

The Future of Fedora

Update on Fedora 4

David Wilcox
Andrew Woods
Michael Durbin

DURA SPACE™

“DuraSpace is an independent 501(c)(3) not-for-profit organization providing leadership and innovation for open technologies that promote durable, persistent access to digital data. We collaborate with academic, scientific, cultural, and technology communities by supporting projects and creating services to help ensure that current and future generations have access to our collective digital heritage.”

DuraSpace and Fedora

Leadership and planning

Budgeting - planning, support,
administration

Staffing - hiring, oversight, coordination

Outreach, marketing, advocacy

Fedora 4 project goals

Improved performance

Flexible storage options

Research data management

Linked open data support

Improved platform for developers

Duraspace, Fedora and Me

Community update

Stats for 2014:

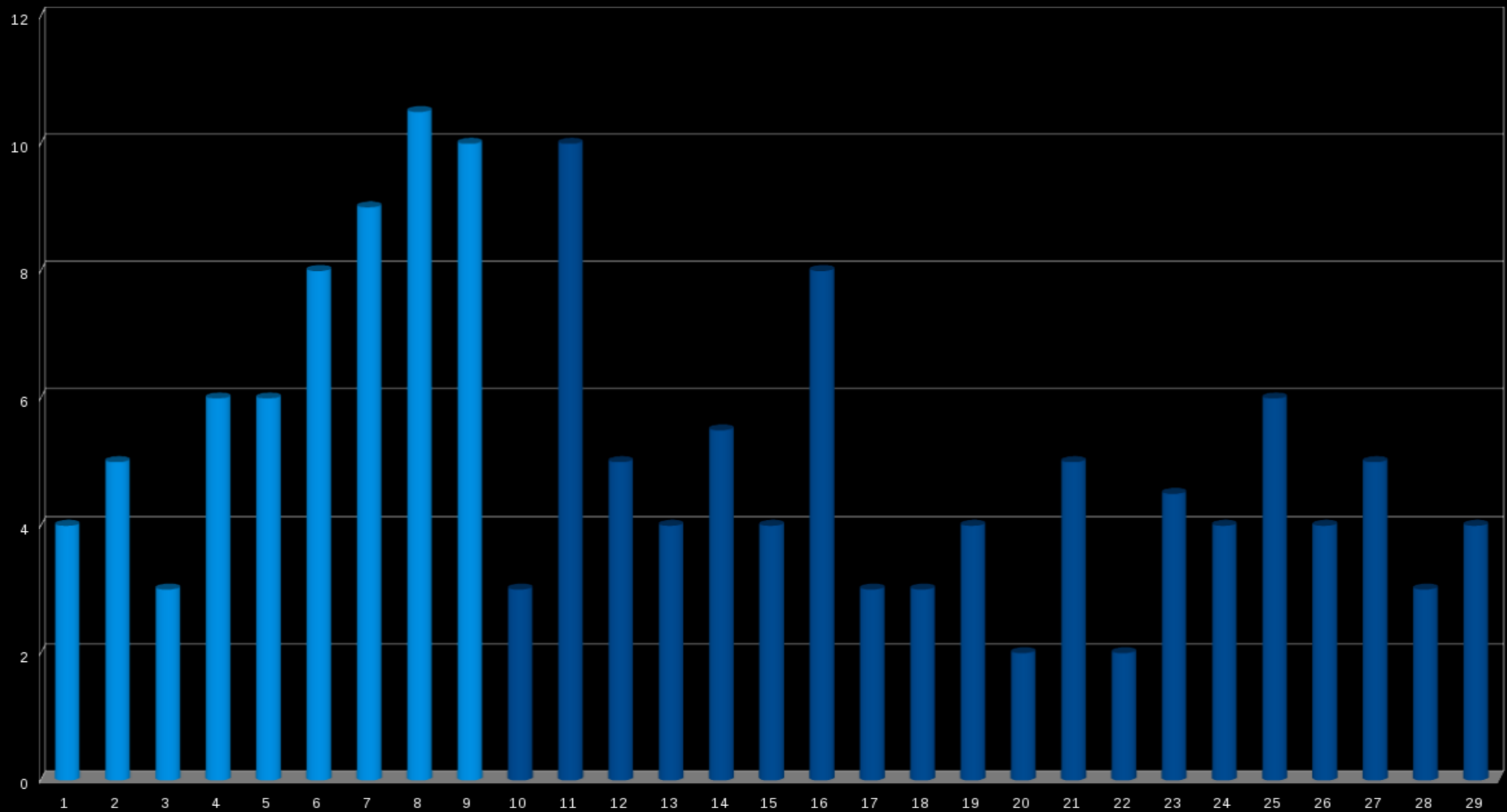
58 Fedora members

17 active developers

18 leadership group members

10 steering group members

Developer Commitments



2014 Fedora Members (58)

Arizona State University Libraries
Brown University Library
Case Western Reserve University Libraries
Charles Darwin University
Colorado Alliance of Research Libraries (CARL)
Columbia University Library
Cornell University
Docuteam GmbH
Durham University
FIZ Karlsruhe
George Washington University
Ghent University Library
Gothenburg University Library
Indiana University
ICPSR
Johns Hopkins University Libraries
La Trobe University
London School of Economics & Political Science
LYRISIS
Macquarie University
National Library of Medicine
National Library of Wales / Llyfrgell Genedlaethol Cymru
National Research Council of Canada
Northeastern University Libraries
Northwestern University Libraries
.Ohio State
Oregon State
Pennsylvania State University
Princeton University
Rutgers University Libraries

Smithsonian Institution, Office of Research Information Services
Stanford University
State and University Library of Denmark
The Art Institute of Chicago
Tufts University
University of Alberta
University of California, Los Angeles
University of California, Santa Barbara
University of Cincinnati
University of Connecticut Libraries
University of Hull
University of Lausanne
University of Manitoba
University of Massachusetts Amherst Libraries
University of New South Wales
University of Notre Dame
University of North Carolina
University of Pittsburgh
University of Oxford
University of Prince Edward Island
University of Rochester Libraries
University of Texas Libraries Austin
University of Toronto
University of Virginia
University of Wisconsin
Uppsala University
Yale University
York University

Building sustainability

We're broadening the funding base

More members at lower funding amounts

Raising the overall level of funding

We've established a governance model

Fedora is governed by a Leadership group and a

Steering group

Governance benefits

***Level of
contribution***

Benefit

\$2,500

Contributors can nominate and elect one contributor at the \$2,500 tier to join the Leadership Group

\$5,000

Contributors can nominate and elect two contributors at the \$5,000 tier to join the Leadership Group

\$10,000

Contributors can nominate and elect four contributors at the \$10,000 tier to join the Leadership Group

\$20,000

Contributors guaranteed one seat on the Leadership Group

***0.5 FTE
Developer***

Contributors guaranteed one seat on the Leadership Group

Training

- 3 one-day workshops in October 2014
- Planning a three-day camp for 2015
- Need to “train trainers”
- Make the training portable

Fedora 4 roadmap

Complete Beta pilots by end of October
Release Fedora 4.0-Production in 2014

Fedora 4.1 focus:

Support Fedora 3.x to 4.x migrations

Polish Fedora 4.0 edges

Getting to production 4.0

Acceptance testing

Beta pilots (4)

Technical Working Group

Performance

Preservation-worthiness

Rest API

Clustering

Community feedback

Developer commitments

What's in Fedora 4.0?

Admin console

Fixity Checks

Backup and Restore

Export and Import

Built-In SPARQL Search

External Search and Triplestore

Content modeling

Authorization (Role-based and XACML)

Durable Storage

File System Federation

Transactions

Versioning

Clustering for High Availability

Scale (large files and many files)

Linked Data Platform

Community participation

Acceptance tests

< 100% complete

Many features for which no acceptance tests
have been performed

Linked Data

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)
4. Include links to other URIs. so that they can discover more things.

Tim Berners-Lee, <http://www.w3.org/DesignIssues/LinkedData.html>

Linked Data Platform

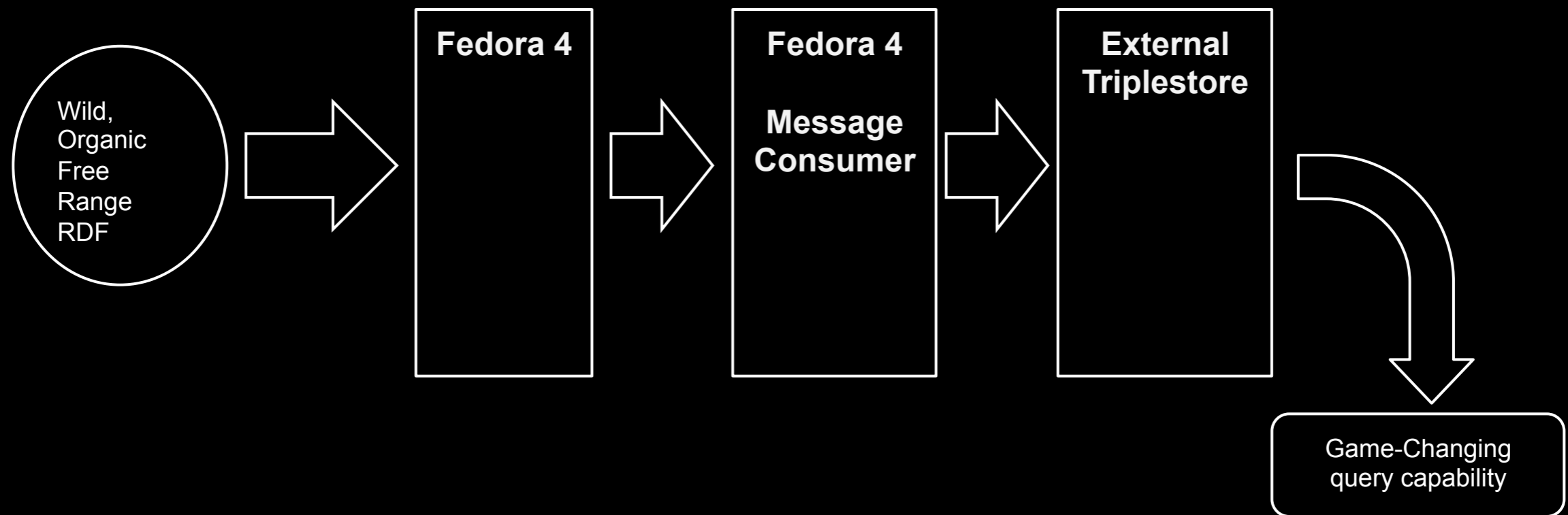
“set of best practices and simple approach for a read-write Linked Data architecture, based on HTTP access to web resources that describe their state using the RDF data model.”

W3C Last Call Working Draft 16 September 2014

<http://www.w3.org/TR/ldp/>

Fedora 4 is an implementation of a LDP
Compatibility Test Kit in Fedora 4 build process
Ensures current conformity
Stays current with standards

Linked Data Platform



API and Implementation

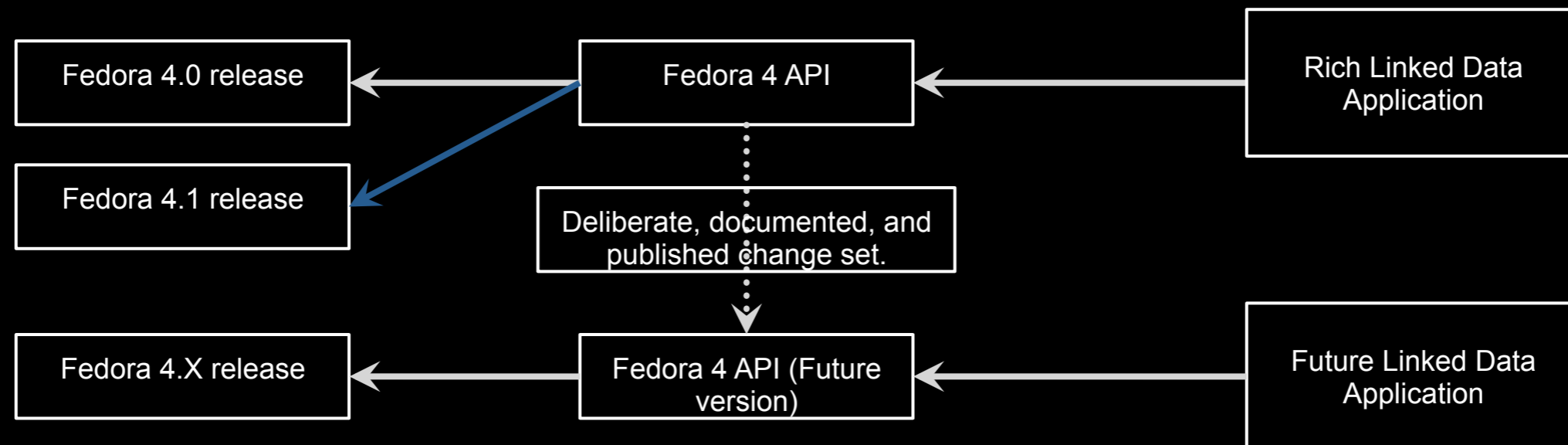
TWG recommendation

Version Fedora API separately from Implementation

Core API = LDP + Versioning + Transactions

Extension APIs

- Optional features
- Community-developed features



Migration Considerations

Features may be similar, but underlying organization and model have changed.

Migration Considerations

Fedora 3

XML Objects + Datastreams

Flat pool of objects

PIDs

Fedora 4

RDF Properties +
Datastreams

Hierarchy of objects

Path and URI

Migration Considerations

Versions

Can't be retained natively.

Must be iterated over and loaded.

Version metadata must be explicitly included.

Migration Considerations

Content Models

- How are content models used in your Fedora 3 repository?
- Do they have any logic built into them, or is that handled at a higher application level (e.g. Islandora, Hydra)?

Migration Considerations

Disseminators

- Does your Fedora 3 repository make use of disseminators?

- What are they used for? XSL transforms?

Derivatives?

- How can we support the existing disseminator use cases in Fedora 4 without re-creating disseminators?

Migration Considerations

Enhancements

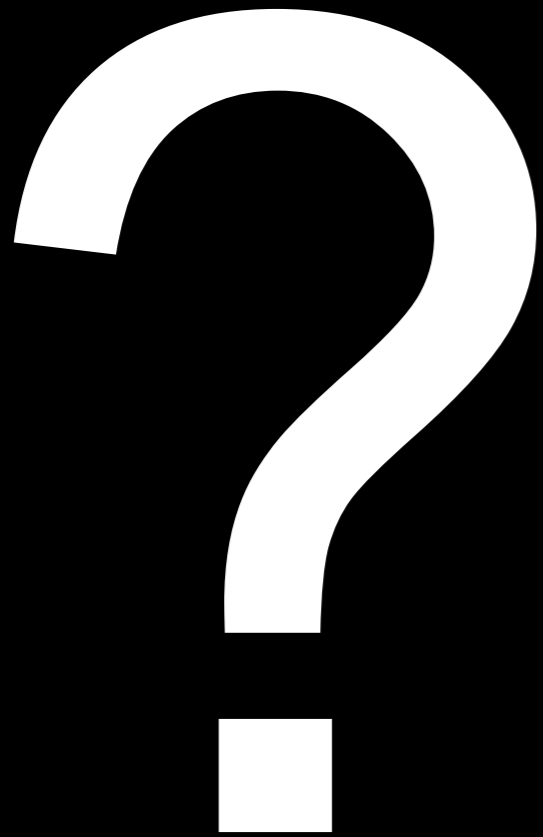
- Take advantage of properties
- Embrace (or ignore) the hierarchy
- Query RDF in powerful new ways
- Build applications over Fedora API with confidence

Community participation

Migration Pilots for 4.1 release

How we can identify and develop strategies for edge case migrations.

Questions



Useful links

Fedora 4 wiki

<https://wiki.duraspace.org/display/FF>

Fedora mailing lists

<https://wiki.duraspace.org/display/FF/Mailing+Lists+etc>

Contact

dwilcox@duraspace.org

awoods@duraspace.org

miked@virginia.edu