Using BIBFRAME, Flask, and Fedora4 as a Catalog Pull Platform

What is a Catalog Pull Platform?
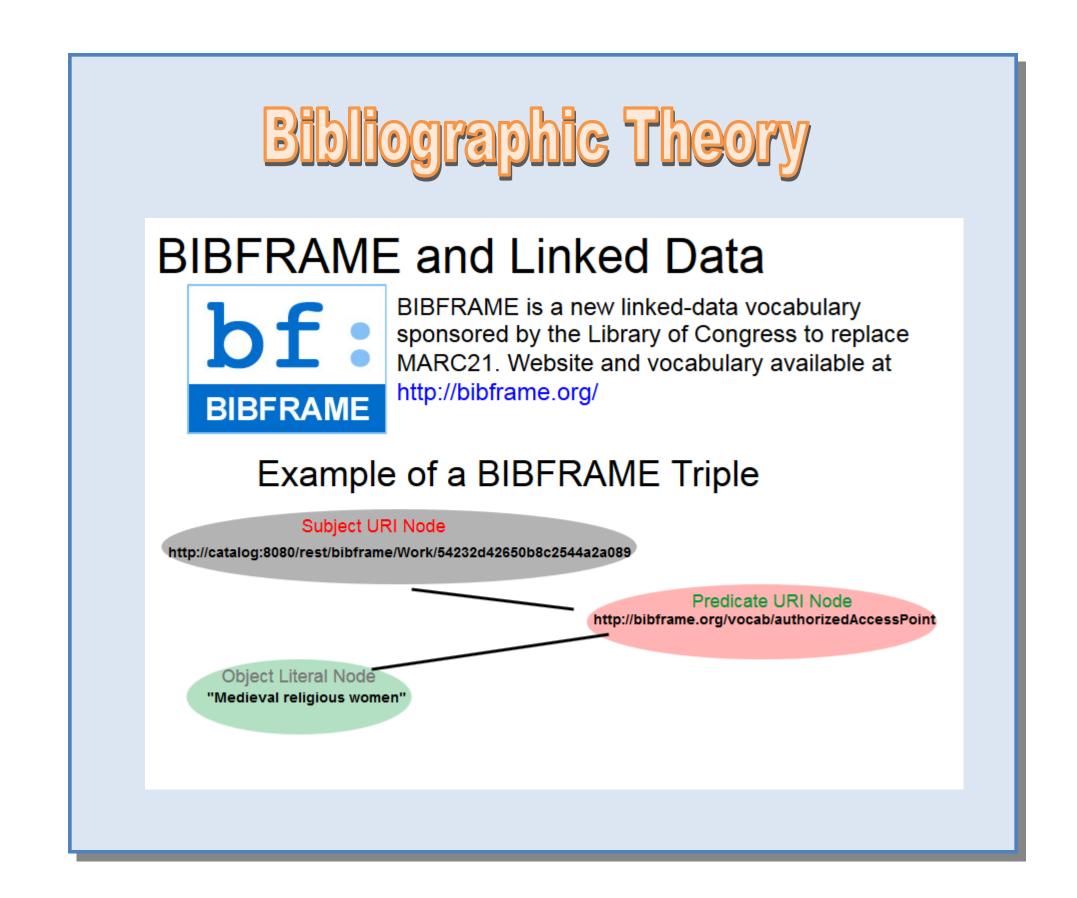
The Tutt Library Catalog Pull Platform:
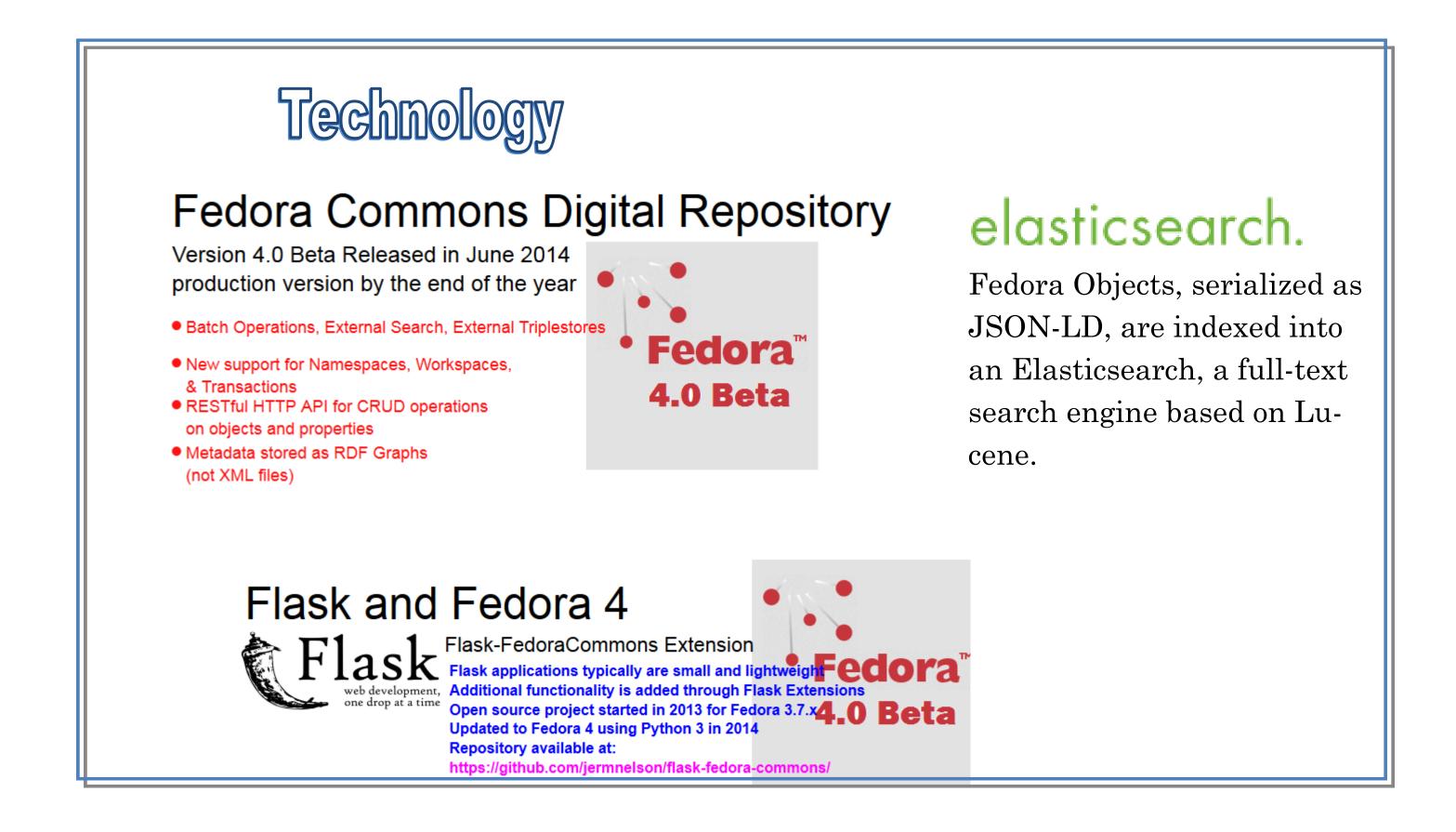
modifies catalogs for libraries and other cultural heritage institutions
enables
builds
creates

The Catalog Pull Platform offers individual libraries a robust and responsive online catalog for accessing and managing their collections. A library, or other memory institutions, should be able to easily customize and integrate their collections with their various web properties, resource management, search-engine and social media. Instead of being dependent on a set of features, that is offered or "pushed" from the vendor, the features in the Catalog Pull Platform are directly "pulled" from the needs of various cohorts of patrons, staff, librarians, library administration, other organizations, and finally the web harvesting algorithms of commercial search engines.

Jeremy Nelson
Metadata and Systems Librarian
Colorado College
Email: jeremy.nelson@coloradocollege.edu
Twitter: @jermnelson

Bibliographic Theory

BIBFRAME and Linked Data

bf: BIBFRAME

BIBFRAME is a new linked-data vocabulary sponsored by the Library of Congress to replace MARC21. Website and vocabulary available at http://bibframe.org/

Example of a BIBFRAME Triple

Subject URI Node
http://catalog:8080/rest/bibframe/Work/54232d42650b8c2544a2a089

Predicate URI Node
http://bibframe.org/vocab/authorizedAccessPoint

Object Literal Node
"Medieval religious women"

Technology

Fedora Commons Digital Repository

Version 4.0 Beta Released in June 2014
production version by the end of the year

Fedora™ 4.0 Beta

- Batch Operations, External Search, External Triplestores
- New support for Namespaces, Workspaces, & Transactions
- RESTful HTTP API for CRUD operations on objects and properties
- Metadata stored as RDF Graphs (not XML files)

elasticsearch.

Fedora Objects, serialized as JSON-LD, are indexed into an Elasticsearch, a full-text search engine based on Lucene.

Flask and Fedora 4

Flask
web development, one drop at a time

Flask-FedoraCommons Extension
Flask applications typically are small and lightweight
Additional functionality is added through Flask Extensions
Open source project started in 2013 for Fedora 3.7.x
Updated to Fedora 4 using Python 3 in 2014
Repository available at:
https://github.com/jermnelson/flask-fedora-commons/

Fedora™ 4.0 Beta

MARC-to-BIBFRAME into Fedora Workflow

Step 1: MARC21 records extracted from legacy ILS

MARC 21

Library of Congress XQuery utility

MARC XML

Step 4. MARC XML is converted to BIBFRAME RDF SML through LOC XQuery utility

RDF XML

bf: BIBFRAME

Step 5. BIBFRAME Resource Graphs extracted and ingested into Fedora 4 and then indexed into Elasticsearch instance

Step 2: MARC21 record ingested into Fedora4

Fedora 4 Repository

BF:Annotation

BF:Person

BF:Work

BF:Instance

elasticsearch.

JSON-LD

Step 3. JSON-LD serializations of MARC Fedora Objects indexed into Elasticsearch

Schema.org Web Editor

Another application of the Catalog Pull Platform is a simple, Flask-based Web editor for Schema.org RDF vocabulary that allows triples to be applied to any new or existing Fedora 4 objects,

Schema.org Editor
localhost:8100

Schema.org Editor

Bo

BodyOfWater   Body of Water
Bone
Book
BookFormatType   Book Format Type
BookStore   Book Store
BookmarkAction   Bookmark Action
BorrowAction   Borrow Action
BowlingAlley   Bowling Alley

https://github.com/jermnelson/schema-org-editor

http://intro2libsys.info/dlf-forum-2014-poster