# Introduction to Fedora 6.0

Overview

David Wilcox, Program Leader, LYRASIS

Arran Griffith, Program Coordinator, LYRASIS



### **Learning Outcomes**

Understand what Fedora can do for you

Understand the key capabilities of the software

Gain hands-on experience with Fedora





Stores, preserves, and provides access to digital objects

Supports flexible content models for objects

Supports semantic relationships between objects

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

### What's new with Fedora 6.0?

Enhanced digital preservation capabilities

**Robust migration tooling** 

Built-in simple search

Improved performance and scale

Metrics collection and reporting

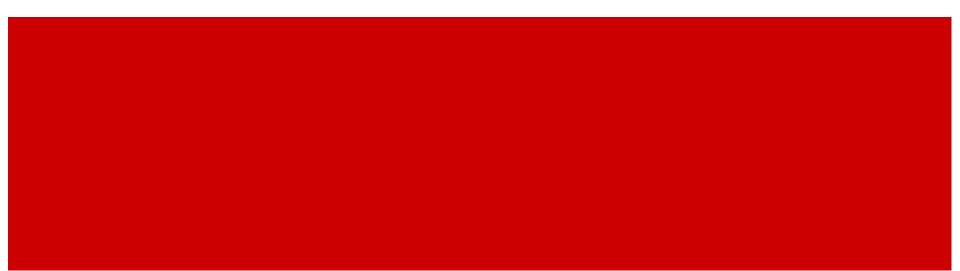
### Fedora Front-Ends

Fedora is *middleware* 

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



# Fedora Examples



### **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** 

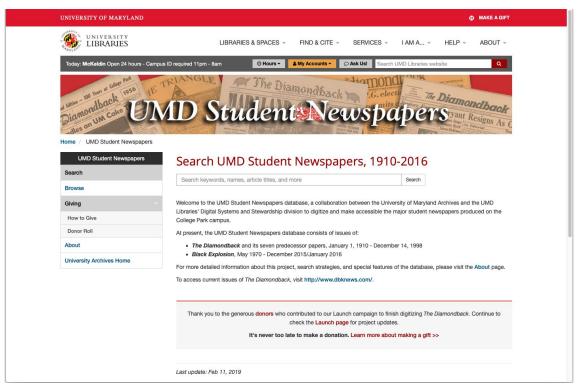
### https://scholarspace.library.gwu.edu/

### **Research Data**

🔛 🛛 Deep Blue Data		About Help Contact +DLogin		
Enter sea	rch terms	Q 60		
University of Michi access and preserv research data that we	repository offered by the gan Library that provides vation services for digital re developed or used in the arch activities at U-M.	Q Browse ⑦ Deposit Your Work		
	Featured Works			
	Featured Works         Recently Uploaded           Retinal fundus images for glaucoma analysis: the RIGA dataset         Depositor: sborda@umich.edu           Keywords: Medical imaging, Optic cup, Image processing, Optic disc, Glaucoma, Optic nerve head, Image segmentation, Fundus images, Automated glaucoma screening system           Neighborhood Effects: Food Environment			

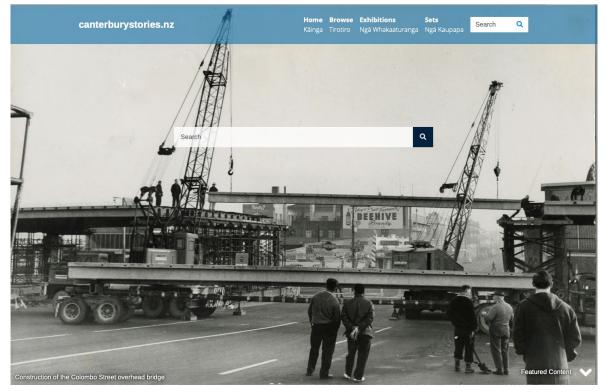
### https://deepblue.lib.umich.edu/data

### **Newspapers**



https://www.lib.umd.edu/univarchives/student-newspapers

### **Special Collections**



#### https://canterburystories.nz

# **Fundamental Concepts**



### Web Resources

Everything is a web resource with an identifier (URI)

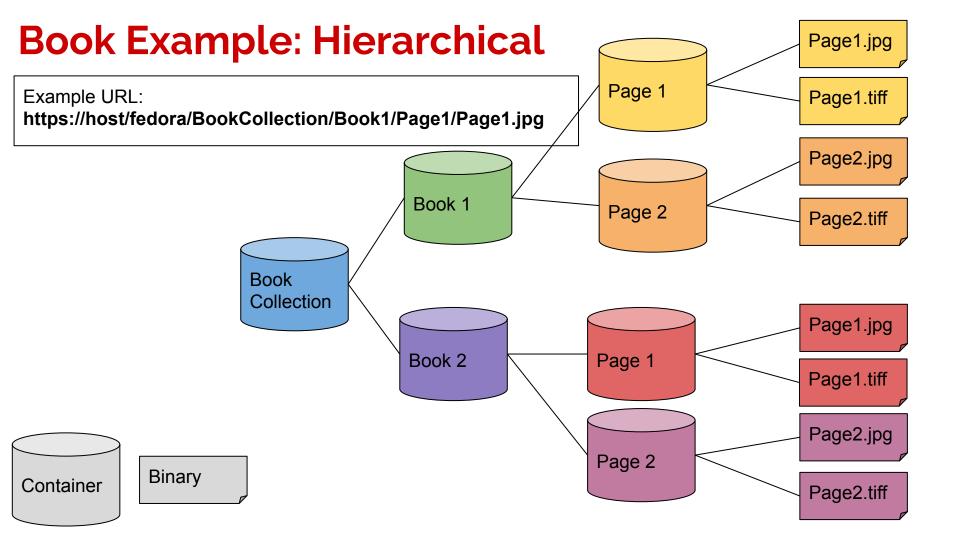
- Container (metadata) resource
- Binary (file) resource

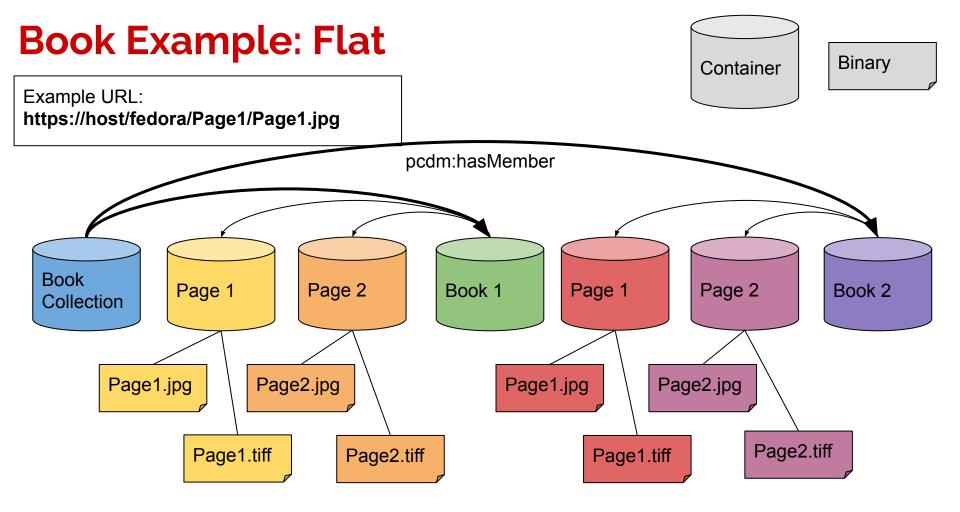
Resources can have properties expressed as RDF triples

Resources can contain other resources

- Containers or
- Binaries







### **RDF Properties**

Fedora 4 Home Transactions -

#### http://localhost:8080/fcrepo/rest/Page1/Page1.jpg

Home / Page1 / Page1.jpg

Download Content Created at 2018-05-07T14:50:42.13Z by fedoraAdmin Last Modified at 2018-05-07T14:50:42.13Z by fedoraAdmin Fixity Children 0 Update Content Properties File Choose File No file chosen ebucore: filename IMG 1328.JPG Update ebucore: hasMimeType image/jpeg fedora: created 2018-05-07T14:50:42.13Z **Update Properties** fedora: createdBy fedoraAdmin PREFIX premis: fedora: hasFixityService <http://www.loc.gov/premis/rdf/v1# http://localhost:8080/fcrepo/rest/Page1/Page1.jpg/fcr:fixity > fedora: hasParent PREFIX test: <info:fedora/test/> http://localhost:8080/fcrepo/rest/Page1 PREFIX rdfs: fedora: lastModified <http://www.w3.org/2000/01/rdf-2018-05-07T14:50:42.13Z schema#> fedora: lastModifiedBy PREFIX xsi: fedoraAdmin <http://www.w3.org/2001/XMLSche

### **Standards**

### **Fedora API Specification**

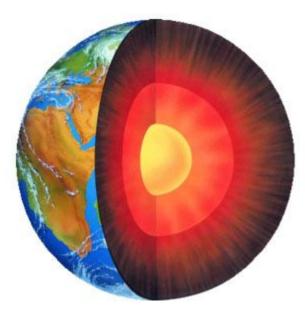
- Focus on existing standards
- Fewer customizations to maintain
- Opportunities to participate in related communities

# **Core Features**



### **Core Services and Standards**

- 1. Create/Read/Update/Delete Linked Data Platform
- 2. Versioning Memento
- 3. Authorization WebAC
- 4. Fixity Digest / Want-Digest
- 5. Simple Search
- 6. Messaging Activity Streams 2.0





Standards-based pattern for creating and managing resources

Basic tools for modeling

content

## **Applied Resource Management (CRUD)**

**Our primary way of managing Fedora resources** 

- Interact with resources via:
  - HTML Interface for browsers (abridged/simplified)
  - HTTP Client (most comprehensive)
    - curl is popular on the command line
    - Every relevant programming language has some sort of http client library

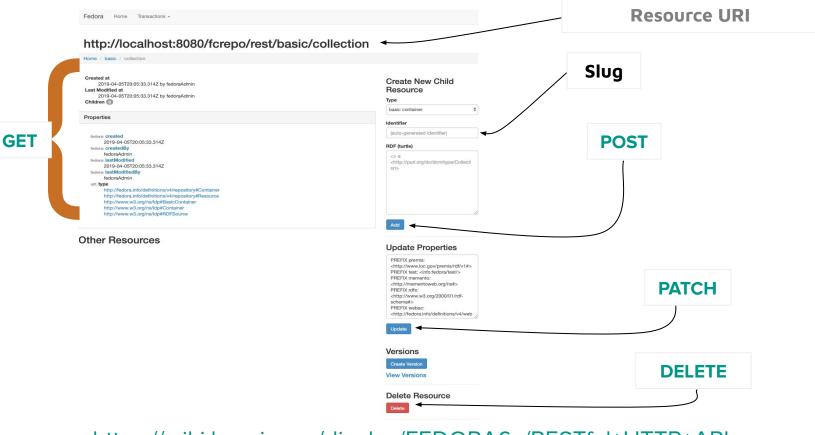


Claim a sandbox instance

Navigate to your sandbox instance

- username: fedoraAdmin
- password: fedoraAdmin

### HTML Interface Cheatsheet

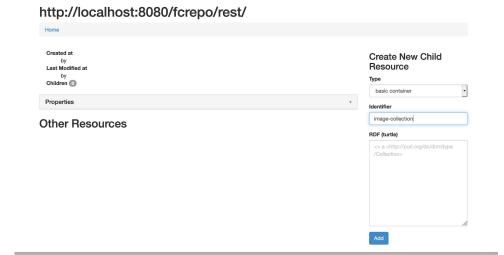


https://wiki.lyrasis.org/display/FEDORA6x/RESTful+HTTP+API

### Step 1a: Resource Creation (POST)

Home Transactions

- 1. Go to http://ip.address:8080/fcrepo/rest
- In "Type" select field choose "basic container" (default)
- 3. In **"Identifier"** text field enter "image-collection"
- 4. Press "add" button



\$ curl -i -X POST -ufedoraAdmin:fedoraAdmin -H "Slug: image-collection"
http://ip.address:8080/fcrepo/rest

## Step 1b: Resource Creation (POST)

edora Home Transactions

- 1. You will be redirected to http://ip.address:8080/fcrepo/rest/image-collection/
- In "Type" select field choose "basic container" (default)
- 3. In **"Identifier"** text field enter "andromeda-galaxy"
- 4. Press "add" button

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

Treated at 2019-08-21T09:57:28.345Z by fedoraAdmin ast Modified at 2019-08-21T09:57:28.345Z by fedoraAdmin Middren 0	Create New Child Resource <sup>Type</sup>
roperties	basic container
feders: <b>created</b> 2010,08.21T09-57-28-3457	andromeda-galaxy
2019-08-21T09:57:28.3452 fedora.createdBy fedoraAdmin fedora.lastModified 2019-08-21T09:57:28.3452 fedora.lastModifiedBy fedoraAdmin reft type http://fedora.lnfo/definitions/v4/repository#Container http://fedora.lnfo/definitions/v4/repository#Resource http://fedora.lnfo/definitions/v4/repository#Resource	RDF (turtle)
http://www.w3.org/ns/dp#BasicContainer http://www.w3.org/ns/dp#Container http://www.w3.org/ns/ldp#RDFSource	

\$ curl -i -X POST -ufedoraAdmin:fedoraAdmin -H "Slug: andromeda-galaxy"
http://ip.address:8080/fcrepo/rest/image-collection

## 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 path that defines a resource and contains user and some server managed RDF triples.

#### http://localhost:8080/fcrepo/rest/imagecollection/andromeda-galaxy

Home / image-collection / andromeda-galaxy

Fedora Home Transactions -

reated at 2019-08-21109:58:16.215Z by fedoraAdmin ast Modified at	Create New Child Resource	
2019-08-21T09:58:16.215Z by fedoraAdmin hildren 0	Туре	
	basic container	
roperties	Identifier	
fedora: created 2019-08-21T09-58:16.215Z fedora: createdBy	(auto-generated identifier) RDF (turtle)	
Tedorat Create Day fedora Admin fedora: BatModified 2019-08-21705-58:16.215Z fedora: BatModifiedBy fedoraAdmin	<> a <http: dc="" dcmitype<br="" purl.org="">/Collection&gt;</http:>	
rdt: type http://fedora.info/definitions/v4/repository#Container http://fedora.info/definitions/v4/repository#Resource http://www.w3.org/ns/ldp#BasicContainer http://www.w3.org/ns/ldp#Cortainer http://www.w3.org/ns/ldp#CoFSource		

\$ curl -i -H "Accept:text/turtle" -ufedoraAdmin:fedoraAdmin http://ip.address:8080/fcrepo/rest/image-collection/andromeda-galaxy

### **Binaries**

We'll use an image for (most) examples of binaries

Download from: https://go.nasa.gov/2Hc15jG

(Save as andromeda.jpg)

Andromeda Galaxy

Credit: NASA



## Step 3: Binary Resource Creation (POST)

edora Home Transactions

ome / image-collection / andromeda-galax

#### 1. Go to

http://ip.address:8080/fcrepo/rest/image-collectio n/andromeda-galaxy

- In "Type" select field choose "binary" In "Identifier" text field enter "andromeda.jpg"
- 3. In "File" choose the andromeda.jpg image
- 4. Press "add" button

#### nttp://localhost:8080/fcrepo/rest/imagecollection/andromeda-galaxy

eated at 2019-08-21109:58:16.215Z by fedoraAdmin st Modified at 2019-08-21109:58:16.215Z by fedoraAdmin sildren ()	Create New Child Resource Type binary
operties	Identifier
fedora: cmated 2019-08-21709:58:16.215Z fedora: createdBy fedora: Advmin 2019-08-21709:58:16.215Z 2019-08-21709:58:16.215Z fedora: IastModifiedBy fedoraAdmin et bype	andromeda.jpg File Browse andromeda.jpg Add Update Properties
http://fedora.info/definitions/v4/repositorytContainer http://fedora.info/definitions/v4/repositorytResource http://www.w3.org/nsi/dp#Gontainer http://www.w3.org/nsi/dp#Container http://www.w3.org/nsi/dp#RDFSource	PREFIX premis: -http://www.loc.gov/premis /rdf/v1#> PREFIX test:-info:fedora/test/> PREFIX mento:

\$ curl -i -X POST -ufedoraAdmin:fedoraAdmin -H"Content-Type: image/jpeg" -H"Slug: andromeda.jpg"
--data-binary "@andromeda.jpg"
http://ip.address:8080/fcrepo/rest/image-collection/andromeda-galaxy

## Step 4: Binary Resource Retrieval (GET)

1. You will be redirected to

http://ip.address:8080/fcrepo/rest/image-collection/andromeda-galaxy/andromeda.jpg/fcr:metadata

- 2. Notice the fcr:metadata part!
  - a. Image content is at "andromeda.jpg"
  - b. Its metadata (rdf properties) in a virtual subpath named /fcr:metadata
  - c. The fcr:metadata part is an implementation detail, but the fact that the RDF that describes a binary has a URI and can be linked to is important.

\$ curl -i -u fedoraAdmin:fedoraAdmin http://ip.address:8080/fcrepo/rest/image-collection/andromeda-galaxy/andromeda.jpg/fcr:metadata

# Step 5: Update RDF Properties (PA'

1. Go to

http://ip.address:8080/fcrepo/rest/image-collection/andromeda-galaxy

- 2. We will add a "dc:title" property using "Update Properties"
  - a. At the end rewrite "DELETE..." to

DELETE {}
INSERT { <> dc:title "Andromeda Galaxy"}
WHERE {}

b. Press "Update"

a-galaxy	Update Content File Browse No file selected.
es"	Update
	Update Properties
	/v4/config#> PREFIX foaf: <http: xmlns.com<br="">/foaf/0.1/&gt; PREFIX dc: <http: purl.org<br="">/dc/elements/1.1/&gt;</http:></http:>
	DELETE {} INSERT { <> dc:title "Andromeda Galaxy"} WHERE {}

\$ curl -i -XPATCH -H"Content-type:application/sparql-update" -ufedoraAdmin:fedoraAdmin -d "INSERT DATA {<> <http://purl.org/dc/elements/1.1/title> 'Andromeda Galaxy'}" http://ip.address:8080/fcrepo/rest/image-collection/andromeda-galaxy/andromeda.jpg/fcr:metadata

### **Our updated RDF Properties from step 5**.

Properties
dc: title
Andromeda Galaxy
ebucore: filename
andromeda.jpg
ebucore: hasMimeType
image/jpeg
fedora: created
2019-08-21T10:07:12.984Z
fedora: createdBy
fedoraAdmin
fedora: hasFixityService
http://localhost:8080/fcrepo/rest/image-collection/andromeda-galaxy/andromeda.jpg/fcr:fixity
fedora: lastModified
2019-08-21T10:08:54.065Z
fedora: lastModifiedBy
fedoraAdmin
iana: describedby
http://localhost:8080/fcrepo/rest/image-collection/andromeda-galaxy/andromeda.jpg/fcr:metadata
premis: hasMessageDigest
urn:sha1:50db9edaaebf54e2afaa801c0d64750591763e57
premis: hasSize
654014
rdf: type
http://fedora.info/definitions/v4/repository#Binary
http://fedora.info/definitions/v4/repository#Resource
http://www.w3.org/ns/ldp#NonRDFSource

## Last step: Delete a resource (DELETE)

1. Go to

http://ip.address:8080/fcrepo/rest/image-collection/andr omeda-galaxy/andromeda.jpg

- 2. Press "DELETE" (the red one)
- 3. Reload

http://ip.address:8080/fcrepo/rest/image-collection/andr omeda-galaxy/andromeda.jpg

What do you see?

#### Versioning

Versioning of binary resources is not currently available through the HTML UI.

#### **Delete Resource**

Delete

\$ curl -i -XDELETE -ufedoraAdmin:fedoraAdmin
http://ip.address:8080/fcrepo/rest/image-collection/andromeda-galaxy/andromeda.jpg



Discovered tombstone resource at /image-collection/andromeda.jpg, departed: 2019-08-21T17:16:17.047Z

Fedora creates tombstone resources at "original/path/fcr:tombstone" URL, in this case "image-collection/andromeda-galaxy/andromeda.jpg/fc r:tombstone"

So, to recreate a resource at that same PATH you need to delete the tombstone placeholder first

# Versioning



Versions: created automatically or on demand via the REST-API

Versions: retrieved via Memento protocol

# time travel

http://			
2004-01-10	02:45:06	Find	Reconstruct

Find Mementos in Internet Archive, Archive-It, British Library, archive.today, GitHub and many more!

### What is a "version"?

It's a resource (so it has a URI, can be linked to, etc)

Contains the contents of a resource at a given point in time

It is immutable



#### What is a "version"?

Has a date (to the second)

Is discoverable (i.e. does this resource have any versions?)

Created by a versioning API, i.e.\*not\* plain CRUD operations

#### **Meet the Memento Specification:**

RFC 7089: https://tools.ietf.org/html/rfc7089

"HTTP Framework for Time-Based Access to Resource States"



## **Key Memento Concepts**

- **Memento** versioned body of the original resource
- **Time Gate** an endpoint supporting time-based queries for Mementos
- **Time Map** list of mementos
- **Original Resource** the resource that was versioned

#### **Creating a Version**

Create a resource in Fedora:

http://ip.address:8080/fcrepo/rest/image-collection/map

### Example 1: Creating a Version

Versions can be created via HTML Interface

Note: Auto-versioning is enabled on the demo site



\$ curl -i -X POST -ufedoraAdmin:fedoraAdmin
http://ip.address:8080/fcrepo/rest/image-collection/map/fcr:versions

## **Modify the Resource**

Update the metadata of http://ip.address:8080/fcrepo/rest/image-collection/map

```
DELETE {}
INSERT { <> dc:title "Map of Nova Scotia" }
WHERE {}
```

### **View the Version History**

View versions via HTML Interface

#### Versions

**Create Version** 

View Versions

\$ curl -i -X POST -ufedoraAdmin:fedoraAdmin
http://ip.address:8080/fcrepo/rest/image-collection/map/fcr:versions

## **Advanced Versioning**

You can do many more things with versions:

- Get a list of all versions
- Find the closest version to a specific date
- And more!

https://wiki.lyrasis.org/display/FEDORA6x/Versioning

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

W3C approach for managing authorization using linked data

Interoperable with other applications that implement the same approach



## 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 digest against the calculated value

## **On demand**, A digest can be recalculated and compared at any time via a REST-API request

## **Limitations of Fixity Checking**

- It's a check. Plan for what to do should you encounter a bad binary!
  - Versioning
  - Alternate Backups
- Only works on Fedora-managed content
  - Doesn't retrieve external content
- Digest algorithms implemented: sha1, sha256, and md5

## Fixity: Hands On

Create a new binary resource parade.jpg inside the /image-collection container. Download from <u>http://bit.ly/2g8GWy7</u>

Check the fixity of parade.jpg

- 1. Goto http://ip.address:8080/fcrepo/rest/image-collection/parade.jpg/fcr:metadata
- 2. Click the button labeled Fixity



\$ curl -i -u fedoraAdmin:fedoraAdmin -X GET
http://ip.address:8080/fcrepo/rest/image-collection/parade.jpg/fcr:fixity

#### Conclusion

**Fixity** works as part of Fedora's preservation tooling to ensure the long term stability of your resources.



#### \*NEW\* Simple Search

Supports simple queries

Built into Fedora

Based on an internal database index

#### **HTML Search Interface**

Click the Search button at the top

Fedora Home Tran



Default query displays all repository resources

Enter a search to perform. Where field can be one of:	Query (currently only supports one condition at a time).			
• fedora_id: The ID or URI of the resource.	Field fedora_id S Operator <= S Query term			
<ul> <li>modified: The last modified date/time.</li> <li>created: The created date/time.</li> </ul>				
<ul> <li>rdf_type: The rdf:type of the resource.</li> </ul>	info:fedora/*			
<ul> <li>content_size: The size of the binary or RDF.</li> <li>mime type: The mime-type of the binary object.</li> </ul>				
pperator can be one of <=, < =, >, >=	Submit			
perator can be one of <=, < =, >, >=				

ID	RDF types	Content size	MimeType	Created	Last Modified
http://localhost:8080/rest/image- collection	http://fedora.info/definitions/v4/repository#Container, http://fedora.info/definitions/v4/repository#Resource, http://mementoweb.org/ns#OriginalResource, http://mementoweb.org/ns#TimeGate, http://www.w3.org/ns/ldp#BasicContainer, http://www.w3.org/ns/ldp#Resource http://www.w3.org/ns/ldp#RDFSource, http://www.w3.org/ns/ldp#Resource	0		2021-03- 10T16:35:58.732Z	2021-03- 10T16:35:58.732Z
http://localhost:8080/rest/image- collection/andromeda-galaxy	http://fedora.info/definitions/v4/repository#Container, http://fedora.info/definitions/v4/repository#Resource, http://mementoweb.org/ns#OriginalResource, http://mementoweb.org/ns#TimeGate, http://www.w3.org/ns/ldp#BasicContainer, http://www.w3.org/ns/ldp#Rostainer, http://www.w3.org/ns/ldp#RDFSource, http://www.w3.org/ns/ldp#Resource	109		2021-03- 10T16:36:31.337Z	2021-03- 10T16:38:25.9192

## **Run a Simple Query**

Find the map resource

Field: fedora\_id

**Operator:** =

**Query:** \*map

Query (currently only supports one condition at a time).

Field	fedora_id	0	Operator	=	0		
Query term							
*map	)						
Subr	nit						
Subr	nit						

## **Run Another Query**

Find the parade image by mimetype resource

Field: mime\_type

**Operator:** =

**Query:** \*jpeg

Query (currently only supports one condition at a time).





Enables asynchronous integrations

All actions that affect a repository change... result in a message



## **External Services**



### **Message Based Integrations**

External services listen for messages from Fedora

Relevant messages trigger services

This system is scalable and fault-tolerant



#### **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







## 2021 Roadmap



#### **Release Timeline**

- 1. December 2020: Fedora 6.0 Alpha
- 2. February 2021: Fedora 6.0 Beta
- 3. June 2021: Fedora 6.0 Release Candidate
- 4. June/July 2021: Fedora 6.0 Release

#### **Available for Testing**

- 1. Fedora 6.0 Release Candidate
- 2. <u>Migration tools</u>
- 3. <u>Migration validator</u>



#### **Useful Resources**

Fedora user docs

Fedora 6.0 Design and Development

**Communication Channels**