2012-02-23 - Special Topic - Fedora, Web, and SemWeb Architecture

Time/Place

This meeting is a hybrid teleconference and IRC chat. Anyone is welcome to join...here's the info:

- **Time:** 9am Eastern Standard Time US (UTC-5)
- **Voice**
  - Via Skype (Dial from anywhere):
    - In Skype, select Add a New Contact
    - In the search box, type `freecallphonehd.8053991200`, hit Enter, and add the contact
    - Call the contact
    - When prompted, bring up the Skype keypad and enter the following by clicking with your mouse – not using your keyboard: 341861#
  - Via Phone (US Number):
    - +1(805)399-1200, Access Code: 341861
- **IRC**
  - Join the #duraspace chat room via Freenode Web IRC (enter a unique nick)
  - Or point your IRC client to #duraspace on irc.freenode.net

Agenda

Today's Special Topic meeting:

**Fedora, Web, and Semantic Web Architecture:**

1. What are the major similarities/differences between Fedora and Web Architecture? The object/resource model. Identifiers, Formats, And Protocols.
2. Where does Fedora's architecture intersect with or diverge from Semweb architecture?
3. Should Memento support be added directly to Fedora's REST API, or considered as a higher-level service (extension) on top of it?

Notes

Attendees:

- Chris, Adam, Dan, Ed, Frank, Steve, Eddie

Notes:

Chris: Start with how Fedora fits/doesn't fit into web architecture. Dan has called Fedora a "Resource Store" vs a "Content Store". How does this concept of resource differ from web arch's concept?

Adam: CMA and dissemination architecture. A key distinction is that Fedora's dynamic representations are highly specified, and the specifications themselves are resources.

Chris: So is Fedora just filling in details, or does it actually go in a different direction from web arch?

Adam: Fedora's object model is different. It's disseminations also look an awful lot like member functions in OOP.

Steve: It's also about fundamental responsibilities. Fedora's about curation and dissemination. Fedora makes a distinction about identifying the representation as a thing to be preserved!

Dan: I believe Adam said "An archive needs to care how a representation is created."

Adam: Sounds good, I'll take credit.

Dan: Hourglass design: push complexity to the edge and have simple stuff (verbs, object models) in the "waist."

Adam: Fedora's dynamic disseminations haven't got a lot of uptake.

Me: Complexity, approachability.

Dan: Maturity.

Adam: History...things have changed with Fedora. More emphasis now on the graph of objects and relationships. Change ok.

Dan: Also CMA. And people haven't experimented with alternate ideas for content modeling as much as we'd hoped.

Adam: ECM has. Others haven't in part because it's too hard.

Chris: What about the REST api? Where do we fight the web architecture & REST?
Adam: Orchestration tools/APIs tend to be RPC-based, not resource-based. Usually you talk about operations with SOAP. A REST API doesn't fit as well there.

Dan: Gigantic strength in seeing Fedora as middleware and not as a repository. Having different interfaces is important.

Chris: REST doesn't define Fedora, nor does SOAP. Still feel both are important. Advantages of REST API over SOAP? Big one in my mind: ease of integration with existing languages/tools. Success with SOAP tool interop has historically been mixed, though it's getting better.

Adam: (agree on SOAP tool interop)

Frank: Also if you adhere to REST, you get some things for free. Caching on top of Fedora via Apache. Can't do with SOAP API.

Dan/Chris: Method-for-method equivalence of REST api and SOAP is not necessary. But overall capabilities should be comparable.

Dan: Fitting Fedora into a grand eventing architecture?

Chris: Steve has suggested that more generic, resource-oriented events would be good to publish from Fedora. Advantage is integration with more generic (non-Fedora-specific) tools. Today we have Fedora-specific event-based messages coming out. Need for both approaches?

Steve/Others: Yes. Seems both have their place.

Adam: On any given interface to Fedora, it seems like we want to consider a Fedora-event-oriented exposure and a more generic resource-oriented exposure.

Chris: Now, where do things seem weird with semweb architecture and Fedora?

Frank: Having ids for all Fedora entities makes sense. Referencing via URI.

Chris: Use of the info: scheme seems to be one place we fight some accepted semweb guidelines (linked data says http-only). The info-uri effort has been discontinued, but thankfully, it wasn't designed with resolution in mind so Fedora info: URIs will continue to be useful.

Adam: A fundamental difference: In Fedora, the only place where I can make an RDF assertion is inside the object that the assertion is persisted.

Chris: I'd qualify that -- you can make whatever RDF assertions you want about Fedora object, but if you want them indexed by today's Fedora RI we have certain restrictions on that.

Adam: Steve presentation at OR: Pattern we've seen with Fedora is that public ids have been exposed through some other system, not Fedora. Fedora has internal ids.

Steve: Was a "what if": Fedora as the curator of the public ids? Valid to have an internal representation and a different public exposure of that model, and a transformation between those things.

Adam: Community seems to prefer not to use info: uri as public identifier.

Dan: Nuanced conversation. Asking to potentially build some new core ideas. Web arch. allows for private resolver systems. Was thinking of ideas that came out of SOA community, notion of Fedora ecosystem providing a private resolver system. A system for being an authority source for identifiers.

Dan: If we provided technology or specifications for the IFAPs (Interfaces, Formats, And Protocols) for federating Fedora, we can facilitate communities who are actually putting federations together.

Steve: Fedora has a role as registry repository.

Dan: Can act as resolver system, but can be made better.

Adam/Chris: Should Fedora's REST api endpoints be considered Durable or Linked Data Friendly?

Dan/Chris/Adam: We should be very clear about that.

Dan: Also Fedora needs to re-address notion of versioning and migration. Fundamental pattern of a bolt-on in Fedora to handle migration. Might not be all of memento.

Chris: If REST API is not intended to provide durable ids for things, memento doesn't belong there.

Dan: We often leave out the Federation of repos.

Steve: Believe it's very important that Fedora have an internal id.

Chris: Using URIs for internal ids works well today because we have standards (rdf) to relate them.

Frank: What about UUIDs? Easy to generate.

Chris: Also URI-compatible. It's convenient in any case for Fedora to continue allowing people to mint ids outside the repository.

Adam: Anti-pattern we see a lot of: putting semantics in ids. But human nature.

Steve: From a long term preservation perspective, you'd want to isolate yourself from current trends, standards.

Chris: Riding waves is good, but we're not defined by them.

See also: notes on IRC.
Action Items

TBD