ability to provide ongoing support for the software:

| **Name** | ***Community is sustainable*** |
|---|---|
| Detail | *Both DSpace and Fedora originated as projects and are now becoming open source communities. Committed community members and mechanisms for distributed governance, development and support are critical to this transition and the sustainability of the community. Selection of DSpace or Fedora may not require capital expenditure on software licences, but will require commitment of time and personnel that may amount to equal expense, albeit for better value. Given the resources required, it is important to choose a solution which can be sustained in the medium- to long-term.* |
| DSpace | Supported |
| Detail | Has a governance structure[3] including a core group of committers who are permitted to commit changes out of a larger group of contributers of.[4] There are several live installations.[5] |
| Fedora | Supported |
| Detail | Fedora users hail from a variety of groups: academic computing groups, research libraries, archives, publishing societies, government agencies and commercial vendors. This is because the repository is intended to act as a foundation for several kinds of information management systems; its design is therefore flexible enough to accommodate multiple usages and their data models. The software has a growing user-base in Europe, Australia and the United States.[6]<br>Community governance, working groups and developer guidelines are being drawn up as the transition to fully open source development takes place. The Fedora Project has evolved into a non-profit organisation called Fedora Commons to act as custodian of the software platform and steer its future direction.[7] It was awarded $4.9M from the Gordon and Betty Moore Foundation in August 2007. |

1 Fedora, 'Fedora Search Interface Documentation', Fedora website. URL: <http://www.fedora.info/download/2.2/userdocs/server/webservices/search/index.html>
2 Fedora, 'Fedora Generic Search', Fedora website. URL: <http://www.fedora.info/download/2.2/services/genericsearch/doc/index.html>
3 DSpace, 'DSpace Federation Governance', *DSpace website.* URL: <http://wiki.dspace.org/index.php/DspaceGovernance>
4 DSpace, 'DspaceContributors', *DSpace website.* URL: <http://wiki.dspace.org/index.php/DspaceContributors>
5 DSpace, 'DspaceInstances', *DSpace website.* URL: <http://wiki.dspace.org/index.php/DspaceInstances>
6 Fedora, 'Fedora Commons Portfolio of Projects', *Fedora website*. URL: < http://www.fedora-commons.org/portfolio>
7 Disruptive Library Technology Jester, 'A Vision for FEDORA's Future, an Implementation Plan to Get There, and a Project Update', *Disruptive Library Technology Jester website.* URL: <http://dltj.org/2007/01/fedora-update/>

| Name | ***Similar users exist*** |
|---|---|
| Detail | *In open source communities, the users of the software are responsible for steering the development of the software. The more users that share your vision, the more likely that the community will add the functionality you need to the software; it is therefore useful to know whether organisations with similar missions have adopted the software. Paradigm was interested in users working in digital preservation, with archival materials or other complex materials.* |
| DSpace | Partial support |
| Detail | Interest in digital preservation has increased in recent years and the DSpace community is interested in adding relevant functionality to the repository software. Archival users of DSpace, or users with complex collections, are smaller in number, but growing. |
| Fedora | Supported |
| Detail | The Fedora community has several members working with complex collections and with an interest in digital preservation. There are also other users working in records management and archival domains. |

| Name | ***Can personnel be recruited/trained easily to support the software in the event of staff turn over?*** |
|---|---|
| Detail | *Is the repository so esoteric that staff turnover is catastrophic? How easy will it be to recruit people to administer, develop and use the software?* |
| DSpace | Partial support |
| Detail | DSpace is largely self-contained and uses mainstream open source components; it should be relatively easy to recruit and/or train appropriate technical personnel. Badly or undocumented esoteric organisation specific modifications may cause problems, but these can be mitigated by following the open source philosophy of feeding back changes (so long as these are accepted by the community) and by documenting local customisations. Training users should be straightforward. Provision of training for new and intermediate DSpace administrators, repository managers and developers would be useful. |
| Fedora | Partial support |
| Detail | Fedora also uses mainstream open source components, but its implementation is more flexible than DSpace and there is a steeper learning curve for technical and non-technical staff. If a well-documented user-friendly installation is present, training technical and non-technical users in its use is less problematic. If a repository is in development, this could pose greater problems. Provision of training for new and intermediate Fedora administrators, repository managers and developers would be useful. |

| Name | ***Support available*** |
|---|---|
| Detail | *Support can and should be provided by a variety of methods:*<br>• *Good quality authoritative documentation appropriate to the audience.*<br>• *Informal, wiki-like, documentation.*<br>• *Commercial enterprise, selling support for the repository.*<br>• *Community leaders.*<br>• *National user groups.*<br>• *International user groups.*<br>• *Special interest user groups.*<br>• *Conferences.*<br>• *Active mailing lists.* |
| DSpace | Supported |

| | |
|---|---|
| Detail | Documentation tends to be technical in nature and it would be helpful if it were re-organised to meet the needs of the different groups within the DSpace community. The consolidation of documentation in one place (clearly identifying which version it supports and who should read it) would help. There is some end-user and management information at <http://www.dspace.org>, but establishing what Dspace can or cannot do can require significant effort as you also need to review the information on the wiki[1] and identify which version the text concerns.<br>For technical support a good starting point is:<br>• <http://wiki.dspace.org/index.php/TechnicalFaq><br>Commercial support of DSpace (for set-up or ongoing support) in the UK or elsewhere is not easy to locate. Some references from the DSpace website would be useful.<br>DSpace User Group meetings are held regularly.<br>Community support in the UK may be provided by the JISC Repositories Support Project.[2]<br>DSpace has an active mailing list. |
| Fedora | Supported |
| Detail | Documentation is predominantly aimed at a technical audience. The rationale for some implementation decisions needs to be explained more clearly for a non-technical audience who are often the key-decision makers. The tutorials need to be updated to reduce the learning curve. As Fedora is a flexible framework oriented approach, more documentation on sample usage scenarios and, for example, content-model how-tos would be beneficial. There are working groups exploring some of these areas and progress reports would be of assistance.<br>Support is via the mailing-list and the Wiki does capture some how-to information.<br>Vendor support is available through VTLS, who provide additional services to Fedora, though these are currently based on an earlier version of Fedora and therefore lack much of the useful preservation- and authenticity-related features available in the current version of Fedora (2.2).<br>The Fedora community does meet through regular conferences, but the larger gatherings tend to be in the United States. A UK and Ireland Fedora Users Group has been established for more local support. Community support in the UK may be provided by the JISC Repositories Support Project.[3]<br>Fedora has an active mailing list.<br>Fedora has established a number of working groups to take forward development of the community and the software. |

| **Name** | ***Has realistic learning curve*** |
|---|---|
| Detail | *Repository implementations are likely to start small, meaning reliance on a small number of curatorial and technical staff who have familiarised themselves with the software. Repository systems which are difficult for new staff to understand present a risk in a fluid labour market.* |
| DSpace | Supported |
| Detail | DSpace administrators are supported by the DSpace System Documentation and the online community. New administrators will still need time to familiarise themselves with the repository, but DSpace provides documentation aimed at new system administrators. Ordinary users should find the existing DSpace user interface reasonably familiar and the online help should prove sufficient. A printable version of the online help user guide does not appear to be available. |
| Fedora | Partial support |
| Detail | The learning curve for Fedora is quite steep, but the system documentation is reasonably clear and Fedora has a supportive online community who can provide assistance when the documentation is unclear. Ingest tools for Fedora are slowly appearing which will reduce the repository specific component of the learning curve for non-technical users which, at present, is still quite high. |

1 DSpace, 'DSpace Wiki', *DSpace website.* URL: <http://wiki.dspace.org/index.php/Main_Page>
2 SHERPA, 'Repositories Support Project - RSP', *SHERPA website.* <http://www.sherpa.ac.uk/projects/rsp.html>
3 SHERPA, 'Repositories Support Project - RSP', *SHERPA website.* <http://www.sherpa.ac.uk/projects/rsp.html>

| Name | ***Availability of information for planning purposes*** |
|---|---|
| Detail | *Managers need access to development priorities and timetables. Knowing when features will be implemented can help institutions decide when they should start working on a local solution to a problem (which should be submitted for consideration of inclusion in the main codebase) and whether they should wait for a central response.* |
| DSpace | Partial support |
| Detail | The DSpace Architecture Review Group has published recommendations which will inform development of version 2. Confirmation of what will be implemented when is not yet available. See DSpace Architecture Review Group, *Toward the Next Generation: Recommendations for the next DSpace Architecture* (January 24, 2007).[1] |
| Fedora | Partial support |
| Detail | The priorities are not explicit and information on the website is out-of-date. The nature of the working groups and diagrams of the proposed Fedora framework provide implicit information. Information about timings is not clear. |

## System administration concerns

System administrators are responsible for administering the infrastructure. They will want to know the answers to a number of questions, such as those listed below:

- What environment – the hardware and software – is needed by the repository?
- How is the environment and repository installed, configured and patched?
- How is the repository to be backed up?
- How is the repository secured? What authentication and authorisation services are available?
- How will the repository fit with existing infrastructure?
- Are upgrades designed to be minimally disruptive?

| Name | ***Provides installation instructions including hardware and software required*** |
|---|---|
| Detail | *System administrators must prepare the environment for installation.* |
| DSpace | Supported |
| Detail | The DSpace wiki[2] has links to the installation documentation in various formats. See the latest DSpace documentation at Sourceforge.[3] |
| Fedora | Supported |
| Detail | Fedora provides clear links to technical documentation for the current release,[4] which includes :<br>• Release notes.<br>• Installation guide.<br>• Update and migration. |

| Name | ***Provides patching information*** |
|---|---|
| Detail | *System administrators must know when patches are available, and whether they should be installed immediately or whether they may adversely affect the repository.* |
| DSpace | Partial support |
| Detail | Patches are communicated via the mailing list. There is the implicit expectation that systems administrators are familiar with the requirements of the repository, monitor the bug-tracking systems and are able to assess the implications of upgrading or patching the system. |
| Fedora | Partial support |

1 DSpace Architecture Review Group, *Toward the next generation: Recommendations for the next DSpace Architecture* (January 2007). URL: <http://wiki.dspace.org/static_files/0/0e/DSpace-recs.pdf>

2 DSpace, 'DSpace System Documentation', *DSpace website.* URL: <http://wiki.dspace.org/index.php/DspaceResources#DSpace_System_Documentation>

3 DSpace, *DSpace System Documentation: Contents.* URL: <http://dspace.svn.sourceforge.net/viewvc/*checkout*/dspace/trunk/dspace/docs/index.html>

4 Fedora, 'Documentation', *Fedora website. URL:* <http://www.fedora.info/documentation/>

| | |
|---|---|
| Detail | Patches are communicated via the mailing list. System administrators are expected to monitor the bug-tracking system and mailing-list and make decisions about the necessity of patches offered there. As Fedora evolves into an open source community this is unlikely to change. |

| **Name** | ***Provides information on backing-up*** |
|---|---|
| Detail | *System administrators need documentation of any special requirements to be aware of in planning backup and disaster recovery strategies for the repository.* |
| DSpace | Supported |
| Detail | Information on how to organise backup for a DSpace repository is available.[1] |
| Fedora | Supported |
| Detail | Fedora 2.2 provides a journaling module[2] that allows a repository to be mirrored, or to restore a Fedora repository to the exact state before failure, rather than the state at last backup, in case of corruption or failure of the repository. The Fedora repository can also be completely rebuilt by crawling the digital object XML source files that are stored on disk. |

| **Name** | ***Provides information on securing repository*** |
|---|---|
| Detail | *What authentication and authorisation services are available? Are they compatible with Shibboleth? How are these documented?* |
| DSpace | Supported |
| Detail | DSpace associates application sessions with specific users and/or groups via a mechanism called Stackable Authentication (permitting custom authentication methods to be stacked on top of the default DSpace username/password method). DSpace also maintains authorisation policies that allow it to understand which credentials are required (if any) to undertake actions on particular resources. Shibboleth can be used with DSpace. |
| Fedora | Supported |
| Detail | Fedora 2.2 supports an Access Control and Authentication module that includes the ability to enforce fine-grained access control policies expressed using XACML at the level of the Fedora web service APIs and down to the object/datastream/dissemination level. Release 2.2 also introduced configurable authentication (Tomcat realms and login modules) with out-of-box support for multiplexing multiple authentication sources, including Tomcat-users and LDAP. Shibboleth can be used alongside Fedora.[3] |

| **Name** | ***Are upgrades designed to be minimally disruptive?*** |
|---|---|
| Detail | *How is updated software supplied to the community? Are update and migration instructions provided?* |
| DSpace | Supported |
| Detail | Upgrade and migration instructions are provided.[4] The nature of the DSpace architecture has led some users to make some local changes to the software that may require some work to migrate to newer versions. |
| Fedora | Supported |
| Detail | Upgrade and migration instructions are provided[5] and Fedora's service architecture should minimise upgrade problems. |

## Developer concerns

Developers are responsible for creating new system functionality. They will want to know:

- Is there good quality developer documentation which explains how to obtain source code, the development conventions for the creation and submission of source code, how to report bugs, etc.?

1 DSpace, ' Backing Up and Restoring a DSpace Instance', *DSpace website.* URL: <http://wiki.dspace.org/index.php/BackupRestore>
2 Fedora, 'Journalling Guide', *Fedora website.* URL: <http://www.fedora.info/download/2.2/userdocs/server/journal/index.html>
3 Fedora, 'Fedora and XACML', *Fedora website.* URL: <http://www.fedora.info/wiki/index.php/Fedora_and_XACML>
4 DSpace, *DSpace System Documentation: Updating a DSpace Installation.* URL: <http://dspace.svn.sourceforge.net/viewvc/*checkout*/dspace/trunk/dspace/docs/update.html>
5 Fedora, 'Fedora Upgrade and Migration Guide', *Fedora website.* URL: <http://www.fedora.info/download/2.2/userdocs/distribution/migration.html>

- What is the quality of the source code and APIs?
- Is there a healthy developer community?

| Name | ***Provides developer guidance*** |
|---|---|
| Detail | *Guidance to get developers started should be available.* |
| DSpace | Supported |
| Detail | Information for developers is available.[1] |
| Fedora | Supported |
| Detail | The documentation is distributed at present. Now the software is to become community driven a developer's guide has been promised. |

| Name | ***Good quality developer documentation*** |
|---|---|
| Detail | *Published APIs, data models, etc., are essential. Explicitly documenting these aids third-party development of decoupled applications.* |
| DSpace | Supported |
| Detail | System documentation is packaged with the source code and Javadocs can be generated from the sourcecode after download. There is also developer documentation for coding with the DSpace system.[2] |
| Fedora | Supported |
| Detail | System documentation[3] and Java docs for v. 2.2 are available.[4] |

| Name | ***Active development community*** |
|---|---|
| Detail | *Both DSpace and Fedora are open source communities with small teams of developers distributed across organisations. Provision of a developer support network is important.* |
| DSpace | Supported |
| Detail | An active and open development community, but the code base is complex. |
| Fedora | Supported |
| Detail | Limited to the development team at Cornell and Virginia until relatively recently. This is beginning to change as the original funding for the Fedora project expired in September 2007 and the project directors have put measures (e.g. governance, working groups, opening of SVN repository on Sourceforge, coding rules, etc.) in place to smooth the transition from funded project to the Fedora Commons open source community. |

## Repository evaluation conclusions

Neither DSpace or Fedora meet the needs of a production repository for personal digital archives as-is. Given the requirement to undertake development work, Fedora is the better choice because of its open architecture, proven scalability, flexible data model and the distribution and interests of its community. Development priorities include an interface which implements a workflow for ingest to the repository, and a dissemination user interface which supports search, retrieval and display for objects in the preservation repository. These interfaces must be designed such that they are usable by archival staff. Pre-requisites for the design of these interfaces is the development of detailed content models which specify how object types (e.g. website, email directory) and formats (html, jpeg, pst, mbox) are to be treated, and what metadata the repository must capture about them.

1 DSpace, 'Guide to Developing with DSpace', *DSpace website.* URL: <http://wiki.dspace.org/index.php/DspaceDeveloping>
2 Richard Jones, *DSpace Developer Documentation*, Version 1.1. URL: <http://wiki.dspace.org/static_files/7/7d/DevelopersDocumentation.pdf>
3 Fedora, 'Documentation', *Fedora website.* URL: <http://www.fedora.info/documentation/>
4 Fedora, 'Overview', *Fedora website.* URL: <http://www.fedora.info/download/2.2/javadocs/index.html>

## Useful resources

DSpace
Main page: <http://www.dspace.org/>
System documentation: <http://www.dspace.org/index.php?option=com_content&task=view&id=151>
Wiki: <http://wiki.dspace.org/index.php/Main_Page>

Fedora
Main page (old): <http://www.fedora.info/>
Main page (current): <http://www.fedora-commons.org>
System documentation for v. 2.2: <http://www.fedora.info/download/2.2/userdocs/>
Wiki: <http://www.fedora.info/wiki/index.php/Main_Page>
Sourceforge: <http://sourceforge.net/projects/fedora-commons>

DSpace Architecture Review Group, *Toward the Next Generation: Recommendations for the next DSpace Architecture* (24 January 2007).
URL: <http://www.dspace.org/index.php?option=com_content&task=view&id=151>
Jantz, Ronald and Giarlo, Michael J., 'Digital Preservation: Architecture and Technology for Trusted Digital Repositories', *D-Lib Magazine*, 11, 6 (June 2005)
URL: <http://www.dlib.org/dlib/june05/jantz/06jantz.html>

Johns Hopkins University, 'A Technology Analysis of Repositories and Services' (2006), *John Hopkins University website.*
URL: <https://wiki.library.jhu.edu/display/RepoAnalysis/>

Lagoze, Carl, et al., 'Fedora An Architecture for Complex Objects and their Relationships', *International Journal on Digital Libraries*, 6, 2 (April 2006).
Preprint availabale at <http://www.arxiv.org/abs/cs.DL/0501012>

Open Access Repositories in New Zealand (OARINZ) project, *Open Access Repositories in New Zealand website.*
URL: <http://www.oarinz.ac.nz/>

Open Society Institute, *A Guide to Institutional Repository Software*, 3rd edition (August 2004).
URL: <http://www.soros.org/openaccess/software/>

Powell, Andy, *Notes about possible technical criteria for evaluating institutional repository (IR) software* (December 2005).
URL: <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/ir-software.pdf>

Staples, Thornton, Wayland, Ross, and Payette, Sandra, 'The Fedora project: An open-source digital object repository management system', *D-Lib Magazine* (April 2003).
URL: <http://www.dlib.org/dlib/april03/staples/04staples.html >