



fedora



DELOS Digital Library Summer School

Rethinking the Role of Repositories

Thursday, May 31, 2007
Settignano, Florence, Italy

Sandy Payette
Director, Fedora Project
Researcher, Cornell Information Science



fedora

Tutorial Schedule

Time	
20 min	Intro: New vision for repositories and digital libraries
40 min	What about semantic web and semantic technologies?
15 min	Break
30 min	Fedora repository and service-oriented architecture
15 min	Discussion and questions



fedora



New vision for repositories and digital libraries





Positioning for New User Requirements

- How are user needs evolving?
- Do we understand expectations of younger generation?
- Can we see current trends"?
 - Behavior
 - Technology
- Can we choose technology appropriately?



fedora



Upcoming Generation of Scholars

- Age 10 - play
 - Yahoo (music)
 - Google (bios, animals)
 - Neo pets (community)
 - Powerpoint (expression)

- Age 20 - social and study
 - Blogging
 - IM
 - Google
 - BitTorrent
 - Craigslist



fedora



Implications of Web 2.0

- Upcoming generations of scholars will have a completely different paradigm and expectations regarding technology
 - Collaborative classification (e.g., flickr)
 - The power of collective intelligence (amazon)
 - Alternative trust models (reputation like in ebay and in the open source software movement)
- Plan now for technologies to be service-oriented, flexible, decomposed



The logo for Fedora, consisting of the word "fedora" in a lowercase, sans-serif font.

Must move from evolutionary to revolutionary...

- Until recently, repositories have been positioned as an evolution of the traditional publishing paradigm
- Mostly a document-centric paradigm
 - Submit documents
 - Store documents
 - Access documents
 - Archive documents
- Focus on descriptive metadata with controlled vocabularies
 - Human metadata creation and indexing to promote search and discovery
 - Some citation analysis to understand relationships of documents

fedora



New Contexts ! - user and technical

User Contexts

semantic digital library

e-scholarship

scholarly publication

e-science

Technical Contexts

service-oriented

web 2.0

semantic web

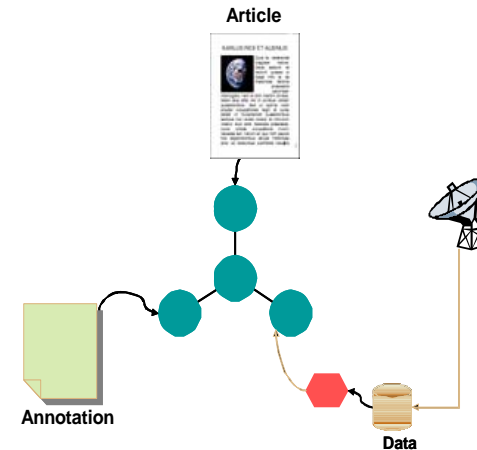
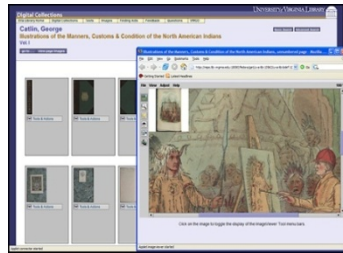
fedora



New positioning for repositories and digital libraries

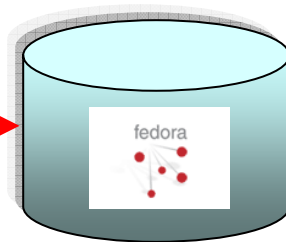
Scholarly and Scientific Workbenches

Linking Data and Publications

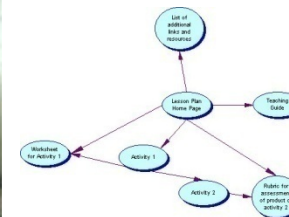


“Web 2.0” Collaborative Repositories

Museum Exhibits with Lesson Plans



blog and wiki



The logo for Fedora, featuring the word "fedora" in a lowercase, sans-serif font above a stylized network diagram of red nodes connected by grey lines.

fedora

Goals for the New Order...

1. Support the creation and publication of new forms of *"information units"*
 2. Knowledge integration: capturing semantic and factual relationships among information entities
 3. Promote information *re-use* and *contextualization*
 4. Accomodate information that is created as a byproduct of *collaborative activity*
(note Web 2.0 and Library 2.0 concepts)
-
1. Integrate with institutional and enterprise *processes and services* that support research, collaboration, and scholarly communication (e.g., workflow)

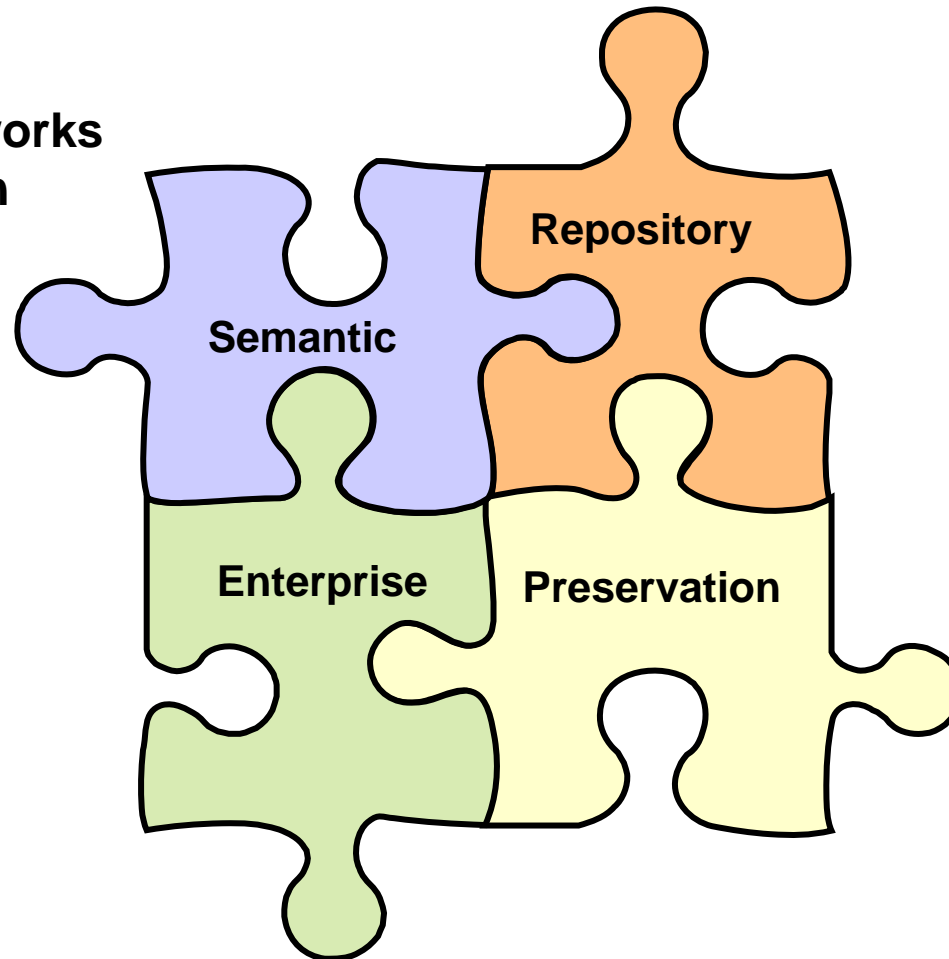
fedora



Requirement: Core Technology Integration

- Information Networks
- Contextualization
- Relationships
- Query
- Inference

- Workflow
- Messaging
- Transactions
- Replication



- Digital Objects
- Manage
- Access
- Versioning
- Storage

- Integrity Check
- Monitoring
- Alerting
- Migration





fedora



What about Semantic Web
and "Web 2.0" Technologies?





PLoS ONE and Topaz Open Access Publishing and Collaboration

PLoS ONE : Publishing science, accelerating research - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://www.plosone.org/home.action

PLoS PATHOGENS Advancing our Understanding of Microbial Pathogenesis

PLoS one BETA Publishing science, accelerating research

Search PLoS ONE... GO

Home About For Users For Authors and Reviewers PLoS Journals PLoS.org

Open Access Join PLoS Submit Today

PARTNERS UnitedLayer fedora science commons

Recently Published Most Annotated

- Acceptance of Anti-Retroviral Therapy among Patients Infected with HIV and Tuberculosis in Rural Malawi Is Low and Associated with Cost of Transport
- Factors Affecting Frequency Discrimination of Vibrotactile Stimuli: Implications for Cortical Encoding
- Modularity and Intrinsic Evolvability of Hsp90-Buffered Change
- Requirement of TORC1 for Late-Phase Long-Term Potentiation in the Hippocampus
- Genetic Influences on Exercise Participation in 37,051 Twin Pairs from Seven Countries

What is PLoS ONE?
A new way of communicating peer-reviewed science and medicine.

- Annotations - add and share your comments
- Discussions - join the conversation
- More functionality coming soon

Your feedback will help us shape PLoS ONE.

Gene Specific Product Directory

Explore by Subject

snag it - Google S... PLoS ONE : Publish... 11 Microsoft Offi... 4 Microsoft Powe... Microsoft PowerPo... Snagit 7 4:00 PM

fedora

NSDL Data Repository (NDR): *An Information Network Overlay*

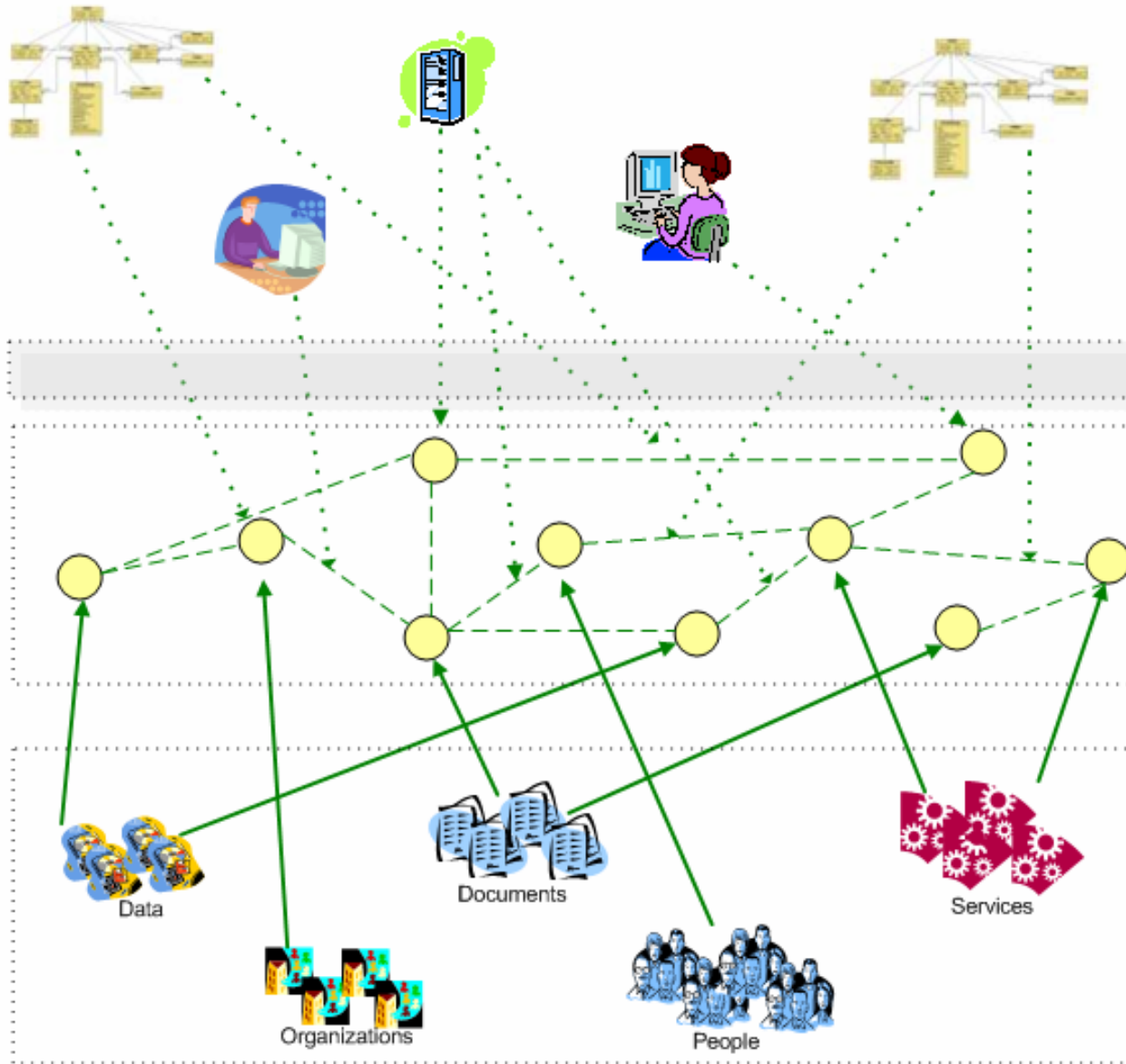
External
Contributors

Access
Controlled
API

Information
Network
Overlay

Primary
Resources

NDR



fedora



NSDL 2.0

Expert Blogs on top of Fedora-based NDR

The screenshot shows the NSDL Expert Voices website. At the top left is the 'EV NSDL' logo with the tagline 'Expert Voices: Speaking of Something Interesting'. A search bar is located at the top right. Below the header, the breadcrumb 'NSDL.org > Expert Voices' is visible. The main content area is titled 'WhoSays' and is divided into four categories: 'K-12 TEACHERS', 'UNIVERSITY FACULTY', 'LIBRARIANS', and 'NSDL LIBRARY BUILDERS'. Each category contains an introductory blog post titled 'NSDL Expert Voices Intro' dated 'Nov. 2nd, 2005 9:35 pm' and written by 'admin'. The text of the first post is 'Welcome to the NSDL Expert Voices blogging project!'. A red arrow points to this text, which is also circled in red. Each post includes a 'Read more >' link and a small image. On the right side, there is a sidebar with a link 'Return to the Annual Meeting Blog', a notice 'Coming to a browser near you in Jan'06', a 'Categories' list (Current News, Education, Health, Mathematics, Science, Social Studies, Technology), and a 'Top Resources' section.

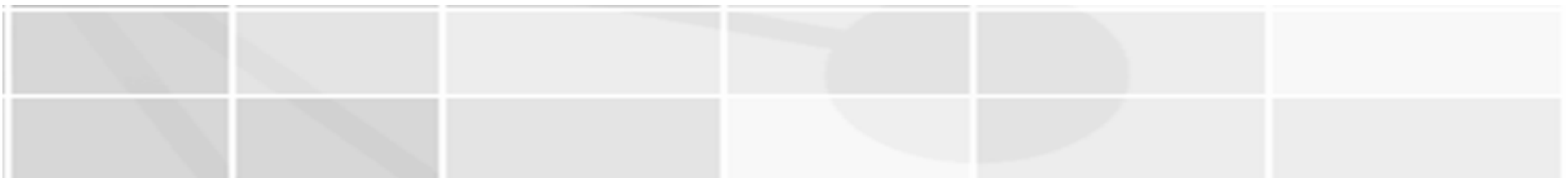
fedora



Semantic Technologies: How can RDF help us?



- Naturally extensible... just assert a new relationships and properties as needed
- Integrate information from different sources... just index all the RDF data together
- RDF allows many points of view
- Enables complex queries and inference





Semantic Modeling Basics

- What is an ontology?
 - A data model that represents some domain-specific reality
 - Objects, their properties, and their relations are modeled
 - Enables reasoning
- Ontologies can be community-oriented
 - Agreement on semantics to meet specific needs
 - Different communities can have their own models
 - But, ultimately we can integrate different perspectives!
- Ontology markup and representation languages:

- RDF and RDF Schema
- OWL

The logo for Fedora, featuring the word "fedora" in a lowercase, sans-serif font above a stylized network diagram of red nodes connected by white lines.

fedora

Semantic Modeling Basics: RDF Schema

- RDF Schema is a simple “meta” vocabulary used to describe ontologies
 - Class, subclassOf, type
 - e.g., Person, Team
 - Property, subPropertyOf
 - e.g., playsFor
 - Domain (the **class for the subjects** of a particular *property*)
 - **Person** *playsFor* Team
 - Range (the **class for values** of a particular *property*)
 - Person *playsFor* **Team**



The logo for Fedora, consisting of the word "fedora" in a lowercase, sans-serif font.

Semantic Modeling Basics: RDF Schema

- RDF Schema is a meta vocabulary can be used to define other domain-specific vocabularies
 - Descriptive metadata vocabulary (e.g., Dublin Core, MODS)
 - See Simile:
<http://simile.mit.edu/wiki/Ontologies>
 - Friend of Friend (FOAF) vocabulary
 - Digital Object Relationships
 - See Fedora:
<http://www.fedora.info/definitions/1/0/fedora-relsext-ontology.rdfs>





RDF: Start with real-world statements



- Statements:

- "The resource <http://www.cs.cornell.edu/payette> was created by Sandy Payette."
- "The SemDL Tutorial was accepted by the event WWW2007."
- "The Fedora semantic digital library session is part of the SemDL Tutorial."
- The Fedora semantic digital library session is presented by Sandy Payette.

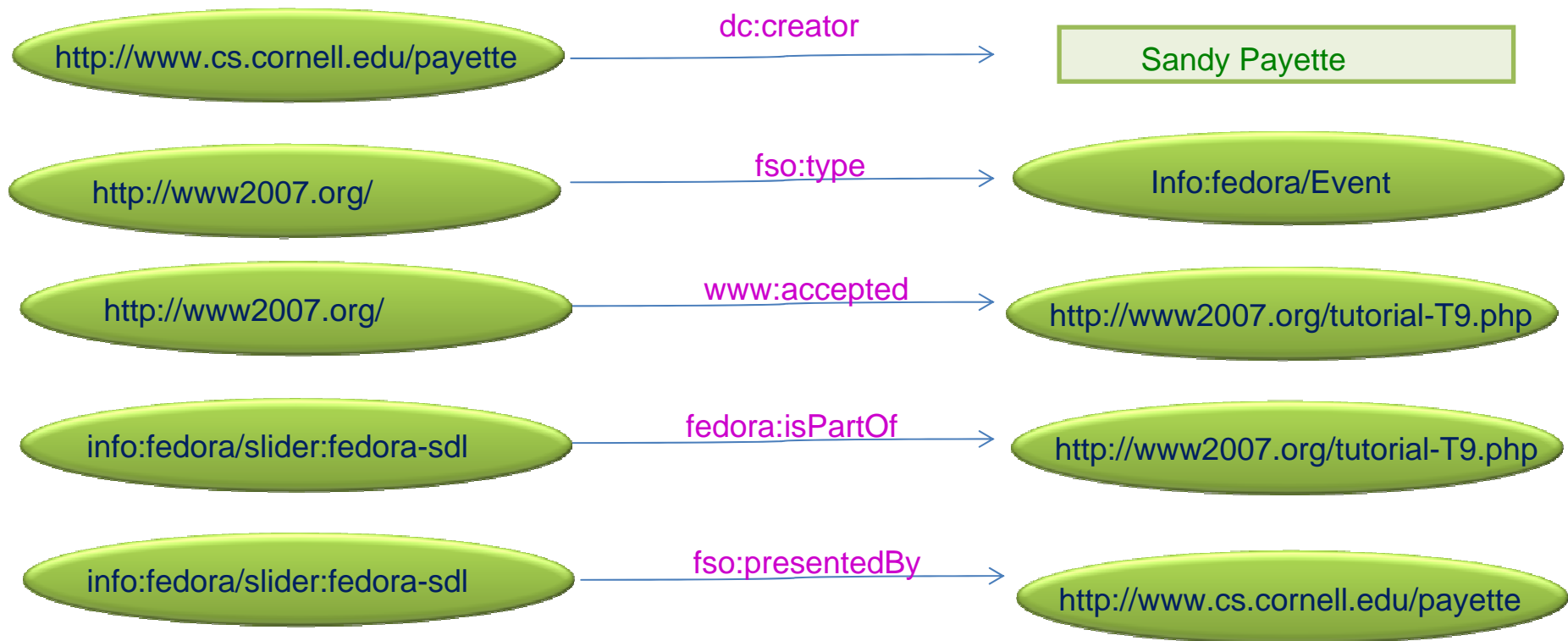


RDF: Express statements as “triples”

Resource (subject)

Property (predicate)

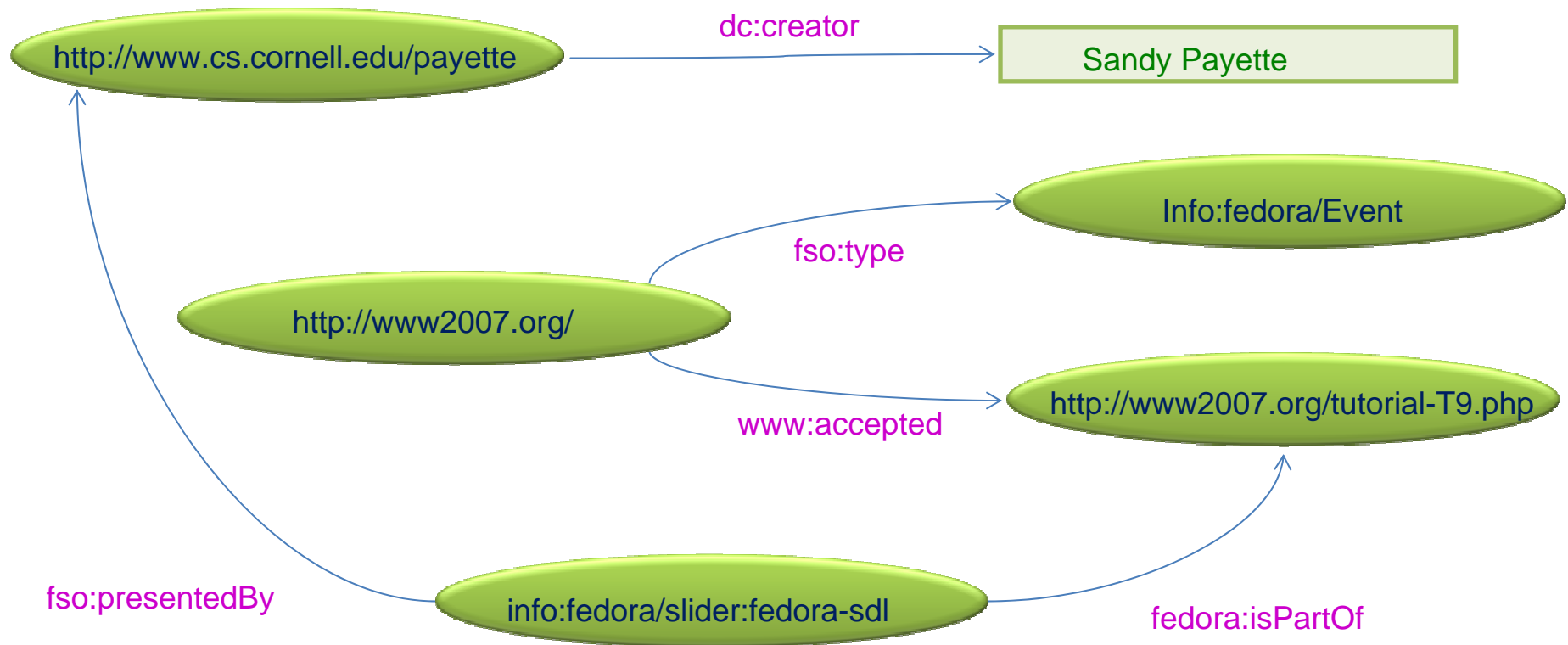
Value (object)



fedora



RDF: Index into graph (triplestore)



The logo for Fedora, featuring the word "fedora" in a lowercase, sans-serif font above a stylized network graph with several red nodes and connecting lines.

fedora

RDF: Query the graph

What events did Sandy Payette present at?

<http://www2007.org/>

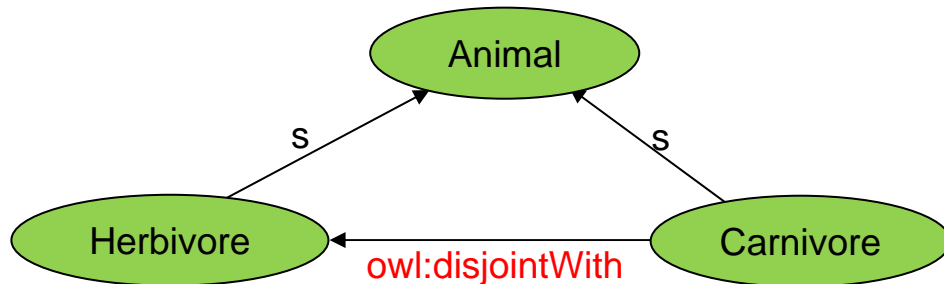


fedora



Semantic Modeling: Web Ontology Language (OWL)

- A richer vocabulary for defining classes, their properties and their relationships among classes.



- Builds on RDF and RDFS
- W3C Recommendation





fedora



Case Study: Fedora for Semantic Digital Libraries



The logo for Fedora, featuring the word "fedora" in a lowercase, sans-serif font above a stylized network diagram of red nodes connected by thin lines.

fedora

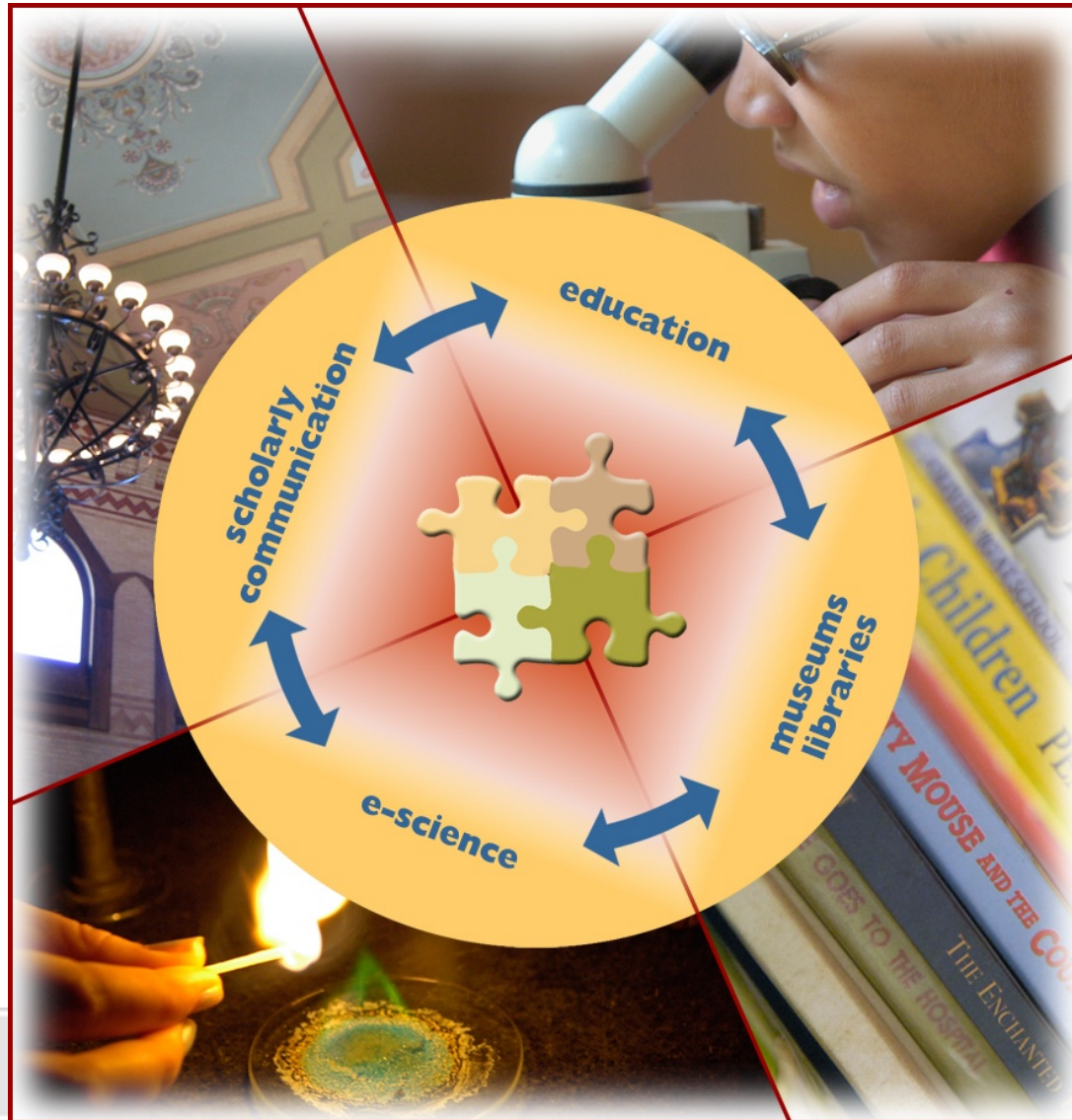
Fedora and Fedora Commons

- Fedora
 - Flexible
 - Extensible
 - Digital
 - Object
 - Repository
 - Architecture
- History
 - Cornell Research (1997-2002)
 - DARPA and NSF-funded research and reference implementations
 - Distributed, Interoperable Repositories (experiments with CNRI)
 - Open Source Project (2002-present)
 - Andrew W. Mellon Foundation (2002-2009)
 - Joint development by Cornell University and University of Virginia
 - Fedora Commons non-profit organization (Summer 2007)

fedora



Fedora - Multiple Communities



fedora





Arts and Humanities

The Arts and Humanities Data Service | tel: 020 7943 1988 | email: info@ahds.ac.uk

ahds arts and humanities data service

Select a subject area:

Search Collections

Enabling Digital Resources for the Arts and Humanities

The Arts and Humanities Data Service (AHDS) is a UK national service aiding the discovery, creation and preservation of digital resources in and for research, teaching and learning in the arts and humanities. Currently, we cover five subject areas:

Archaeology, History, Visual Arts, Literature, Languages & Linguistics, Performing Arts

The Arts and Humanities Data Service is funded by JISC

Tufts University Libraries & Archives

Tisch Library is the main library on the Medford/Somerville campus. With a shelving capacity of a million volumes, the library supports the teaching and research mission of the Schools of Arts & Sciences and Engineering. The library has an extensive film and microform collections, a selective depositary of US federal government documents, and the collection is a full and ever-growing resource to the entire University community. The library subscribes to more than 300 online databases and over 15,000 print and electronic journals. A call, conference and group study rooms, as well as a Media Lab for viewing DVDs and videotapes are also available. Tufts also is a charter member of the Boston Library Consortium and participates in its Virtual Catalog, which allows users to search for and directly request books from 19 area college and university libraries.

The Edwin Ginn Library, located in the Cabot Intercultural Center on the Medford/Somerville campus, reflects the specialized curriculum and research interests of The Fletcher School. The library collects books, journals, and documents, in print and online format, in the fields of international law and

Perseus
Tufts is home to the Perseus Digital Library, the world's largest online database of Latin and Greek texts and archeological finds. The digital collection contains 6.5 million words of Greek text, 4 million words of Latin material, more than 64,000 images and 3-D representations of ancient Greek and Roman sites. Receiving up to 420,000 hits a day from more than 75 countries, Perseus aims to increase accessibility of primary and secondary source-texts, and bring a wide range of materials to a larger audience.

The Perseus Digital Library

Gregory Crane, Editor-in-Chief, Tufts University

Perseus is an exciting digital library, engineering interactions through time, space and language. Our primary goal is to bring a wide range of source materials to as large an audience as possible. We anticipate that greater accessibility to the sources for the study of the humanities will strengthen the quality of questions, lead to new answers, lead to new avenues of research, and connect more people through the connection of ideas.

Classics in a Digital World

Curious about where classics might go in a digital world? See the [recent](#) of a [new article about ePhilology](#) that will appear in The Blackwell Companion to Digital Library Studies.

National Library of Scotland

catalogue • news • collections • info for users • catalogues • digital library • professional info • search

Catalogue Search:

Look for a book
Search our main catalogue for books, maps, music and more

Search other catalogues
Further online resources

Site Search:

Browse site by subject
An A to Z list of key subjects

Site search
Find information using keyword search or site map

An information treasure trove for Scotland

The National Library of Scotland is an information treasure trove of Scotland's knowledge, history and culture, with millions of books, manuscripts and maps covering every subject.

Maps and engravings in special editions
Ayrshire student wins bookbinding prize
Saltaire Society Literary Awards

In the spotlight

An Arts and Crafts treasure from Proobbe Anna Traquair
James Clerk Maxwell tops our Scottish scientists poll

Quick Links:

Contact/enquiries
Celebrating 50 years NLS strategy
John Murray Archive
Support NLS
Freedom of information

Help For:

First-time site visitors
Researchers
Librarians
Publishers
Teachers
Community groups
People with disabilities
Donors

Spoken Word at Northwestern

Spoken Word Research

A substantial portion of Western cultural heritage in the 20th century rests in spoken-word form. Only a tiny fraction of the vast content has been transcribed, and even that is inadequate as the spoken word - and not a transcription of it -- remains the best source. While text versions may capture the words spoken, they cannot capture the emotional or ambient qualities that provide insight into events and participants.

Even today, access to audio still remains limited to physical visits to archival collections or secondary sources contained in films and video played for passive learners in class. For example, vast portions of the secret recordings of American presidents John F. Kennedy, Lyndon B. Johnson and Richard M. Nixon are accessible to the public only through actual visits to archival collections throughout the United States. Unfortunately, such public release does not translate to public access today.

The National Gallery for the Spoken Word (NSGW), Historical Voices, The Oyst Project, and BBC - Information - Archives have digitized and accessioned thousands of hours of spoken word resources into digital repositories. This project will build on the current research in digital library development, much of which was enabled through NSF or JISC support.

As with any primary source, these materials do not literally "speak" for themselves and impart wisdom; they require interpretation and analysis. We will create a set of tools to enable rigorous analysis for research using the new content. Because digital audio libraries now permit fast and simple audio segmentation, researchers can now use online tools to listen and locate a precise section of audio within a larger audio file. Our project will build on this

Digital Collections

Catlin, George

Illustrations of the Manners, Customs & Condition of the North American Indians

View 1

Illustrations of the Manners, Customs & Condition of the North American Indians, unannotated page

Click on the image to toggle the display of the image/viewer. Tool menu bars

ENCYCLOPEDIA of CHICAGO

CHICAGO HISTORY MUSEUM

CHICAGO HISTORY MUSEUM, THE NEWBERRY LIBRARY, AND NORTHWESTERN UNIVERSITY

EDITORS: JANICE L. REIFF, ANN DURKIN KEATING, AND JAMES R. GROSSMAN

Search

Full List

ChicagoHistoryMuseum

ENCYCLOPEDIA of CHICAGO

CHICAGO HISTORY MUSEUM, THE NEWBERRY LIBRARY, AND NORTHWESTERN UNIVERSITY

EDITORS: JANICE L. REIFF, ANN DURKIN KEATING, AND JAMES R. GROSSMAN

SEARCH

Full List

ChicagoHistoryMuseum

ENCYCLOPEDIA of CHICAGO

CHICAGO HISTORY MUSEUM, THE NEWBERRY LIBRARY, AND NORTHWESTERN UNIVERSITY

EDITORS: JANICE L. REIFF, ANN DURKIN KEATING, AND JAMES R. GROSSMAN

SEARCH

Full List

ChicagoHistoryMuseum



Sciences



HOME PAGE	Homepage English Deutsch
THE PROJECT ESCIDOC	Welcome to the eSciDoc project!
PLANNING AND STATUS	eSciDoc is as a shared project of the Max Planck Society and FIZ Karlsruhe, funded by the Federal Ministry of Education and Research (BMBF), with the aim to realize a platform for communication and publication in scientific research organizations.
ORGANIZATION	
BACKGROUND	
MATERIALS, DOWNLOADS	
CONTACT	
IMPRINT	

AWI
Siftung Alfred-Wegener-Institut für Polar- und Meeresforschung in der Helmholtz-Gemeinschaft

About us | Research | Resources | Click&Learn | Ships&Stations | Press | Jobs

deutsch

Headlines

New group of algae discovered: Picobiliphytes

An international group of researchers has succeeded in identifying a previously unknown group of algae. As currently reported in the scientific journal *Science*, the newly discovered algae are found among the smallest members of photosynthetic plankton - the picoplankton (Picobiliphytes: A marine picoplanktonic algal group with unknown affinities to other Eukaryotes" *Science*, Vol. 316"). On account of the minute size of the organisms (no more than a few thousandth of a millimetre) and the appearance of phycoobiliproteins, researchers have termed the new group Picobiliphytes. [Read more about Picobiliphytes...](#)

How fish species suffer as a result of warmer waters

Ongoing global climate change causes changes in the species composition of marine ecosystems, especially in shallow coastal oceans. This applies also to fish populations. A new investigation, just published in the scientific journal *Science*, reveals that a warming induced deficiency in oxygen uptake and supply to tissues is the key factor limiting the stock size of a fish species under heat stress. [Read more about science article...](#)

PLoS ONE - Publishing science, accelerating research - Mozilla Firefox

PLoS ONE PATHOGENS Advancing Our Understanding of Microbial Pathogenesis

PLoS one 2006

Search PLoS ONE...

Home About For Users For Authors and Reviewers PLoS Journals PLoS.org

Open Access
Join PLoS
Submit Today

What is PLoS ONE?

- Annotations - add and share your comments
- Discussions - join the conversation
- More functionality coming soon

Your feedback will help us shape PLoS ONE.

Explore by Subject

- Research**
- Climate Sciences
 - Geosciences
 - Biosciences
 - New technology
 - Research program
 - Cooperations/Societies

- Resources**
- Observatories
 - Technology Tran
 - Databases/Softw
 - Library

- Click&Learn**
- Sealce for Begin
 - Animal Migration
 - School Projects

Memorial Sloan-Kettering Cancer Center

HOME ABOUT US HOW TO HELP NEWSROOM CALENDAR LOCATIONS JOBS CONTACT

Search: Find a Doctor:

Establishing standards of care for the treatment of cancer

Laboratory and Clinical Research:

- Sloan-Kettering Institute
- Memorial Hospital Research

Treatment Information by Cancer:

Select a Type of Cancer:

Graduate Education

Learn more about our PhD and MD/PhD programs in the biomedical sciences [more >>](#)

NEWSROOM

- New Online Search Tool Helps Scientists Explore Genes in Medical Literature
- Oncology Nurse Receives National Honor
- Ludwig Fund Gives \$120 Million for Cancer Research

FOR THE NEWLY DIAGNOSED

Getting the correct diagnosis and the most appropriate treatment from the start is crucial [more >>](#)

MAKE A GIFT

Support our programs of patient care, research, and education [more >>](#)

Memorial Sloan-Kettering Cancer Center | Sloan-Kettering Institute
Louis V. Gerstner, Jr., Graduate School of Biomedical Sciences, Memorial Sloan-Kettering Cancer Center

EV NSDL
Expert Voices: Speaking of Something Interesting

NSDL.org > Expert Voices

WhoSays

47 TEACHERS

NSDL Expert Voices Intro
Nov. 2nd, 2005 9:35 pm
Written by: admin
Welcome to the NSDL Expert Voices blogging project! [Read more >](#)

UNIVERSITY FACULTY

NSDL Expert Voices Intro
Nov. 2nd, 2005 9:35 pm
Written by: admin
Welcome to the NSDL Expert Voices blogging project! [Read more >](#)

LIBRARIANS

NSDL Expert Voices Intro
Nov. 2nd, 2005 9:35 pm
Written by: admin
Welcome to the NSDL Expert Voices blogging project! [Read more >](#)

NSDL LIBRARY BUILDERS

NSDL Expert Voices Intro
Nov. 2nd, 2005 9:35 pm
Written by: admin
Welcome to the NSDL Expert Voices blogging project! [Read more >](#)

Return to the Annual Meeting Blog

Coming to a browser near you in Jan'06

Categories

- Current News
- Education
- Health
- Mathematics
- Science
- Social Studies
- Technology

Top Resources

Resource Title
Resource Description

Education



fedora

Fedora's RDF-enabled Repository

- **RDF provides a natural model for exposing repository as graph**
 - Object-to-object relationships
 - Networks of digital objects
 - Relationships to external entities
 - Query the graph; traversal to discover related stuff
- **RDF provides a generalizable, extensible data model**
 - Graph-based data model is a common reduction
 - Avoid fixed schema problems and metadata mud wrestling
 - Extensible enrichment of object descriptions
 - Freedom to add and interleave statements from multiple ontologies
 - Organic evolution
- **Powerful queries and inference for repository management**
 - Transitive relationships among objects
 - Dependency analysis;
 - Detection/Extraction of sub-graphs
 - Provenance of disseminations

fedora

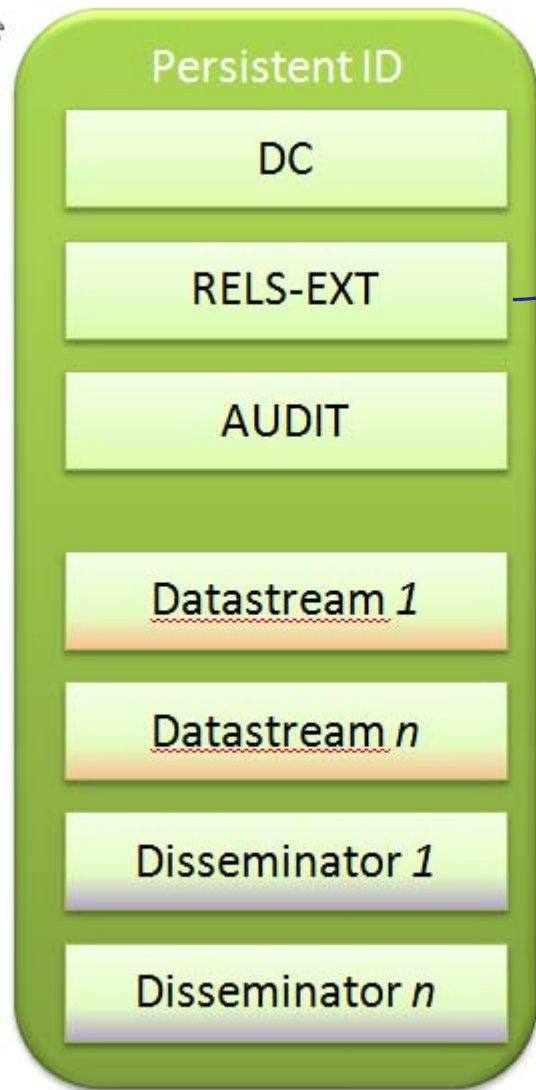


Fedora Digital Objects

- Flexible object model can support
 - Documents, articles, journals
 - Electronic Scholarly Texts
 - Digital Images
 - Complex multimedia publications
 - Datasets
 - Metadata
 - Learning objects
 - More...
- Create "networks" of objects using RDF
 - Define object relationships and other properties via RDF
 - Collection/member; part/whole; etc.



Fedora Digital Object Model with RDF



Reserved Datastreams

Custom Datastreams (Managed or By-Reference)

Service Binding Metadata for Dynamic Views
(optional)

The logo for Fedora, consisting of the word "fedora" in a lowercase, sans-serif font.

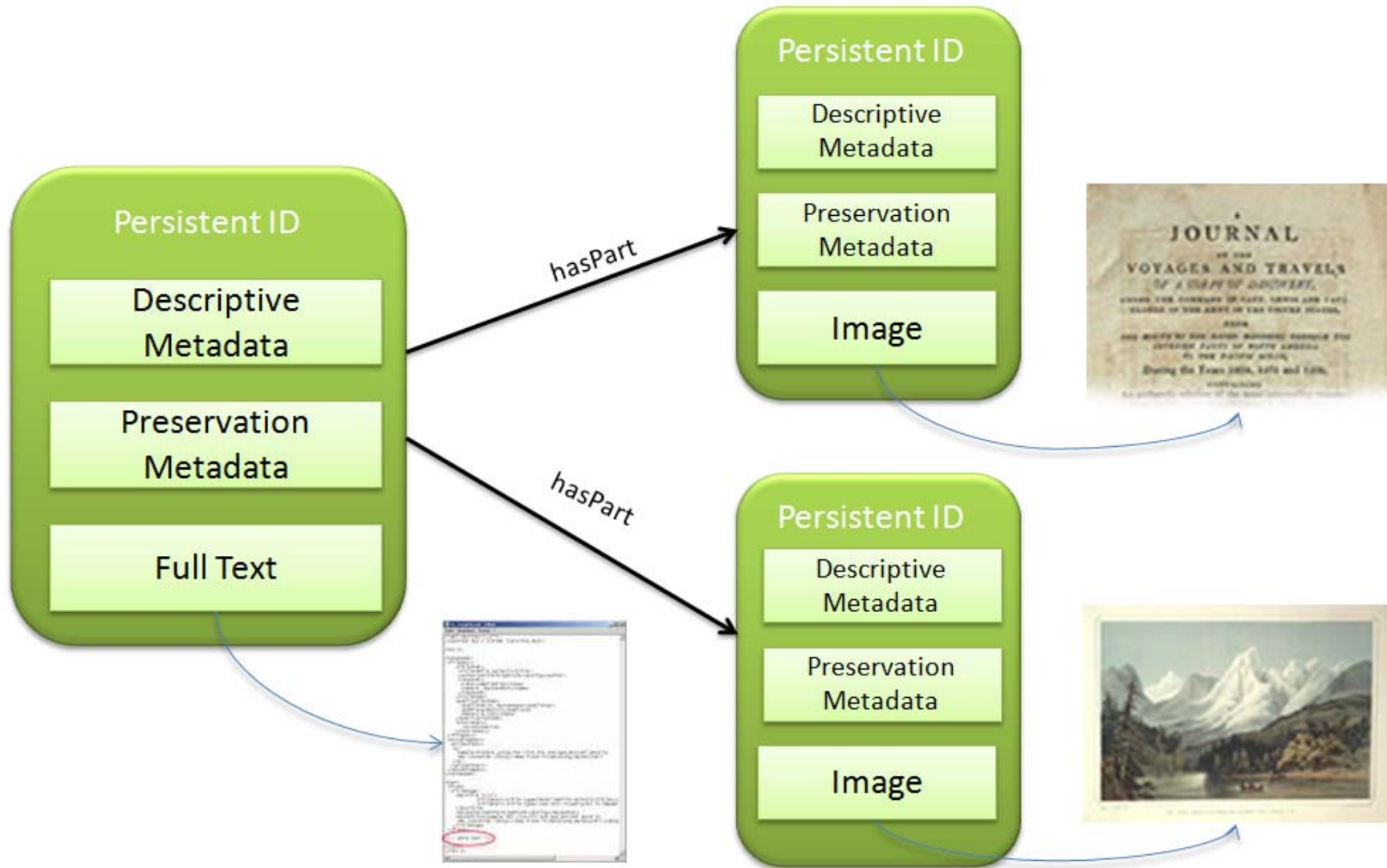
Digital Objects contain RDF assertions

- **Assert relationships from Fedora base ontology**
 - Collection - member
 - Whole - part
 - Equivalence
 - Description Of
 - More...
- **Assert relationships/properties from community ontologies**
 - isAnnotationOf
 - isRecommendedBy
 - isCertifiedBy
 - More





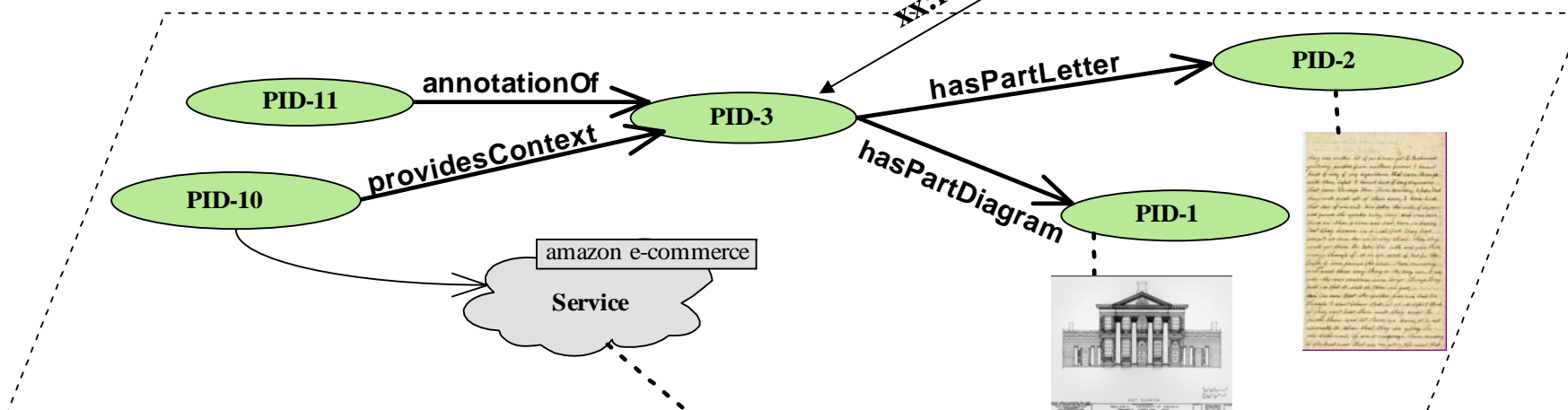
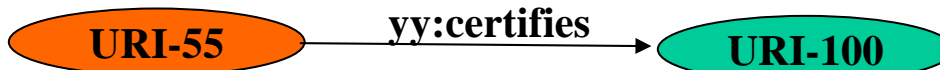
Digital Objects with "compositional semantics"



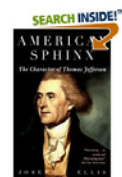
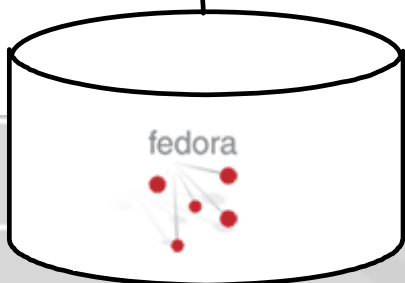


Use Case: scholarly objects and annotation in humanities

scholarly objects



library and museum objects



American Sphinx: The Character of Thomas Jefferson (Vintage) (Paperback)
 by Joseph J. Ellis "IT WAS A PROVINCIAL version of the grand entrance..." (more)
 \$19.95 natural rubber section: neonational appearance: pure nubilation: midtime: judata.
 CAPA: American Revolution: Declaration of Independence: Continental Congress: John Adams: New England (more)
 ★★★★★ (104 customer reviews)

List Price: \$15.95
 Price: \$10.95 & eligible for FREE Super Saver Shipping on orders over \$25. [See details](#)
 You Save: \$5.10 (32%)

Availability: Usually ships within 24 hours. Ships from and sold by Amazon.com. [See more on this item.](#)

Want it delivered **Monday, December 19**? Order it in the next 4 hours and 29 minutes, and choose **One-Day Shipping** at checkout. [See details](#)

72 used & new available from \$7.39

commercial web content



3 Objects - 3 RDF "Relationships" Datastreams

```
<rdf:Description rdf:about="info:fedora/uva:pid-11">  
  <ais:annotationOf rdf:resource="info:fedora/uva:pid-3"/>  
</rdf:Description>  
</rdf:RDF>
```

```
<rdf:Description rdf:about="info:fedora/uva:pid-10">  
  <ais:providesContextFor rdf:resource="info:fedora/uva:pid-3"/>  
</rdf:Description>  
</rdf:RDF>
```

```
<rdf:Description rdf:about="info:fedora/uva:pid-3">  
  <uva:hasPartLetter rdf:resource="info:fedora/uva:pid-2"/>  
  <uva:hasPartDiagram rdf:resource="info:fedora/uva:pid-1"/>  
</rdf:Description>  
</rdf:RDF>
```

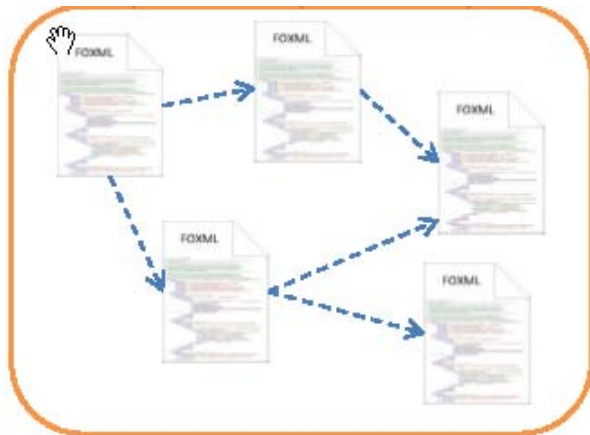
fedora



Fedora RDF-based Resource Index (RI)

- NOT the core object store - RI is a graph-based index of the repository
- Automatic, incremental indexing into triplestore
- Search/query the repository via Fedora RI Query Interface

Digital Object Store



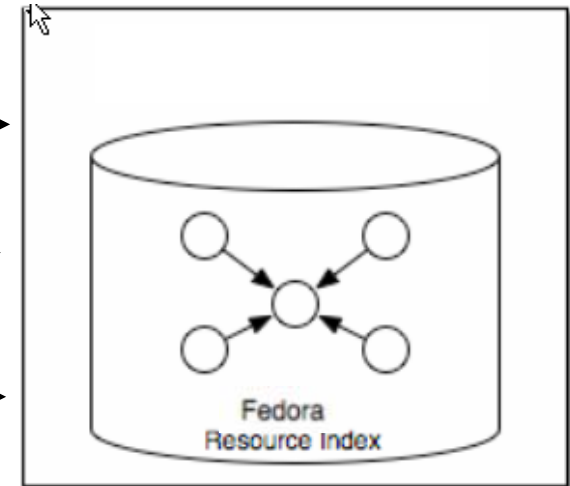
Fedora Repository

RDF datastream

Fedora object properties

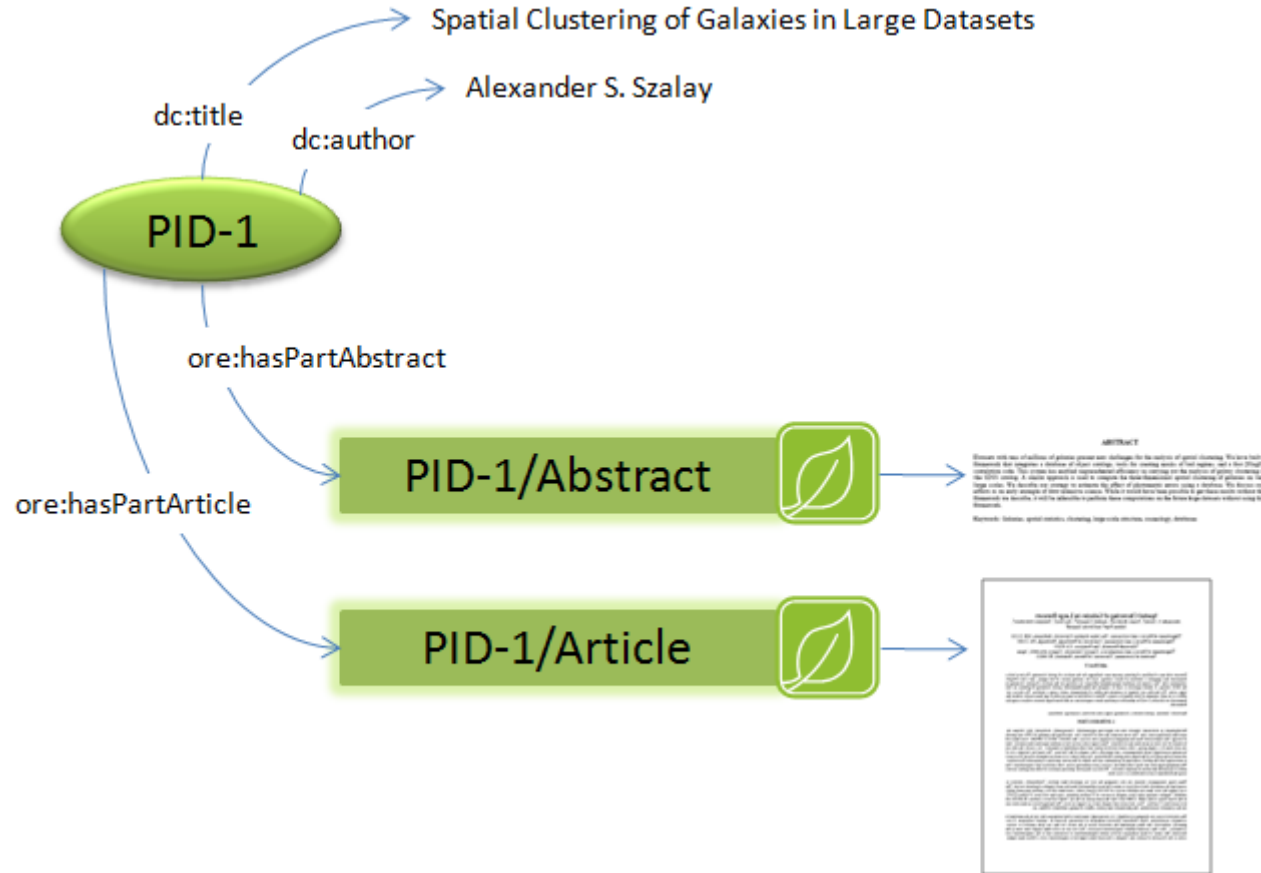
DC datastream

RDF Index of Repository



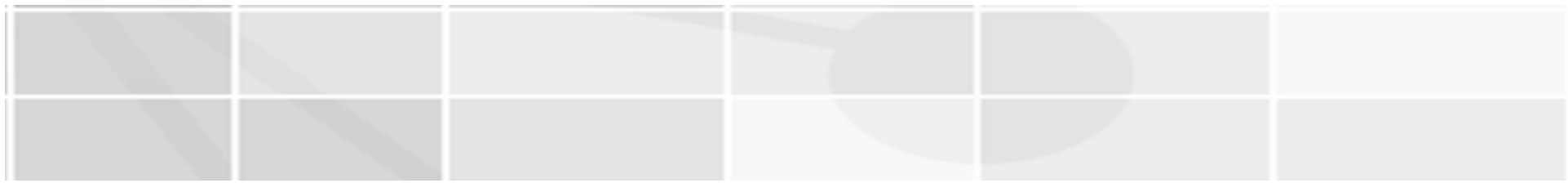
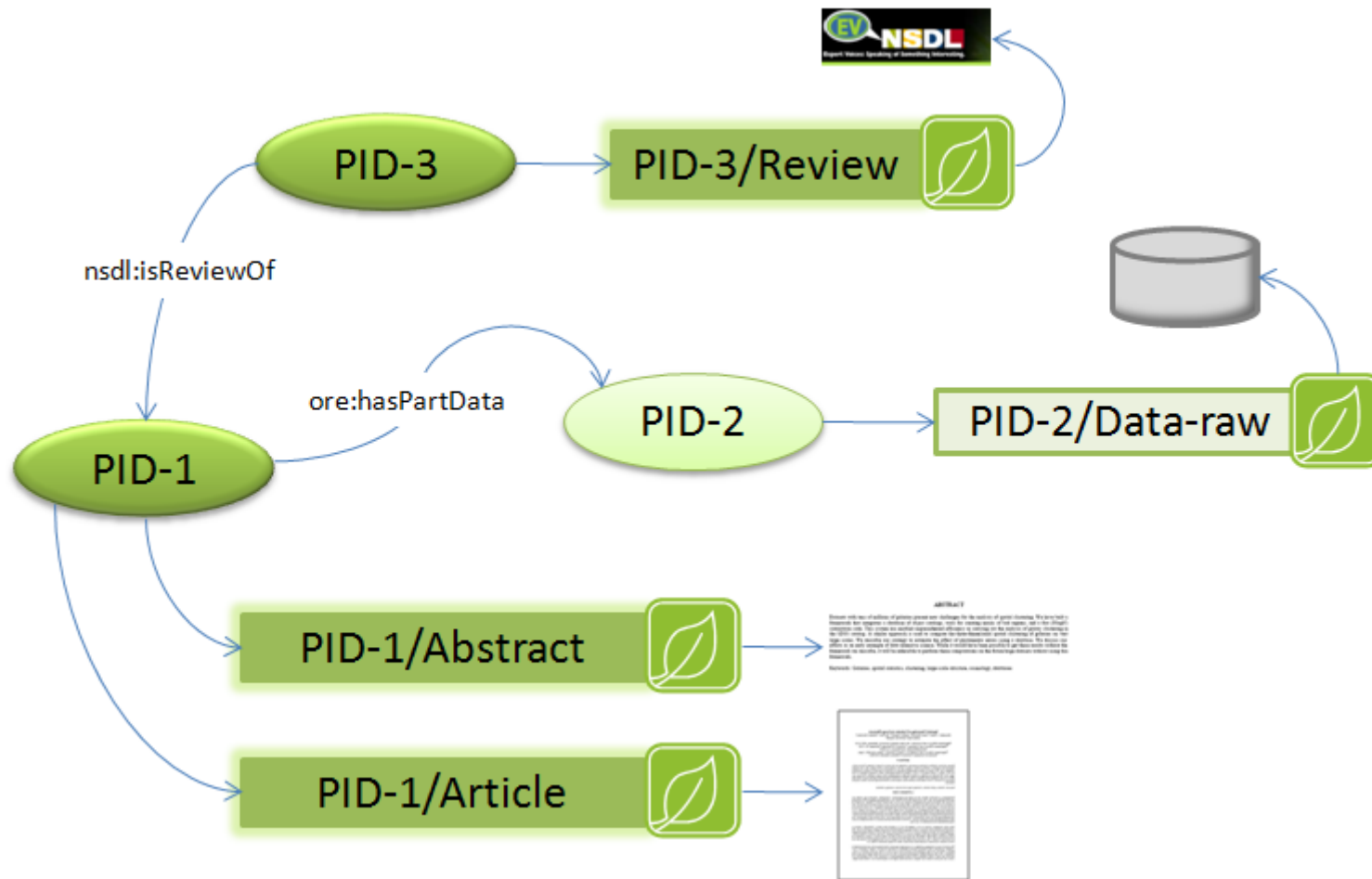


Use Case: Scholarly publication...





Enrich with datasets and reviews...

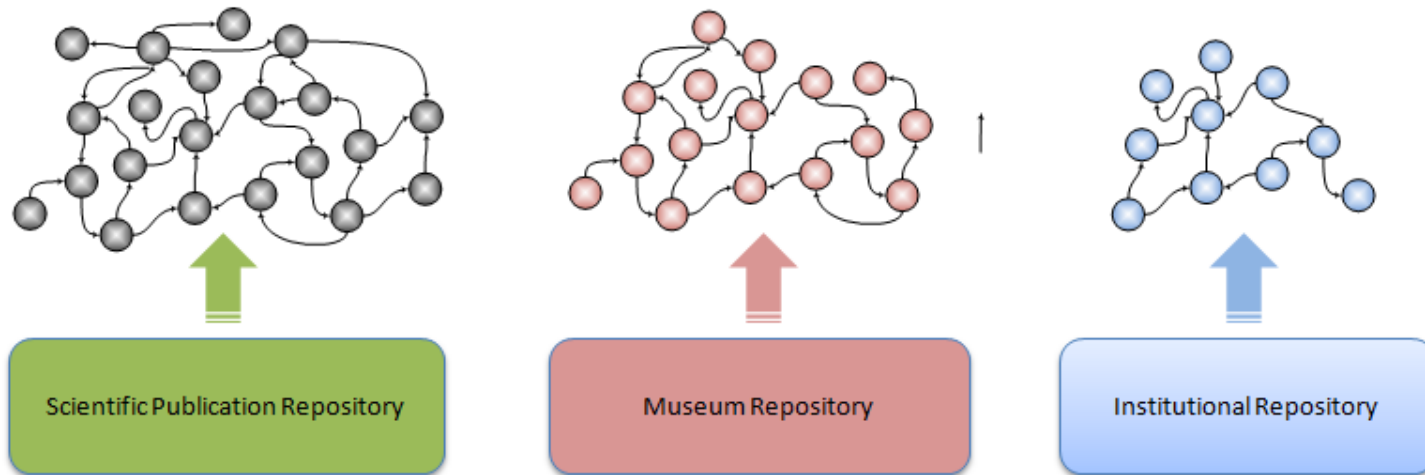




Looking Ahead: OAI-ORE Object Re-use and Exchange



Graph Integration and Interoperability



<http://www.openarchives.org/ore/>

fedora



Triplestore Challenge

- Scalability
 - Few triplestores perform well for 100M+ triples
 - Kowari - we tested to 180M triples
 - MPTStore - we tested to 250M triples
- Performance
 - Jena - easy to get out of memory
 - Sesame Native - slow for complex queries
 - Kowari
 - Fast queries and full-featured query language (iTQL)
 - Instability and corruption problems
 - MPTStore
 - Very fast for SPO queries (limited support for complex queries)
 - Add/modify significantly faster than Kowari
 - Mulgara
 - Fork of Kowari; complex queries; models; inference
 - Major bug fixes to fix stability and corruption problems
 - XA2 transactions
 - Claims support for billions of triples



fedora



Quick demo: "Slider" repository





fedora



Fedora Repository and Service Framework

Robust Middleware for Digital Libraries



fedora

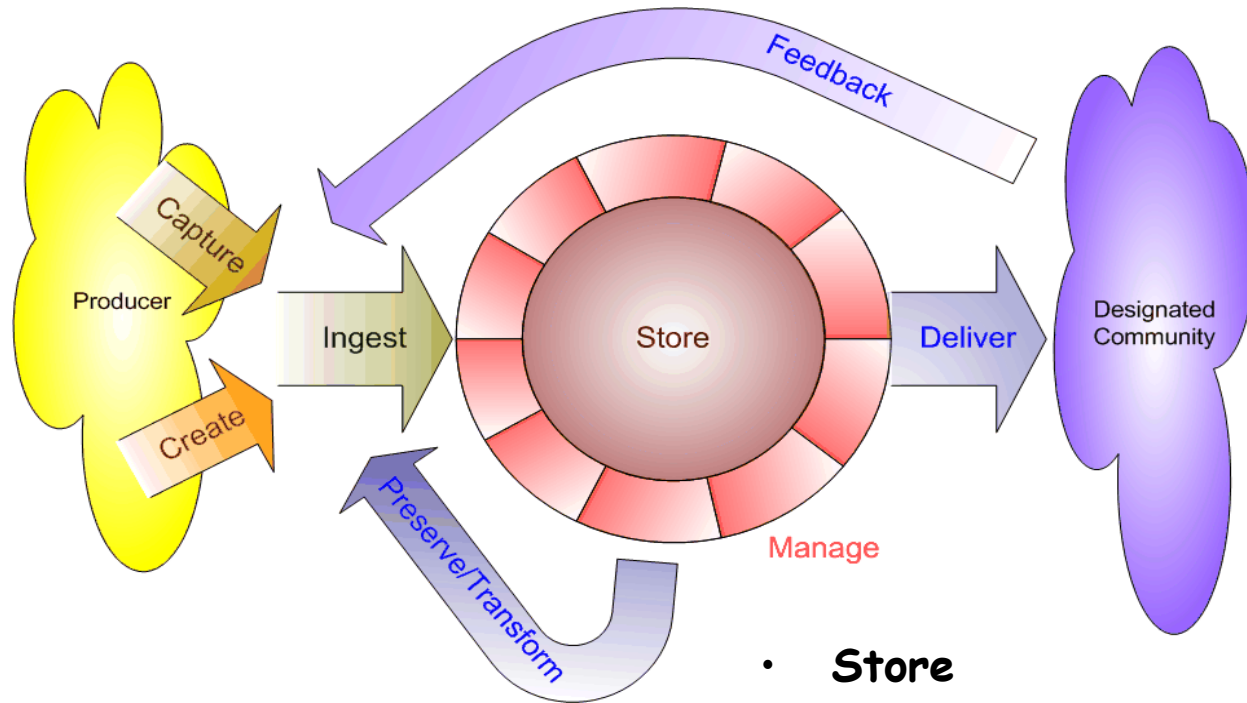


Focus on Full Information Lifecycle

For structured, unstructured or semi-structured data:

- Augment
- Annotate
- Enable collaboration

- Capture
- Create
- Ingest



- Share
- Reuse
- Fuse
- Repurpose
- Connect

- Store
- Manage
- Deliver
- Preserve
- Ensure security and integrity

Courtesy of Dan Davis, Harris Corp.

fedora



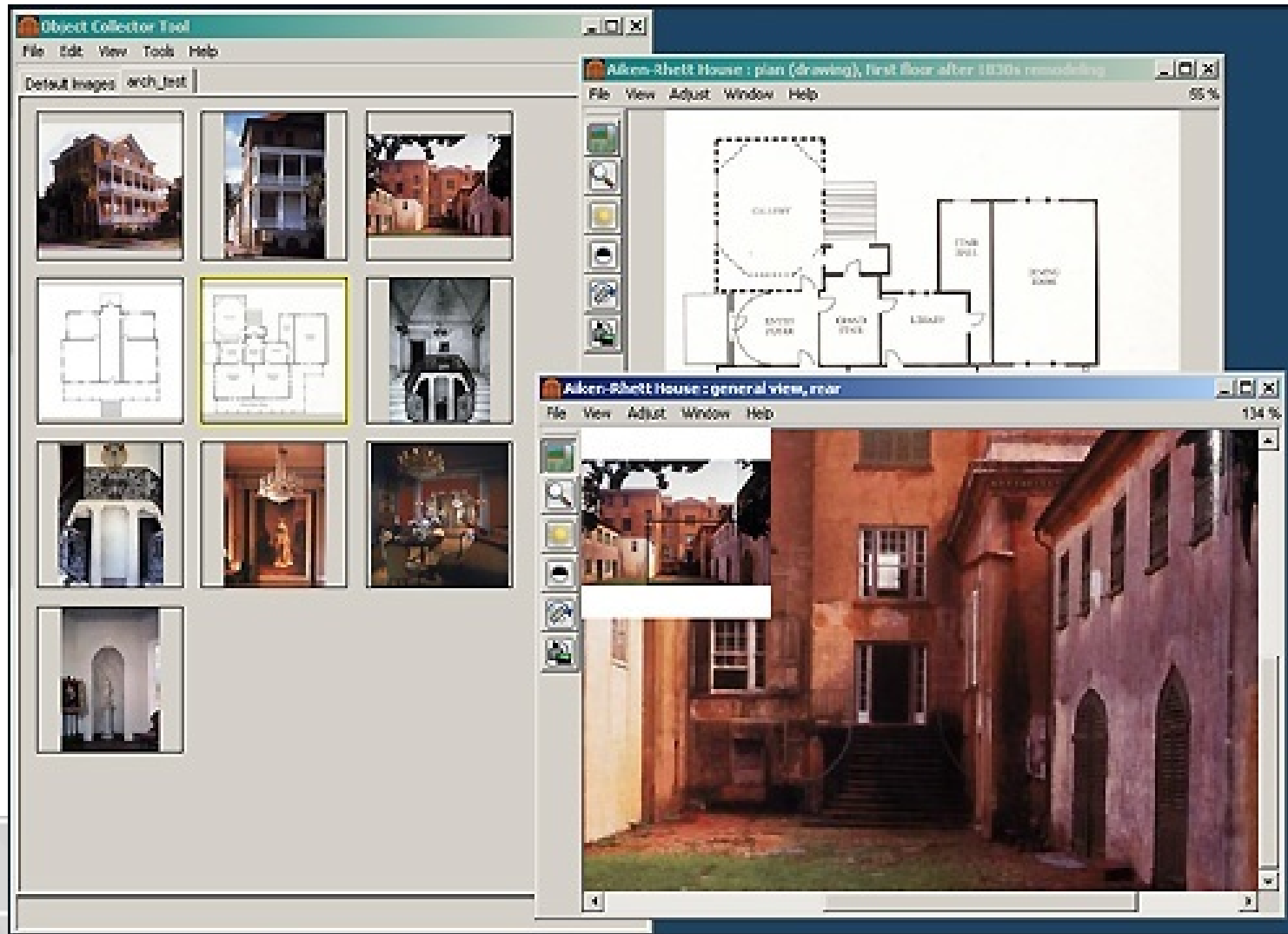
University of Virginia Image Collections

A screenshot of a web browser displaying the University of Virginia Digital Collections website. The browser window title is "Untitled Document - Microsoft Internet Explorer". The website header includes "UNIVERSITY OF VIRGINIA LIBRARY" and a navigation menu with links: "UVA Library Home", "Digital Collections", "Texts", "Images", "Finding Aids", "Feedback", "Questions", and "YBGO". The main content area is titled "Search Results" and shows "1 - 25 of 62 results". There are three tabs: "images & files", "images only", and "files only". The first search result is for "Colonial Bell Chapel". It features a thumbnail image of a stone fountain with a dragon sculpture. Below the thumbnail is a menu with options: "Get Information", "Open in Image/Viewer", "Add to Collection", "Download Image", and "Get Citation". To the right of the main search results, there is a larger view of the same fountain image, with a toolbar containing icons for zooming and other image manipulation functions. The browser's address bar at the bottom shows the URL "uva.lib:79647/uva.lib:79646".

fedora



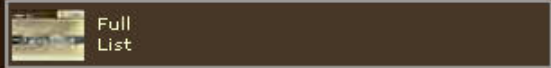
UVA "Collector Tool"



ENCYCLOPEDIA of CHICAGO

Entries|Historical Sources|Maps|Special Features|User's Guide

SEARCH



ENCYCLOPEDIA of CHICAGO

CHICAGO HISTORICAL SOCIETY, THE NEWBERRY LIBRARY, AND NORTHWESTERN UNIVERSITY

EDITORS: JANICE L. REIFF, ANN DURKIN KEATING, AND JAMES R. GROSSMAN



The Electronic Encyclopedia of Chicago © 2005 Chicago Historical Society.
The Encyclopedia of Chicago © 2004 The Newberry Library. All Rights Reserved. Portions are copyrighted by other institutions and individuals. Additional information on copyright and permissions.

ENCYCLOPEDIA of CHICAGO

Entries | Historical Sources | Maps | Special Features | User's Guide

SEARCH



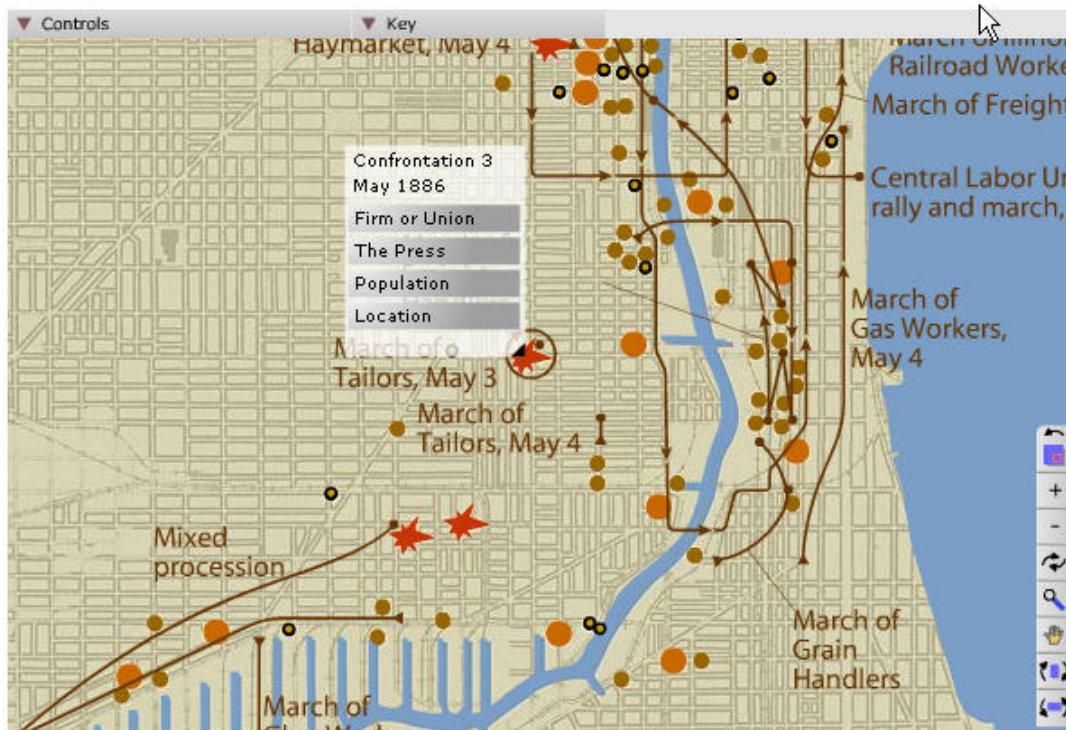
SEE ALSO

- Eight-Hour Movement
- Haymarket and May Day
- Street Life
- Strikes
- Unionization

HISTORICAL SOURCES

RICH MAP : LABOR UNREST IN CHICAGO, APRIL 25 - MAY 4, 1886

Labor Unrest in Chicago, April 25-May 4, 1886 | The Press and Labor in the 1880s | Manufacturing Climate | Chicago's Social Geography



Shouting Amazons,
Chicago Tribune, 4 May 1886 | Page 1 | Page 2
Transcription | Related Newspaper Stories

SHOUTING AMAZONS.
WOMEN EMULATE THEIR BROTHERS AND GO ON STRIKE.

Hundreds of Them Quit Work and Parade the Streets Amid Much Enthusiasm—How the Day Was Spent in Various Parts of the City Where Labor Troubles Exist—Concessions Made by the Brewers—City Streets Left Impassable by Disappointed Workmen.

Between 300 and 400 girls and women were affected with a malignant form of the eight-hour melody yesterday morning. An hour before noon a number of them who have been employed in the tailor shops about Division and Sedgwick streets gathered together, formed into a procession, and marched around to the different establishments, calling on those who work to join them.

The advance of the strike was evidently so from the fact, however, approximately them. 'a little out of a Elevators "A" all the four models. In the Chicago, near Eleventh cars loaded with a lot of work by the fitting work demanded no in a day's labor is in they work for by the hour, a door, and blind ran and Clayton morning, having pay for eight of the men led by men from this factory and work. The patrol crowd were released. The

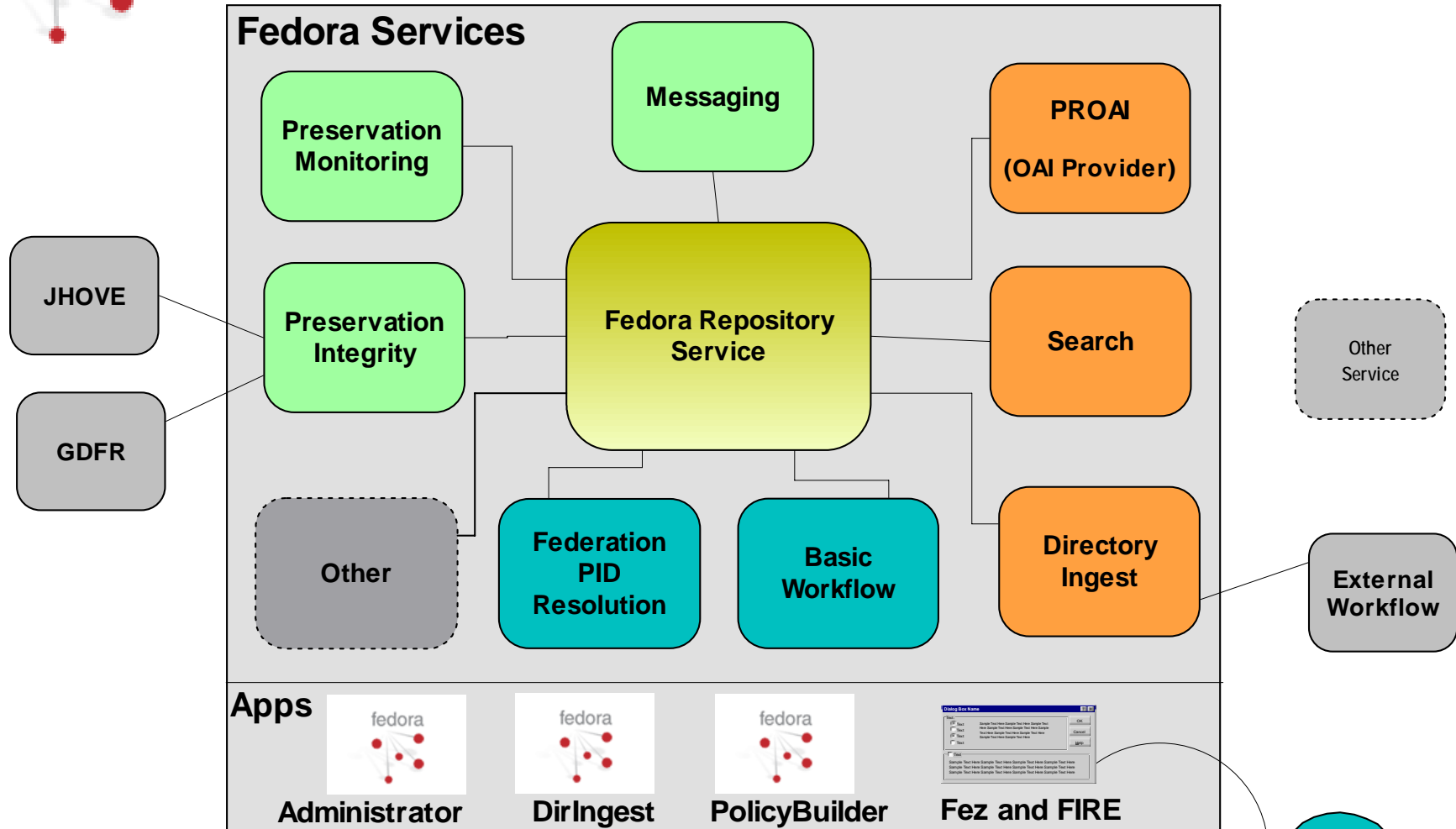
now out on strike who end of lumber rather than to and Mr. Van Schuyler has thing to the contrary.

Concessions by the The Brewers' Association yesterday International Breweries, John H. McAvoy president of the union and the publication in yesterday's over at length and included the demand that the statute a working day Sunday instead of 1. The advance of was amounts to \$5 a not to employ men to every twenty be any pressure on the their boarding piece beer for all the employees. The employer each individual to \$2.50, the hours for drink 2, and 4 with a final time. Every matter



Fedora Service Framework

SOA

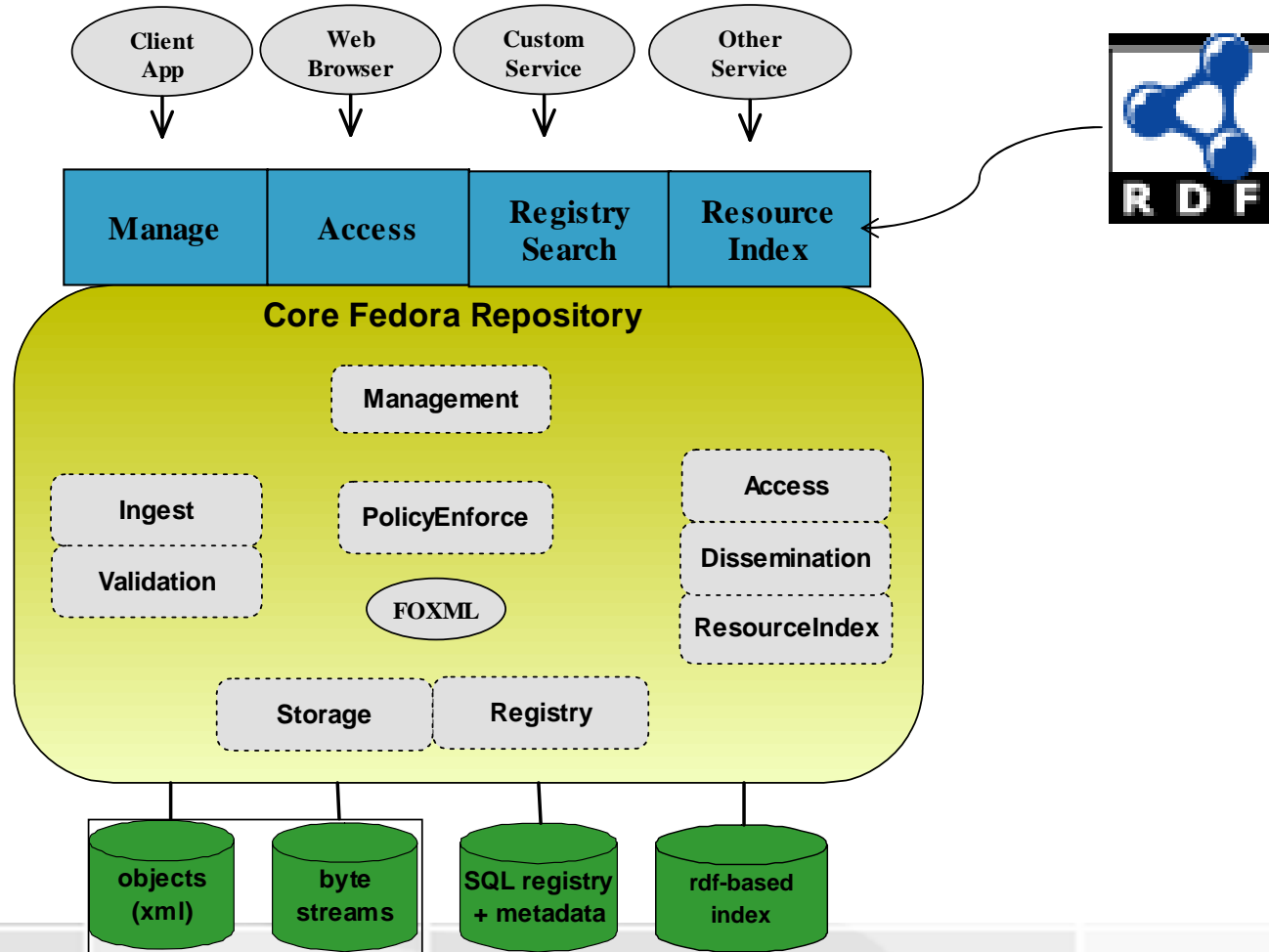


Institutional Repository Clients



At the Core: Service-Oriented Repository

Web Service Exposures





Fedora Service Framework

Supporting Services

- Directory Ingest Service
 - Submit a directory of files with a METS manifest
 - Hierarchy preserved as object-to-object relationships
- OAI Service
 - Based on PROAI
 - Multiple metadata formats; oai sets
 - Configure to harvest Fedora dissemination types
- Fedora Search Service (GSearch):
 - Configure to index datastreams or disseminations
 - Lucene or Zebra



The logo consists of the word "fedora" in a lowercase, sans-serif font, positioned above a stylized network diagram. The diagram features several red circular nodes connected by thin, light-colored lines, suggesting a distributed or interconnected system.

fedora

Fedora Service Framework Supporting Services

- Repository Replication (Journaling)
 - Set up a "following" repository by replaying all operations upon the primary repository
- Message Brokering
 - Publish and subscribe
 - Core repository service can publish events
 - Services can subscribe and listen to events
 - Services can publish their own events






fedora

Fedora Service Framework

Supporting Services - Upcoming

- Validation and Integrity
 - Bytestream format validation
 - Content model validation
- Preservation - Monitoring and Alerting
- Preservation - Migration
- Basic Workflow



fedora



Ongoing Challenges

- Low barrier to entry
 - Simple protocols (e.g., like OAI)
 - Light-weight (REST vs. SOAP?)
 - Simple tools to create overlays
- Service matching (object-to-service)
 - Ontologies to expose objects with formats and semantics
 - OWL-S for semantic service description
 - Matching-making algorithms
- Security and Trust
 - Authentication and trust among repositories and services
 - Interoperability of authorization policy
- Preservation
 - Distributed and dynamic digital objects a challenging *reality*

The logo for Fedora, featuring the word "fedora" in a lowercase, sans-serif font.

Choosing technology to evolve with user needs... now

- Think in terms of flexible service frameworks
- Define fundamental services for libraries
 - Repositories as services
 - Others: workflow, search, monitoring, migration, ...
- Support for complex digital objects
 - Local and remote content
 - Mixed genre → documents, data, images, everything...
 - Dynamic views
 - XML expressions (esp. for ingest/export and migration)
- Model entities with ontology-based metadata
 - RDF allows mix and match of vocabularies
 - Relationships among objects are key - think graphs
 - Organic and evolutionary metadata (freedom from fixed schema)
 - Semantic technologies can enable interoperability, esp. via equivalence assertion and inference

fedora



Discussion and Questions...





fedora



Fedora Web Site

www.fedora.info

Community Open Source Tools

www.fedora.info/tools

Fedora Wiki

<http://www.fedora.info/wiki>

