## **Introduction to Fedora**

Overview, examples, and core features

David Wilcox, DuraSpace @d\_wilcox



### **Learning Outcomes**

Understand the purpose of a Fedora repository

Learn what Fedora can do for you

Understand the key capabilities of the software

Our community is part of an interconnected, worldwide, scholarly ecosystem.

Imagine the WORLD as a DIGITAL RESOURCE ... 0 #= Content X = Users >= Stakeholders ♦ = Payment A= Legal frameworks \* = Technology platforms k infrastructure = links to other n digital resources ECOSYSTEM

Source: The Digital Ecosystem in the Balanced Value Impact Model (illustrated by Alice Maggs); <a href="http://simon-tanner.blogspot.com/2014/09/do-you-understand-your-digital-ecosystem.html">http://simon-tanner.blogspot.com/2014/09/do-you-understand-your-digital-ecosystem.html</a>



#### **DuraSpace open source projects**





#### **DuraSpace** services

DURACLOUD DSPACEDIRECT

@rchivesDIRECT



#### Flexible Extensible Durable Object Repository Architecture

#### Concept

Implementation

Community

#### Fedora...

Stores, preserves, and provides access to digital objects

Supports flexible and complex content models for objects

Supports complex semantic relationships between objects inside and outside the repository using RDF

Supports millions of objects, both large and small

Interoperates with other applications and services

## Why use Fedora?

Fedora is **flexible**: it can handle both simple and complex use cases

Content in Fedora is durable: Fedora supports long-term preservation

Fedora powers successful digital repository and DAM applications

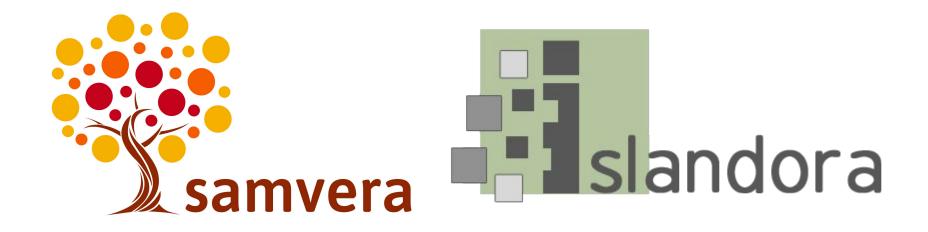
Fedora is standards-based

Fedora is backed by a thriving community

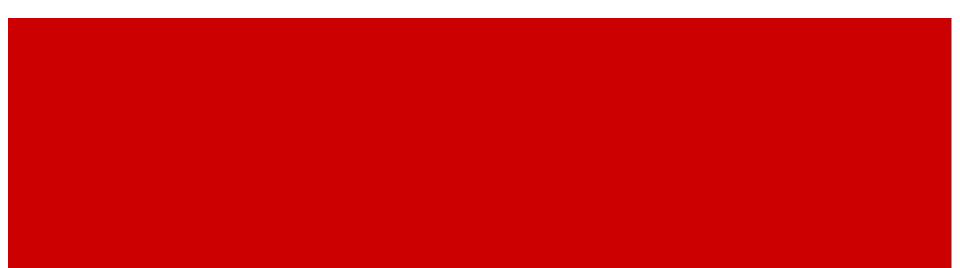
#### Fedora Front-Ends

Fedora is middleware

You can build a custom framework, or join a broader community:



## **Fedora in Production**



## **Institutional Repository**



Featured Works Recently Uploaded

Featured Works



The George Washington...

GWU Pre-Law Student Association

Pre-Law, Student works, Undergraduate works

Privileged Place: Race,...

Squires, Gregory Other

Public policy

Featured Researcher

View other featured researchers

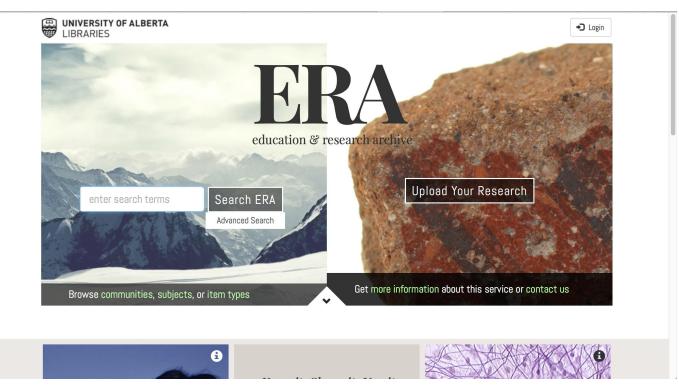


The Project on Middle East Political Science, POMEPS, is a collaborative network designed to enhance the broader Middle East's political science field and its engagement with the broader academic discipline, public policy, and the public sphere. POMEPS seeks to shape innovative new approaches to the political science of the region, support the field's publication in leading academic journals, and ultimately build the number of tenured political scientists specializing on the Middle East. It also aims to increase their contribution to the public foreign policy debate and to the policy-making process, in order to allow their expertise to have more of an impact on vital decisions about the Middle East.

**Browse the POMEPS Collection** 

#### https://scholarspace.library.gwu.edu/

#### **Research Data**



#### https://era.library.ualberta.ca/

#### Manuscripts



York's Archbishops Registers Revealed provides free access to over 20,000 images of Registers produced by the Archbishops of York, 1225-1650, in addition to a growing searchable index of names, subjects, places and organisations. The registers are a valuable, and in many cases, unexploited source for ecclesiastical, political, social, local and family history - covering periods of war, famine, political strife and religious reformation in the Archdiocese of York and the wider Northern Province.

You can browse images using the options below, or search the 2688 indexed entries via the search interface. Further information, guidance and supporting material relating to the registers will be added to the site as time goes on.

Pilot imaging and development, as well as the indexing of the register of Archbishop Neville https://archbishopsregisters.york.ac.uk

#### **Archives and Special Collections**

#### BARNARD

#### **Barnard Digital Collections**

Historical materials from t<mark>he Barnar</mark>d Archives and Special Collections, at your fingertips. Learn more <u>about</u> the digital collections.

Search...

Browse all materials Browse collections Newspaper Yearbook Photos Exhibits

http://digitalcollections.barnard.edu

## **Basic Concepts**

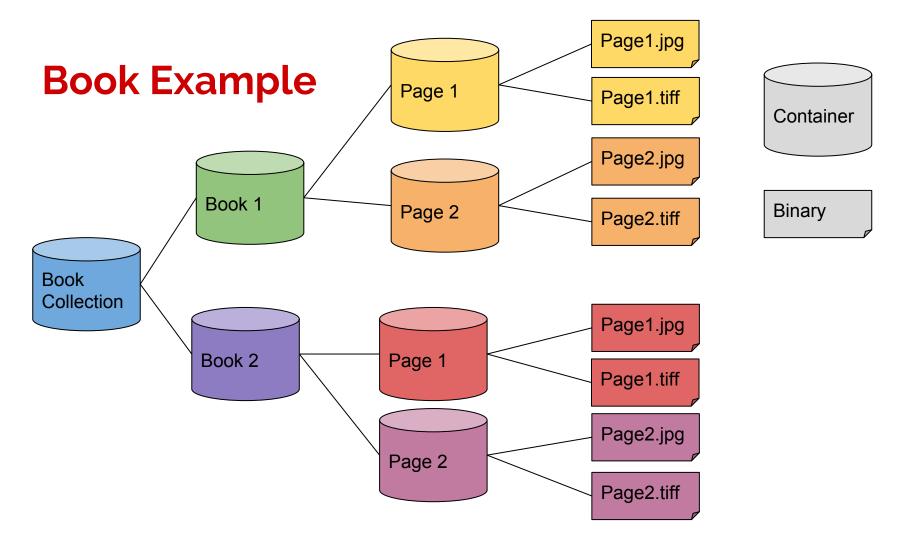


#### Web Resources

Everything is a web resource with a URI

Resources have properties expressed as **RDF triples** 

Resources can contain other resources (containers) or files (binaries)



#### **RDF Properties**

Fedora 4 Home Types Transactions -

#### http://localhost:8080/fcrepo/rest/davidwilcox/book-collection/book1

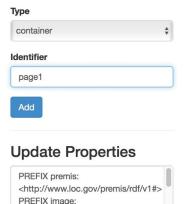
Home / david-wilcox / book-collection / book1

#### **Created at**

2016-05-16T19:06:02.475Z by fedoraAdmin Last Modified at 2016-05-16T19:06:02.475Z by fedoraAdmin Children

# Properties fedora: created 2016-05-16T19:06:02.475Z fedora: createdBy fedora: createdBy fedoradmin fedora: exportsAs http://localhost:8080/fcrepo/rest/david-wilcox/book-collection/book1/fcr:export?format=jcr/xml fedora: hasAccessRoles http://localhost:8080/fcrepo/rest/david-wilcox%2Fbook-collection%2Fbook1/fcr:accessroles fedora: hasParent http://localhost:8080/fcrepo/rest/david-wilcox/book-collection fedora: hasParent http://localhost:8080/fcrepo/rest/david-wilcox/book-collection

#### Create New Child Resource

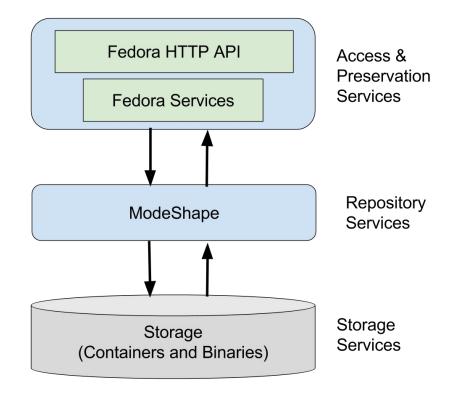


<http://www.modeshape.org/images/

## **Core Features**



#### Fedora system architecture





Focus on existing standards

Fewer customizations to maintain

Opportunities to participate in related communities

#### **Core Services and Standards**

- 1. Create/Read/Update/Delete Linked Data Platform 🖌
- 2. Versioning Memento
- 3. Authorization Web Access Control 🖌
- **4. Fixity** http://tools.ietf.org/html/rfc3230#section-4.3.2 ✓ <sup>1</sup>⁄<sub>2</sub>
- 5. Messaging Activity Streams 2.0 🗸

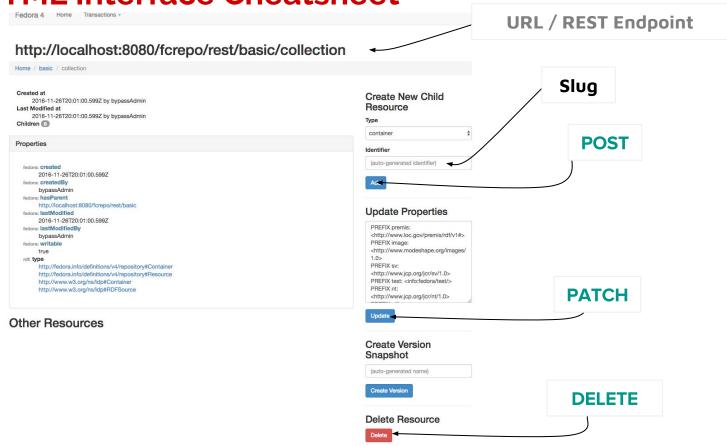
## Hands-on: CRUD

http://localhost:8080/fcrepo/rest/ user/pass: fedoraAdmin/secret3

## Available Operations via HTML UI

- **GET/HEAD/OPTIONS** (Retrieval)
- **POST**/PUT (Creation)
- **PATCH** (Update)
- **DELETE** (Removal)

#### **HTML Interface Cheatsheet**



#### **GET**

#### **Step 1a: RDF Resource Creation** (POST) Fedora 4 Home Transactions -

- Goto 1 http://localhost:8080/fcrepo/rest (root node)
- In "**Type**" select field choose 2. "container" (default)
- In "Identifier" text field enter "basic" 3
- 4. Press "add" button

This will create a new RDF Resource (LDP Basic Container) and redirect us to our next slide!

Username: fedoraAdmin Password: secret3

#### http://localhost:8080/fcrepo/rest/ Home Created at Create New Child by Resource Last Modified at 2016-11-26T20:09:08.874Z by Type Children 1. http://localhost:8080/fcrepo/rest/audit container Properties

Other Resources





basic			
	8	i	
	bas	basic	basid

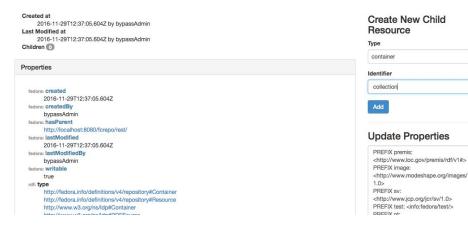
## Step 1b: RDF Resource Creation (POST)

- 1. You will be redirected to <u>http://localhost:8080/fcrepo/rest/basic</u>
- In "Type" select field choose "container" (default)
- In "Identifier" text field enter "collection"
- 4. Press "add" button

This will create a new RDF Resource (LDP Basic Container) and redirect us to our next slide. This way we matched what we had in our cheat sheet! Fedora 4 Home Transactions -

#### http://localhost:8080/fcrepo/rest/basic

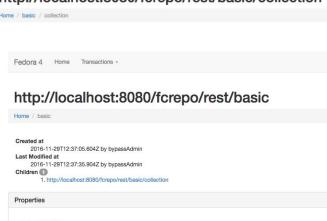
Home / basic



#### **Step 1c: RDF Resource Creation** (POST) Home / basic / collection

- You will be redirected to http://localhost:8080/fcrepo/rest/basic/ collection
- Use "breadcrumb" to go back to 2. http://localhost:8080/fcrepo/rest/basic
- In "**Type**" select field choose 3. "container" (default)
- In "Identifier" text field enter "images" 4.
- 5 Press "add" button

http://localhost:8080/fcrepo/rest/basic/collection





Create New Child

This will create a new RDF Resource (LDP Basic Container) and redirect us to our next slide!

## Step 2: Resource Retrieval (GET)

- Every time you got redirected after creating a Container you were using GET.
- 2. Retrieval is accessed directly via the LDP Path that defines a resource and contains user and some server managed RDF triples.

http:	://localhost:8080/fcrepo/rest/	
Home		
Created	1.4	
Last Mc	dified at	Create New Child Resource
20 Children	016-11-26T20:09:08.874Z by	Туре
	1. http://iocalhost.8080/fcrepo/wst/audit	container \$
Propert	ies	+ Identifier
	_	basic
Othe	r Resources	Add
		Update Properties
		PIETE Mannetic offstrukturen (d. gewijnemiskufder i fa- PIETE Kin mage) Manneticker (d. gewijnemiskufder i fa- PIETE Kin dettrukturen (d. gewijnemiskufder) PIETE Kin offstrukturen (d. gewijnemiskufder) PIETE Kin P
	Fedora 4. Home Transactions -	
	http://localhost:8080/fcrepo/resil Neme / ball: Orested at 2016-11-2012.237.06.60.2by hypess/dmn Last Modified at	/basic Create New Ch Resource
	2016-11-29T12:37:05:604Z by bypassAdmin Children	Туре
	Baseville	container
	Properties	Identifier
	fedora: created 2016-11-2017(2:37:05.604Z fedora: createdBy bypess-Admin	collection Add
	tedami hasParent http://toalinast.8080/tospo/rest/ tedami: lastModified 2016-11-281712:37:05.604Z	Update Proper
	twise lastRedifictly bparAdmin feature writible true et type http://idoox.in/obstrittons/v4/repository/Container http://idoox.in/obstrittons/v4/repository/Container http://idoox.in/obstrittons/v4/repository/Essource http://idoox.in/obstrittons/v4/repository/Essource	PPEIX permit: -http://www.lcg.ov/pm PPETX image: -ttp://www.modelage. 1.0- PPETV image: -petrix image:

## Step 3: Binary Resource Creation (POST)

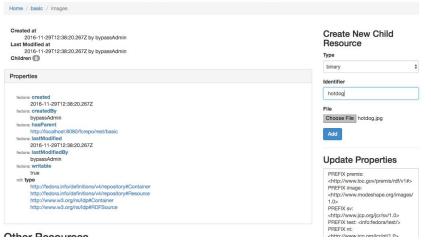
1. Go to

#### http://localhost:8080/fcrepo/rest/basi c/images

- In "Type" select field choose "binary" In "Identifier" text field enter "hotdog"
- 3. In "File" choose any small image
- 4. Press "add" button

This will create a new Binary Resource (LDP Non RDF Source) and redirect us to our next slide!

#### http://localhost:8080/fcrepo/rest/basic/images

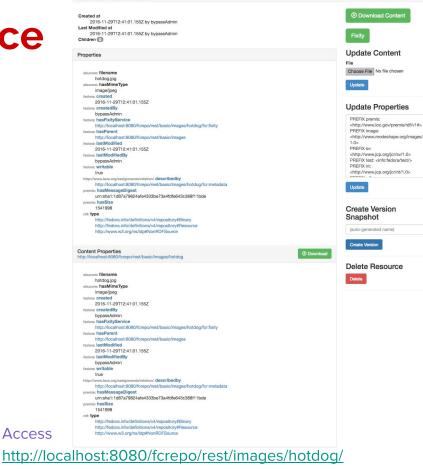


## Step 4: Binary Resource Retrieval (GET)

- 1. You will be redirected to <u>http://localhost:8080/fcrepo/rest/basi</u> <u>c/images/hotdog/fcr:metadata</u>
- 2. Notice the fcr:metadata part!
  - a. Image is LDP contained in "/hotdog"
  - b. Its metadata (rdf properties you can manipulate) in a virtual subpath named /fcr:metadata

Why? That way you can keep operations separated and you can also directly describe via RDF properties binary content. http://localhost:8080/fcrepo/rest/basic/images/hotdog

Home / basic / images / hotdog



directly to download binary!

## Step 5: Update RDF Properties (PATCH)

1. Navigate to

#### http://localhost:8080/fcrepo/rest/basic/collection

- We will add an "pcdm:Object" property using "Update Properties"
  - a. Make sure "PREFIX pcdm" is there
  - b. At the end rewrite "DELETE..." to

DELETE {} INSERT { <> ebucore:width "100"} WHERE {}

c. Press "Update"

	<a href="http://www.w3.org/XML/1998/namespace&gt;">http://www.w3.org/XML/1998/namespace&gt;"&gt;http://ww</a>
<u>on</u>	PREFIX audit: <http: audit#="" definitions="" fedora.info="" v4=""> PREFIX jcr: <http: 1.0="" jcr="" www.jcp.org=""> PREFIX ebucore#&gt; <http: ebucor<br="" metadata="" ontologies="" www.ebu.ch="">e/ebucore#&gt; PREFIX jgr: <http: ldp#="" ns="" www.w3.org=""> PREFIX jgr: <http: ldp#="" ns="" www.w3.org=""> PREFIX fsdora.onfo/definitions/v4/config#&gt; PREFIX fsdora.onfo/definitions/v4/config#&gt; PREFIX mix: <http: 1.0="" jcr="" mix="" www.jcp.org=""> PREFIX mix: <http: ns="" prov#="" www.w3.org=""> PREFIX fggr: <http: 0.1="" foaf="" xmins.com=""></http:></http:></http:></http:></http:></http:></http:></http:>
Download	PREFIX dc: <http: 1.1="" dc="" elements="" purl.org=""></http:>
	DELETE { } INSERT { <> <u>ebucore</u> :width "100" WHERE { }
	Update
	Create Version Snapshot
	(auto-generated name)
	Create Version

## Last step: Delete a resource (DELETE)

1. Stay at

http://localhost:8080/fcrepo/rest/basic/im ages/hotdog/fcr:metadata

- 2. Press "**DELETE**" (the red one)
- 3. You will be redirected to the parent resource after deletion.
- Go again to <u>http://localhost:8080/fcrepo/rest/basic/im</u> <u>ages/hotdog</u>

What do you see?



(auto-generated name)

**Delete Resource** 

Create Version

Delete

#### **Departed**

Fedora creates a tombstone resource at "original/path/fcr:tombstone" URL, in this case "basic/images/hotdog/fcr:tombstone" (try that last path in your Browser)

So, to recreate a resource at that same PATH you need to delete the tombstone placeholder first and that can not be done via HTML UI

## Authorization: Web Access Control



#### **Authorization - Web Access Control**

Authorization is optional and pluggable

WebAC is a W3C approach for managing authorization using linked data

Interoperable with other applications that implement the same approach

Implemented in Fedora 4 by community stakeholders

## Versioning

Versions can be created on demand via the REST-API

A previous version can be restored via the REST-API

# Hands-on: Versioning



#### Create a container named "Book"

#### Create version "vo" of "Book"

#### Add "dc:publisher" to "Book"

INSERT {

<> dc:publisher "University Press"

WHERE { }

#### Create version "v1" of "Book"

Inspect and revert to v0



Over time, digital objects can become corrupt

Fixity checks help preserve digital objects by verifying their integrity

On ingest, Fedora can verify a user-provided checksum against the calculated value

A checksum can be recalculated and compared at any time via a REST-API request

## **External Services**



### **External Component Integrations**

Leverages the well-supported Apache Camel project

Camel is middleware for integration with external systems

Can handle any asynchronous, event-driven workflow



## **External - Indexing**

Index repository content for search

Indexing is configurable - could be based on any property

Solr and Elasticsearch have been tested



### **External - Triplestore**

An external triplestore can be used to index the RDF triples of Fedora resources

Any triplestore that supports SPARQL-update can be plugged in

Fuseki, RDF4J, and BlazeGraph have been tested



rdf4j





Maintains a history of events for each repository resource

Both internal repository events and events from external sources can be recorded

Uses the existing event system and an external triplestore

# **Performance and Scalability**





Testing large files, many files, and many containers

Tests are performed by community members

Have concerns about performance and scale? Join the group!



A number of scalability tests have been run:

Uploaded a 1 TB file via REST API

17 million objects via REST API

3.5 million files via REST API

# **Supporting and Sustaining Fedora**





Managed by DuraSpace (not-for-profit)

Funded by the community

Collaboratively developed by the **community** 

Supported by 2 full-time staff members (not developers)



#### **Useful Resources**

Fedora 4 documentation

https://wiki.duraspace.org/display/FEDORA4x/Fedora+4.x+Documentation

Fedora 4 wiki https://wiki.duraspace.org/display/FF

Fedora 4 mailing lists

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