

Digital Repository Models

DPC Briefing Day

July 5, 2006

Sayeed@jhu.edu

Outline of Presentation

- What is a repository?
- Motivation for Johns Hopkins approach
- Lessons learned from AIHT
- Technology Analysis of Repositories and Applications
- Discussion

Motivation

A repository by any other name would be just as...

...confusing?

Different people mean different things when they speak about “repository”, even when they refer to “institutional repository”

A Working Definition

In my view, a university-based institutional repository is a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members. It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation where appropriate, as well as organization and access or distribution.

- Cliff Lynch (<http://www.arl.org/newsltr/226/ir.html>)

What do Libraries care about?

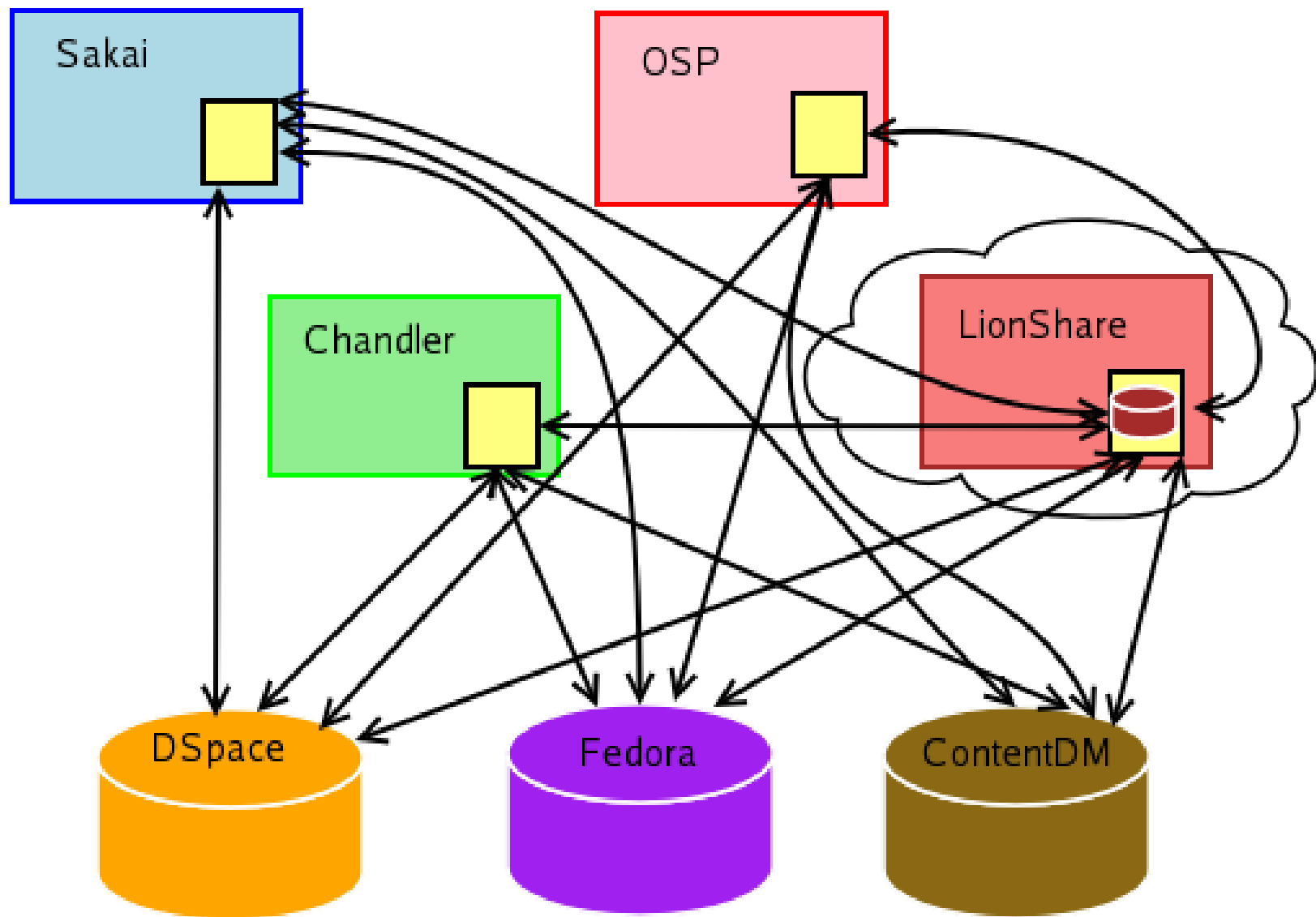
- Libraries spend a great deal of time considering digital access (with good reason)
- However, it's surprising to note the relative lack of emphasis on digital preservation (at least until recently)

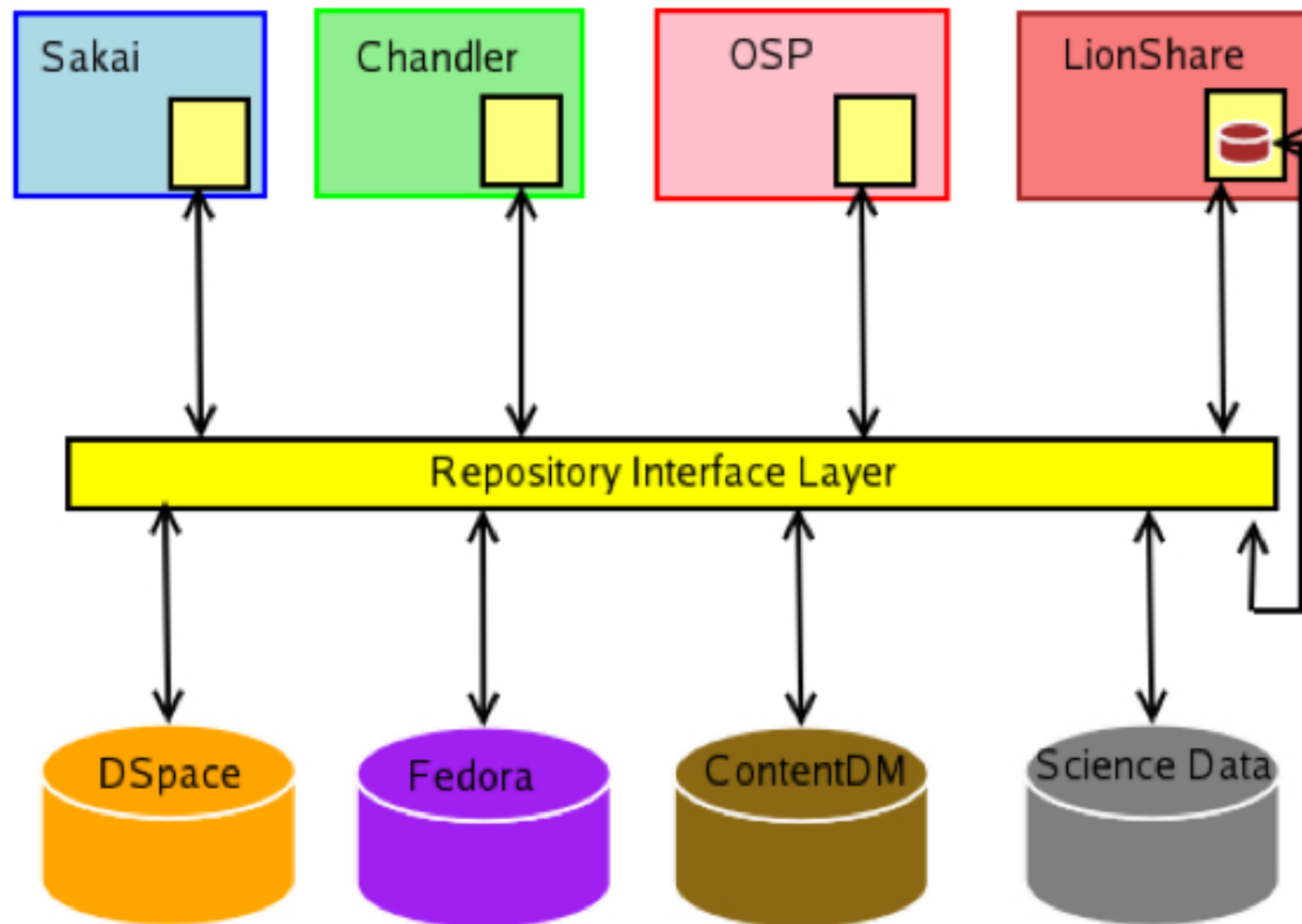
Archive Ingest and Handling Test

- <http://www.dlib.org/dlib/december05/12contents.html>
- Part of the Library of Congress' NDIIPP Program
- Test content consisted of 9/11 Archive from George Mason University
- Approximately 57,000 objects that comprised about 12 gigabytes of content
- Diverse, complex range of content types and metadata
- Each institution was to ingest archive into local environment, share with another institution, and conduct a format migration

What do our customers [or insert your favorite term] need?

- Use (even re-use?), not just discovery
- Seamless access to content and services
(removing silos or stovepipes)
- Collaborative spaces
- Keep it simple (at least from their
perspective)
- Technology providers should embrace
complexity





Analysis

<https://wiki.library.jhu.edu/display/RepoAnalysis/ProjectRepository>

- Dspace
- Fedora
- Digital Commons

- JSR-170
- OKI DR OSID
- ECL

Initial methodology

Scenarios



Use Cases



Repository
Features

Some advice...

- There is no **easy** road to repositories...
- Focus on user requirements, not on specific systems
- Empower technologists to respond with in-house or external resources or even multiple systems
- Consider the repository in a “distributed” manner

Future Work

- Revise, revisit the scenarios, perhaps creating a set of archetypal scenarios
- Work with updated systems
- Use different “lenses” for analysis (e.g., add-on modules, UI-based, application driven access)
- Other systems (e.g., ePrints, ContentDM, VDC, Nesstar, and CDSWare)

Questions or Comments?

- Sayeed@jhu.edu
- <http://www.arl.org/newsltr/226/ir.html>
- <http://www.dlib.org/dlib/december05/12contents.html>
- <https://wiki.library.jhu.edu/display/RepoAnalysis/ProjectRepository>