Powering Linked Open Data Applications

With Fedora and Islandora CLAW

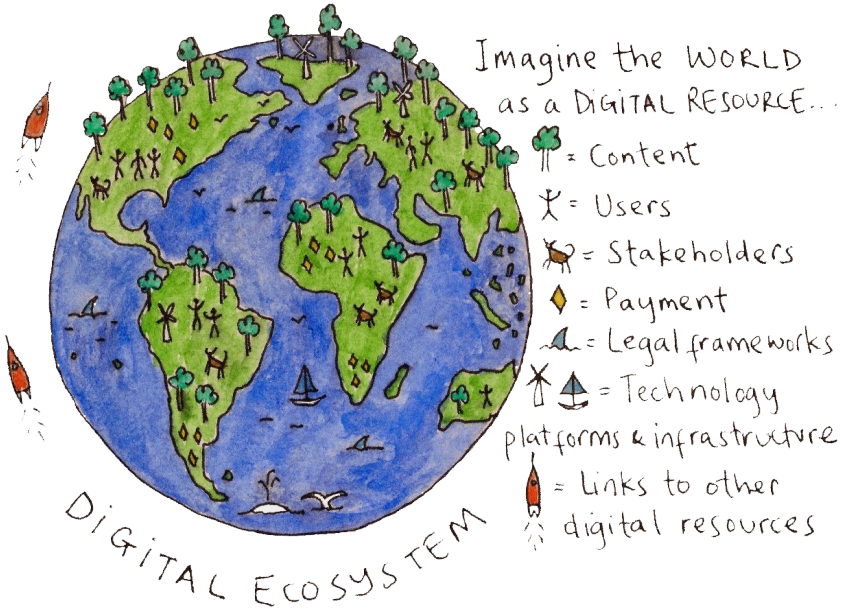
David Wilcox, DuraSpace
dwilcox@duraspace.org

About DuraSpace

DuraSpace is not for profit organization that provides leadership and innovation for **open technologies**

We work to ensure that **current and future generations** have **durable** and **persistent access** to our **collective digital heritage**

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



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

DURASPACE

DuraSpace open source projects

Fedora...

Stores, preserves, and provides access to digital objects

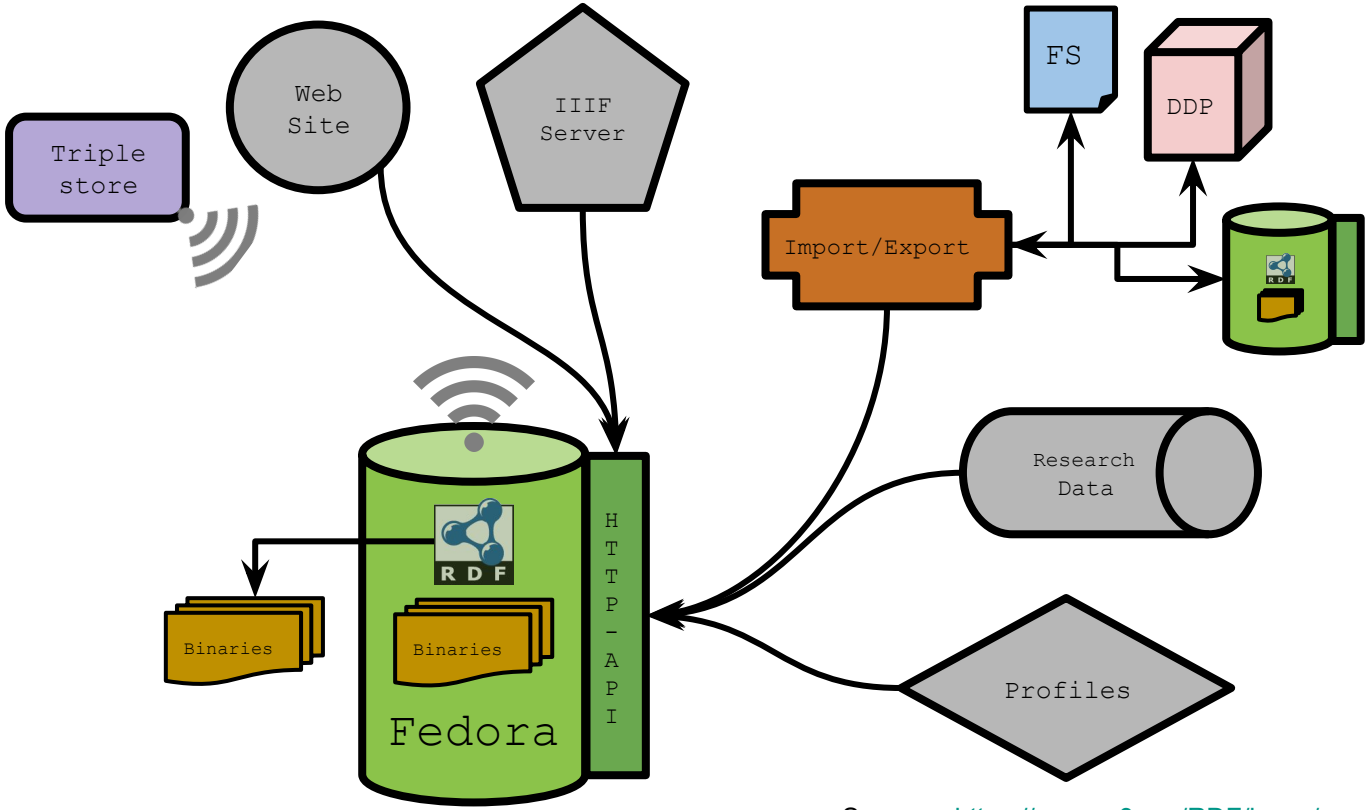
Supports flexible content models for objects

Supports semantic relationships between objects using RDF

Supports millions of objects, both large and small

Interoperates with other applications and services

Component of an Ecosystem



Sources: https://www.w3.org/RDF/icons/
https://pixabay.com/en/icon-communication-sender-antenna-157359/

Core Services and Standards

1. **Create/Read/Update/Delete** - Linked Data Platform

2. **Versioning** - Memento

3. **Authorization** - Web Access Control

4. **Fixity -** HTTP Headers

5. ***Messaging*** - Activity Streams 2.0

https://fedora.info/spec

Fedora API Specification and 5.0 Release

Proposed Recommendation for API Specification 1.0 now available

Fedora 5.0 release candidate available for testing

Fedora 5.0 represents alignment with the specification

- Release numbers now based on semantic versioning

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

elastic

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

Islandora CLAW

Islandora is...

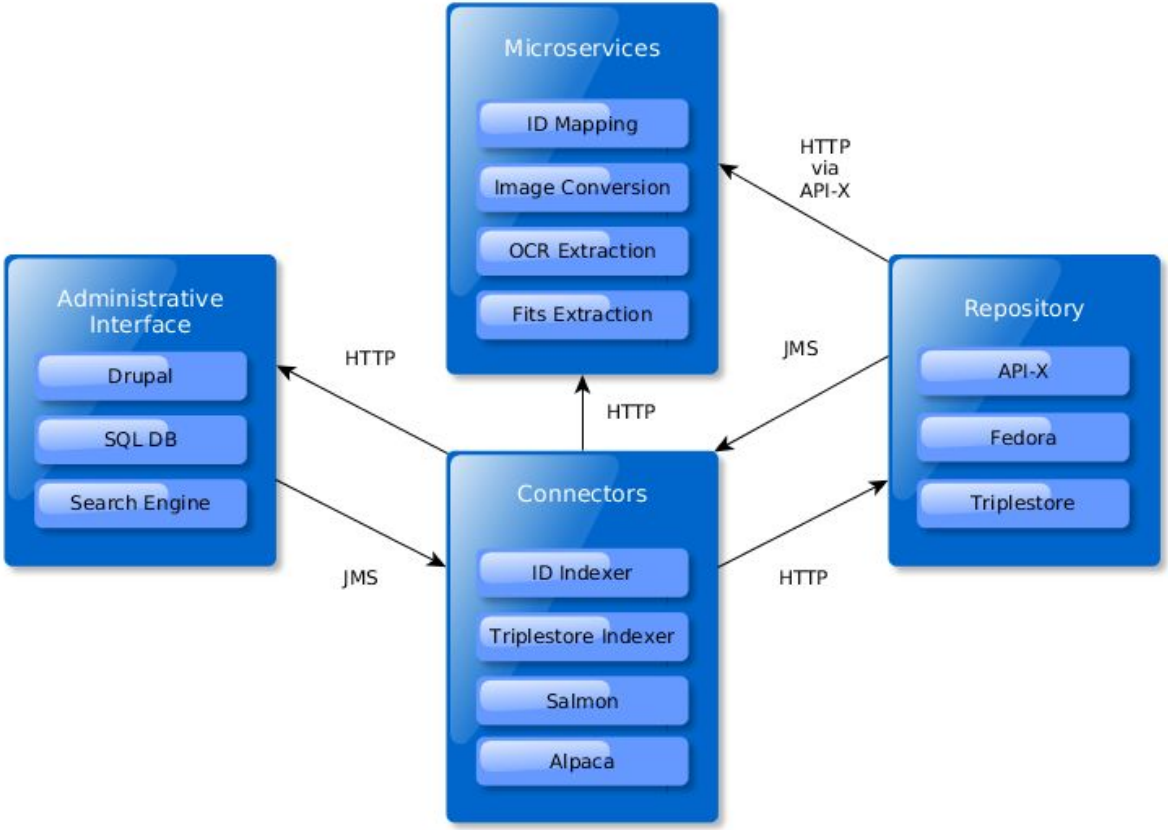
A collection of Drupal modules that enables you to, ingest, manage, and display objects stored in a Fedora repository.

The ease and flexibility of Drupal on top of the robust storage of Fedora, with Solr for discovery.

What is Islandora CLAW?

- Drupal 8
- Solr
- Fedora 4+
- Modules
- Apache Camel
- Connectors
- Microservices
- YAML Configuration

CLAW Architecture

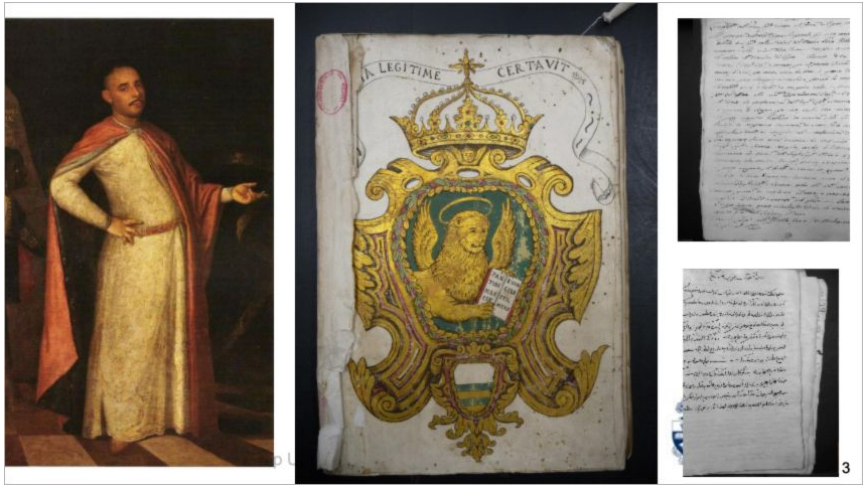


Islandora CLAW Use Case

Linked data modeling for humanities data at The Digital Scholarship Unit at the University of Toronto Scarborough Library

Dragoman Renaissance Research Platform

Collaboration with Prof. Natalie Rothman: Dragomans, diplomatic interpreter-translators ca. 1550-1750



Project Stages

1. ontology development
2. CLAW implementation
3. data migration, refinement
4. data analysis

Why Islandora CLAW?

Linked Data Application Platform

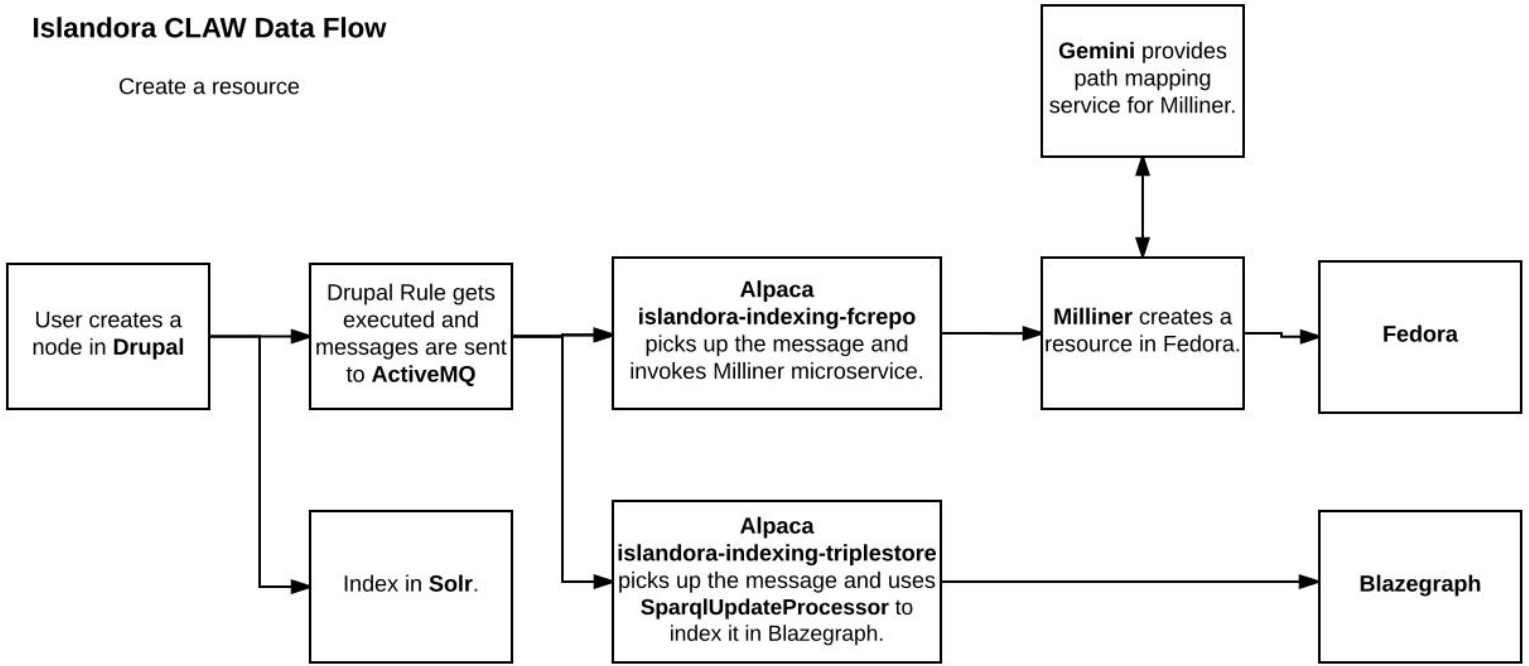
Existing skills and expertise

Leverage content modelling features

Leverage flexible views and search in Drupal

Experiment with different modules in Drupal

Islandora CLAW Architecture



Implementing in CLAW

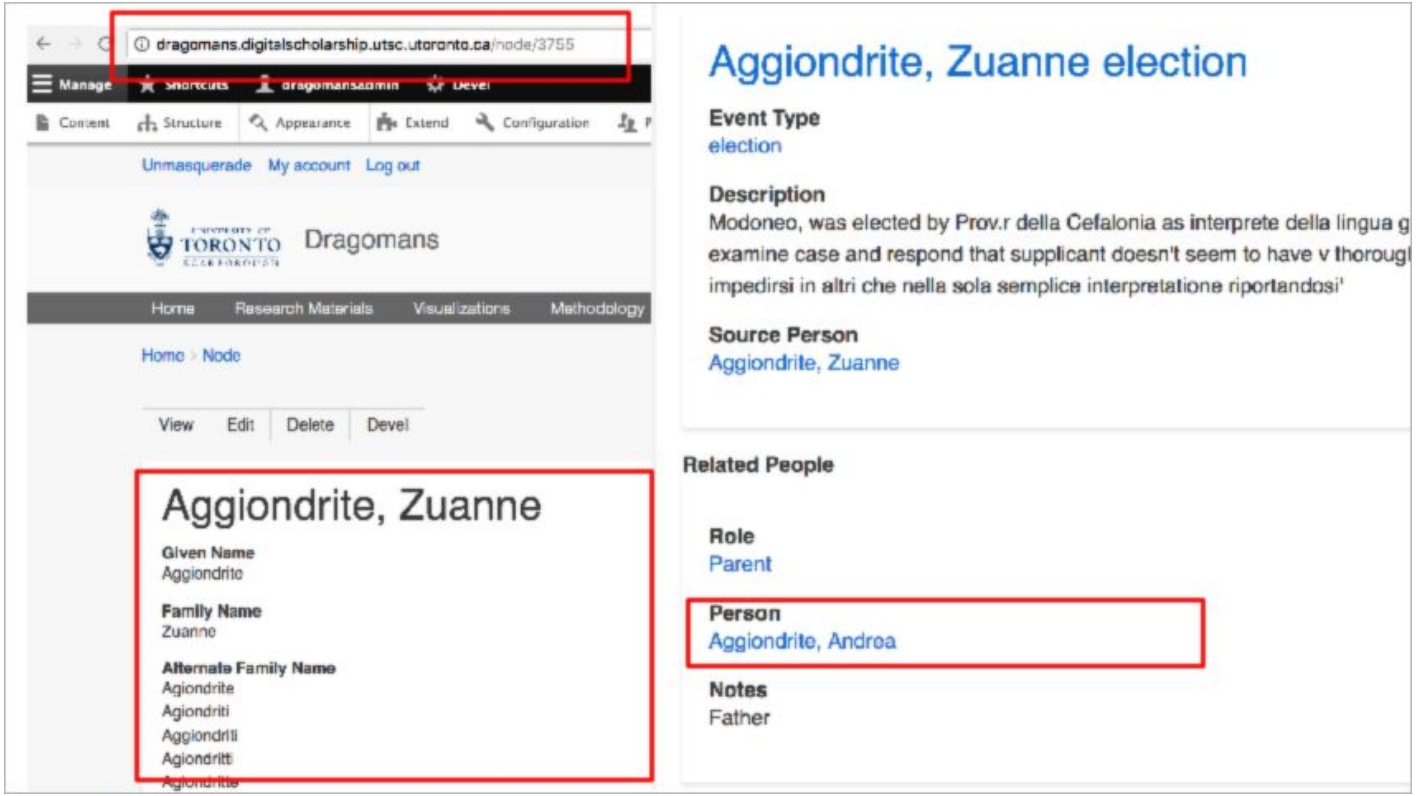
Content Modeling based on ontology

- Classes and Taxonomies into Content Types
- Views to show related content
- Search configuration for faceting

RDF Mapping

Inferences

Resource Structure



Underlying JSON-LD

Querying the Triplestore

Visualization with Palladio

Next Steps

Interoperability with other linked data platforms (enrichment, expansion of data)

Add flexible scholarship tools that address needs of multiple projects

Speed up deployment of new projects on the platform without increasing staffing needs within the unit

Simplify on-ramping of researchers who may not be experts in linked-data

Supporting Islandora and Fedora

Fedora facts

Managed by DuraSpace (not-for-profit)

Funded by the community

Collaboratively developed by the community

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

ASU
BROWN
CASE WESTERN RESERVE UNIVERSITY EST. 1826
think beyond the possible
VASSAR
RUHR UNIVERSITÄT BOCHUM
RUB
DEFF
Danmarks Elektroniske Fag- og Forskningsbibliotek
COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK
Carnegie Mellon University
docuteam
UNIVERSITY of York
LYRASIS
Advancing Knowledge
Durham University
EMORY UNIVERSITY
THE GEORGE WASHINGTON UNIVERSITY WASHINGTON, DC
UNIVERSITY of VIRGINIA
INDIANA UNIVERSITY
NATIONAL LIBRARY OF MEDICINE
UCL
Université catholique de Louvain
LA TROBE UNIVERSITY
LAFAYETTE COLLEGE
ICPSR
JOHNS HOPKINS UNIVERSITY
universität wien
McMaster University
UNIL | Université de Lausanne
UCONN
UNIVERSITY OF CONNECTICUT
LSE
THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE
ART INSTITUTE CHICAGO
slandora
Northwestern
Northeastern
UNIVERSITY of DENVER
UNIVERSITY OF HULL
UCLA
UNSW AUSTRALIA
PennState
UC San Diego
Tufts UNIVERSITY
National Research Council Canada
OSU
Oregon State UNIVERSITY
RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY
Amherst College
UNIVERSITY of Cincinnati
THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL
UNIVERSITY of ROCHESTER
UCSB
UNIVERSITY OF OXFORD
Smithsonian
UPPSALA UNIVERSITET
MACQUARIE UNIVERSITY
UNIVERSITY OF MICHIGAN
UNIVERSITY OF NOTRE DAME
STANFORD UNIVERSITY
UNIVERSITY of MARYLAND
UNIVERSITY of Prince Edward ISLAND
UNIVERSITY of HOUSTON
University of Manitoba
THE UNIVERSITY of TEXAS AT AUSTIN
NIMS
THE UNIVERSITY of TENNESSEE KNOXVILLE
The UNIVERSITY of OKLAHOMA
YORK UNIVERSITY
UNIVERSITY OF OREGON
Czech Academy of Sciences
WISCONSIN UNIVERSITY OF WISCONSIN–MADISON
UNIVERSITY OF TORONTO
University of Pittsburgh
UNIVERSITÄT BERN
University of Michigan Medical School
THE OHIO STATE UNIVERSITY
VILLANOVA UNIVERSITY
University of the Sunshine Coast Queensland, Australia
Yale

Islandora

230+ public sites, 1414 listserv members, 144 GitHub members, 28 Committers, 26 Contributed Modules, 29 Foundation members, 12 Interest Groups, 9 Community-led software releases, 19 Camps, 3 conferences, 2 major versions.

And 2 full-time employees. So we need the Community.

Useful Resources

Fedora documentation
https://wiki.duraspace.org/display/FF/Documentation

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

Islandora CLAW
https://github.com/Islandora-CLAW