

Hello

- Probably most accurate to say this is my current take on what's happening with RDF and digital library technology here at IU with our use of Fedora's digital repository, but I do think this is going to effect things in libraries moving forward

NEXT SLIDE

What makes a good model?

- We're trying to figure out what makes a good metadata model
- Is it how you walk
- The fierce-ness in your eyes
- Or something else that really just defies explanation
- And the only other thing I know about that show is something about always extending your neck - so I'm sticking my neck out here and we'll see how it goes

NEXT SLIDE

What makes a good model? (cont'd)

- RDF is kind of the latest in metadata models
- I'm saying that because it's not in common use yet, particularly as part of how systems function
 - mostly used as way to interchange or offer data for sharing, for others to use
- Fedora Commons Digital Repository software that we use here in Bloomington is making RDF core to how it functions - that's also a bit of a neck-extending leap, at least to those of us still pretty new to doing things in RDF

NEXT SLIDE

What isn't RDF

- Things I've learned in teasing out what is what with RDF (in some ways, saying "RDF" is like saying "Web 2.0" a few years ago - it's not clear what you are talking about)
- Resource Description Framework
- It can contain linked data or XML but it is not equivalent
- It is not a vocabulary of terms or even a syntax, although there is something called an RDF Schema that does define elements you can use when applying the Resource Description Framework

NEXT SLIDE

What is RDF?

-
- Going to show some real life examples soon that will hopefully give a better sense of what I'm talking about when I say RDF
- RDF Primer - Takes you through all of the basics of the framework with terminology and examples
- Rdfabout - talks about using RDF as a way to store data, which is interesting to me in terms of Fedora 4's use of RDF - not just about sharing data with others but using RDF internally (instead of a database of tables or set of XML files)
- RDF for Librarians - Jenn Riley's brown bag talk from 2010 does a great job of introducing RDF and discusses how things differ between the library world and the RDF world
 - libraries are record-centric and everything about that thing is described in a single place (MARC record, MODS record, static XML files)
 - RDF is more distributed and can incorporate information from multiple sources to describe a thing; additionally, that thing can be related to other things wherever those other things are located
 - I also linked her handout because even though the

Scholarworks recording says it will be linked at the end of the presentation, that slide never showed up and the PPT isn't on Scholarworks and the DLP site isn't there anymore, the point is that the handout still exists

NEXT SLIDE

IU's Digital Collections, ca. 2005 - present

- Digital objects in Fedora often have some combination of files (the actual digital thing, possibly represented in different versions, but a single thing somehow)
- Along with the digital thing are other files also attached to that object (called datastreams)
- There's always a Fedora DC data stream, which gives an identifier and sometimes title info in Dublin Core and nothing else, it's required to have a Fedora object but it's not something we do much with
- Other data streams contain descriptive metadata (MODS file), and, if necessary, RELS-EXT (Relationships External) to connect this object to other objects in Fedora
- can also have other data streams of information, such as Excel spreadsheets, text files, other XML files that are not in any particular standard
- All of this together makes the thing we call a Fedora object, at least up through Fedora 3

NEXT SLIDE

Sample Digital Object in Fedora 3

- Wylie House image
- in Fedora - multiple derivative images, metadata, required Fedora DC, RELS-EXT connecting this object to the Wylie House collection and declaring content models (notice

RDF/XML)

NEXT SLIDE

IU's Use of RDF: Avalon Example

- Political Papers collection has recordings in MCO with multiple sections
- Entire recording is a Fedora object (avalon:2360 is the identifier) and each of these sections, such as May 4, 1971's recording, is also an object in Fedora
- Connecting sections of a single audio/video item together using those relationships defined via RDF
- one explicitly-described thing is related to another explicitly-described thing

NEXT SLIDE

IU's Use of RDF: Finding Aid Example

- More complicated but mostly because it breaks down into more parts
- Finding aid has folder of items with one item that is 199 pages and each page has an image
- Those are all represented as objects in Fedora
- RELS-EXT is relating these things together using RDF, a graphical representation like this is possible, and these relationships can be queried

NEXT SLIDE

- Managing and maintaining order and sequence is done using METS at the "folder of items" level
- [show METS file]
<http://fedora.dlib.indiana.edu/fedora/objects/iudl%3A135409>

NEXT SLIDE

- New Fedora 4 shows things differently - properties are more directly available on the object instead of as attached datastream XML files
- This first screen shows a Fedora object that is a collection or Container
- It contains 2 children
- Julie Talks to Squirrels is one child object
 - properties on this child object include the title and the subject, both in Dublin Core
 - all of these properties are being stored in RDF
 - title is example of a literal, subject is a URI for the subject “squirrels” from LOC - example of using linked data
 - the image is a binary file that can be downloaded for this object
- Cliff on a Cliff is the other child object of the Image Collection container
 - 2 subjects on this, the LOC linked data subject of “cliffs” and what I am proposing as the identifier for Cliff Ingham, his Github profile
 - over on the right, the Update Properties allows you to immediately make changes to these properties using SPARQL (SPARQL Protocol and RDF Query Language according to Wikipedia), a query language specifically for querying RDF triple patterns

NEXT SLIDE

- This image is from the Duraspace wiki documentation for Fedora 4 and shows the full Update Properties box - these are all of the namespace prefixes that are available as part of a Fedora object

- So far from what I can tell, it's possible to add other namespaces and properties from those namespaces using this SPARQL dialog box, Fedora 4 seems to be able to handle it

NEXT SLIDE

- For example, I tried a SPARQL update on this object to change the subjects from being expressed as Dublin Core subject to being expressed as MODS RDF, which is subjectTopic - this worked without any problems in Fedora
- This is actually not how a topical subject is expressed in MODS RDF according to their ontology
- I didn't have a chance to work out how to properly express that within Fedora as an object property while still showing subjects as they should be in MODS - a subject further defined as a topical subject or name subject or geographic subject
- Main point here is that Fedora can manage mixing and combining of namespaces to define whatever you are needing to define about the object in RDF
- Interest in using MODS as we often source our descriptive info from library catalog records; MODS gives a little more definition and context than is available in DC

NEXT SLIDE

- This seeming capability in Fedora 4 is making us consider things like how we store our descriptive and structural metadata
- Descriptive metadata in RDF might give us better control over managing that information in Fedora, making editing easier and quicker
- I don't even really know what I'm doing yet and that SPARQL query dialog box worked for me to make

changes to the object's description

- Structural metadata in RDF is definitely possible using ontologies from different namespaces
- There are also ontologies that allow you to express hierarchy (what is part of what) and sequence (the order of those whats); seems to be more complex when trying to include particular order
- It's expressible using RDF, you can model it in RDF, but querying that information or making use of it might be a different story - I'm still learning

NEXT SLIDE

What is RDF?

- Precisely defined as a family of W3C specifications, it is being used as a general method for conceptual description or modeling of information using different format expressions - from Wikipedia
- It is, with respect to how it's shaping up in Fedora 4 and how we seem to be viewing its capabilities for digital library technology, another way of storing data that we can use
 - allows relationships to be defined in different ways than within hierarchical XML or even relational databases
 - I see it as more similar to relational databases without all of the overhead and server requirements of a database system

NEXT SLIDE

RDF in a lunchbox

- Maybe that's what all the lunch boxes are for

NEXT SLIDE

Questions? Comments!