

2016-06-12 - OR2016 Fedora Leadership Group Meeting

Time/Place

- Time: 10am - 4pm IST (UTC+1)
- Location: Pearse Room, [Maldron Hotel Pearse Street](#) ([map](#))
- Dial-in Number: (712) 775-7035
 - Participant Code: 479307#
 - International numbers: [Conference Call Information](#)

Attendees

- [Julie Allinson](#)
- [Aaron Choate](#)
- [Tom Cramer](#)
- [Jon Dunn](#)
- [Karen Estlund](#)
- [Declan Fleming](#)
- [Jonathan Markow](#)
- [Glen Robson](#)
- [Robin Lindley Ruggaber](#)
- [A. Soroka](#)
- [Evviva Weinraub](#)
- [David Wilcox](#)
- [Andrew Woods](#)
- [Neil Jefferies](#)

Agenda

Topic	Time
Technical priorities and required resources (reference: 2016 project plan) • What needs to get done this year? 1. Functionality to support standard import/export 2. Migrations a. ModeShape 5 b. Towards JDBC, away from LevelDB 3. API specification • How do we get it done? 1. Hackathon on performance and scalability 2. Scheduled code sprints • Funding developer contributions	45 minutes
Fedora in 5 years • What is the steady state of Fedora? • Assuming stable software that is easy to get in and out of, should we allow engagement and funding to decline?	30 minutes
Fedora 4 Adoption • Why has adoption been slow? • Are concerns around stability and scalability/performance hindering adoption?	20 minutes
Learning from other projects • Hydra-in-a-box uses 1 week sprints and posts weekly YouTube videos. Should we adopt similar practices? • Are there other projects we can learn from?	10 minutes
Fedora Training • Assessment of training activities and initiatives to date ◦ 13 workshops + Fedora Camp in 2015 ◦ 12 workshops + 2 Fedora Camps in 2016 (planned) ◦ Trained instructors and re-usable materials • Future plans for training	20 minutes

Post-merger thoughts	15 minutes
- What did we (Fedora) hope to gain from the merger (e.g. RSP relationships)? - How should we achieve these goals now that the merger is dissolved?	
Fedora integrations	25 minutes
VIVO, SHARE, OSF	
European, government, museum membership	20 minutes
Engaging with attendees at OR2016	10 minutes

Minutes

[Google Doc for taking notes](#)

Technical priorities and required resources (reference: [2016 project plan](#))

- What needs to get done this year?
 - Functionality to support standard import/export
 - Migrations
 - [ModeShape 5](#)
 - [Towards JDBC, away from LevelDB](#)
 - [API specification](#)
- How do we get it done?
 - Hackathon on performance and scalability
 - Scheduled code sprints
- Funding developer contributions

2 big things floating right now:

- Support for Migrations. Proposed Focus for 2016 = Tooling around import/export
 - ModeShape has undergone a significant change which requires a migration to go from ModeShape 4 to ModeShape 5.
 - Default DB (LevelDB) has issues; recommending an alternative DB.
 - Principles for Fedora:
 - Easy to get stuff in a standardized way
 - Easy to get stuff out in a standardized way.
 - A focus for 2016 should be tooling around import/export.
- API specification
 - Lots of benefits and goodness for this.

ModeShape questions / status.

- Performance & scalability. Now have a WG. Meeting 1x per month.
 - Storage & storage transparency.
 - Federation/Projection.
 - API-X
 - ModeShape viability
- Current reference implementation relies hugely on ModeShape; any shifts in ModeShape have massive effects on Fedora. But it's not custom code that we have to maintain; by adopting ModeShape, we got to move very quickly.
 - Standard import/export. RDF for objects & binaries for files. Iterate over repo and make get requests.
 - Can Fedora provide a Bagging layer that could serve as a middle layer between Fedora and remote preservation services.
 - Some sites (IU) have large binaries not managed by Fedora; would need to accommodate this.
 - Would export for preservation be the same / highly overlapping with export for migration?
 - NU & UCSD have IMLS proposal in the works to investigate technical mechanisms for tracking content / location of bags over long term.
 - Need to gather some use cases.

API Work

Lots of organic engagement. But there are gaps.

<https://wiki.duraspace.org/display/FEDORA/FEDORA+Specification> (drafts)

5 RESTful services + Messaging. Services aligned with web standards.

- CRUD -> LDP
- Authz -> WebAC
- [Versioning](#) -> Memento
- Batch atomic operations (transactions) -> no known standard to align with
- Fixity -> LDP (sort of)

These started as Google Docs, rough & quick first drafts have now taken place for all of them. Versioning is now getting more formalized, now has [W3C-style spec doc](#). What can we do to drive progress on the remaining 5? Do some sprints? Have an event? What are the next steps?

Need to add some structure here. What is the status of each thing? Who is the owner? What are the next steps? Is the API spec work aligned with local work?

What does HyBox need?

- WebAC
- Multitenancy (but not API-related)
- Storage (S3, plugin work)

Take away points;

1. BagIt / Import / Export
2. Make status of API spec generation more clear (identify people willing/able to shepherd specs)
3. Maybe some WebAC work (?)
4. Need to identify alignment of this work with what people / institutions are already working on

Getting Developer Contributions

There are a handful of folks who are working on Fedora. Classic Fedora hobbyists.

We've been talking about this for years. Ideas/comments

- Hire developers to do contract work to fix this.
- Get a funded project (like HyBox)
- Get more Fedora 4 implementations; then institutions might be willing to buy time from a contractor / central agencies
- "I'm out of developers"; don't have anyone to assign this to
- Migration isn't a priority for me right now; the repo works, and my focus is on ingest streams.
- Access is easier to demonstrate, easier to get resources for
- Could import/export be a more visible focal point for Fedora?
- Fedora's ties to other projects / its place in the stack can be an asset:
 - Energy and attention that comes from Hydra, Islandora, PCDM, etc. has helped generate significant attention to Fedora when activity aligns (WebAC, e.g.)
 - Ties to OSF, Vivo, etc. could have similar benefit
 - A Fedora strategy should be to orchestrate the alignment of interest and standards from other communities (Hydra & Islandora, e.g.)
- What if we were bold, and we're going to say that we're not going to do "independent fedora" any more?
 - That would be problematic.
 - Java vs. PHP / Ruby (this problem is dissolved by a complete API spec)
 - Would also cut out a lot of RYO (roll your own) shops who don't use Islandora or Hydra
- Keep interest / edge around integrations. Value becomes visible. Tie Fedora into all sorts of different sites -- Goobi, OSF, Vivo, etc.

Fedora in 5 years

- A well codified, eminently respectable API set that defines awesome digital resource management
- Also useful software.
- And would be nice to have multiple implementations.
- Note:
 - "Reference implementation" is a risky marketing term. Seems [wifty](#). Could we use "standard implementation" or "supported implementation" instead?
 - Lots of discussion on this. Really, don't call it a Reference Implementation in public. Really. People are feeling good about this.
 - Also all the attention on the API spec may make people leery of doing a migration now until things settle down.
 - Maybe this could be recast as a "roadmap"?
- A whole new set of functionalities will be available/important in 5 years. With really good APIs, we shouldn't have as much of a disruption between 3 & 4.
- In 5 years, Fedora should be integral to emerging data management.
 - Let's do mega-intensive work now on Vivo, SHARE, OSF, CRIS, ORCID, etc.
- In 5 years, when we evaluate new tools, we want "do you integrate with Fedora?" to be a given, and important criteria
 - "Do you do 'Fedora'?"
- Will we still be on Java in 5 years?
 - Hard to have a vibrant OSS dev community around Java.
 - Except actually there are way more Java people.
 - This seems highly regional.
- A dynamic & expert community where people gather to solve problems in concert
 - Barrier to Fedora is high -- organization & also technical
 - Fewer opportunities to meet face to face, build energy

Fedora 4 Adoption

Happening, but not necessarily quickly.

Why isn't this happening more quickly? What are the barriers?

More constructively: How do we think Fedora 4 adoption is going? Is it on track or ???

Why?

- Not moving due to focus on Spotlight & access. Waiting for PCDM to robustify. Waiting for CurationConcerns to solidify.
- Rearchitecting whole repo ecosystem; need to figure out how to consolidate everything.
- Things changing rapidly
- "Do you run on Fedora 4" now a checkbox for many new institutions; adding incentive to Avalon's migration. Latest strategy is to move to Fedora 4 ASAP.
- Upgrading things is a major operational concern. Needs planning, funding, project window.
- How do we know if PCDM works; need more implementers.
- Fedora 3 is working really well; moving to a new repo is going to take a huge amount of time--both operationally, and also for data modelling.
- Want to see someone doing it first; someone at scale
- It takes ~5 years for large ships to turn, and move to adopt a new infrastructure

Let's look at these questions as opportunities for institutions that Fedora 4 is serving as a catalyst to lever them into the future...

- New data models
- Embracing RDF
- Consolidating digital library infrastructure

Huge consulting opportunity for helping people think about moving to linked data.

Make Declan do a webinar. Ha.

Are we moving fast enough with adoption?

- Probably, if you think about number of F3 new adopters (few to none)
- Greenfield sites are picking up F4
- Where is Fedora leaking potential adopters? Invenio or Figshare or...?

Learning from other projects

- HyBox does webcast demos at end of each sprint
- Get people other than developers into the community; archivists, data curators
 - How would they work with Fedora? No UI.
 - Talk about content and pipelines.
- Tie into other events; do a Fedora day (or something) at HydraConnect
 - Pursue getting agreement on where in the stack things might happen. Islandora/Hydra vs. Fedora.
- Do some storytelling
- Fedora events good
 - Can we make OR better for Fedora? (But not a lot of devs go)
 - Regional events are good. Can we get more?
 - Training events are good. 2-3 a year.

Training & Events

- 12-13 events/workshops per year
- Training events are good. Get people up to speed.
- But also consider marketing Fedora with a less technical approach .
 - "Curate your data (using Fedora)"
 - "Demonstrating Value to Upper Management (using Fedora)"
 - "Leverage linked data (using Fedora)"
- Have a directory of people who can/will speak about Fedora
 - Slot people on Fedora at these conferences
- Get other people to market for us -- have Fedora listed on their pages as an official integration
- Let's also put up a list of products that Fedora integrates with
- Put up a "Powered by Fedora" badge
- Surprise! Fedora! (Visibility for the many unknown Fedora sites)
- Have a European training event.

The Merger

- Why is it so hard to find RSPs for Fedora?
 - Hydra and Islandora seem to be more attractive for vendors to support.
 - Does Fedora require too much initial investment for a vendor?
- Note taker got distracted here.
- Main perceived benefits of merger:
 - Reach to more institutions (lots of Lyrasis members)
 - Different institutions (small colleges, e.g.)
 - Sales, mktg, training channels of Lyrasis
 - Full lifecycle support: digitization, digital preservation, DuraSpace projects, ArchivesSpace, CollectionsSpace
 - Sibling projects: combined
 - Mass: 70 or 80 staff members could give more heft for exploratory / R&D / new ventures

Integrations

- ArchiveMatica & Fedora 4
- Research data library
- SHARE
- DPN / DuraCloud Vault / APTrust / etc.

Whither Fedora in Europe?

- Few pan-European events to attach to
- Focus on helping organize regional user group meetings
 - Set aside funds to help people travel to these events?

Actions