Fedora Relationships and Information Network Overlays

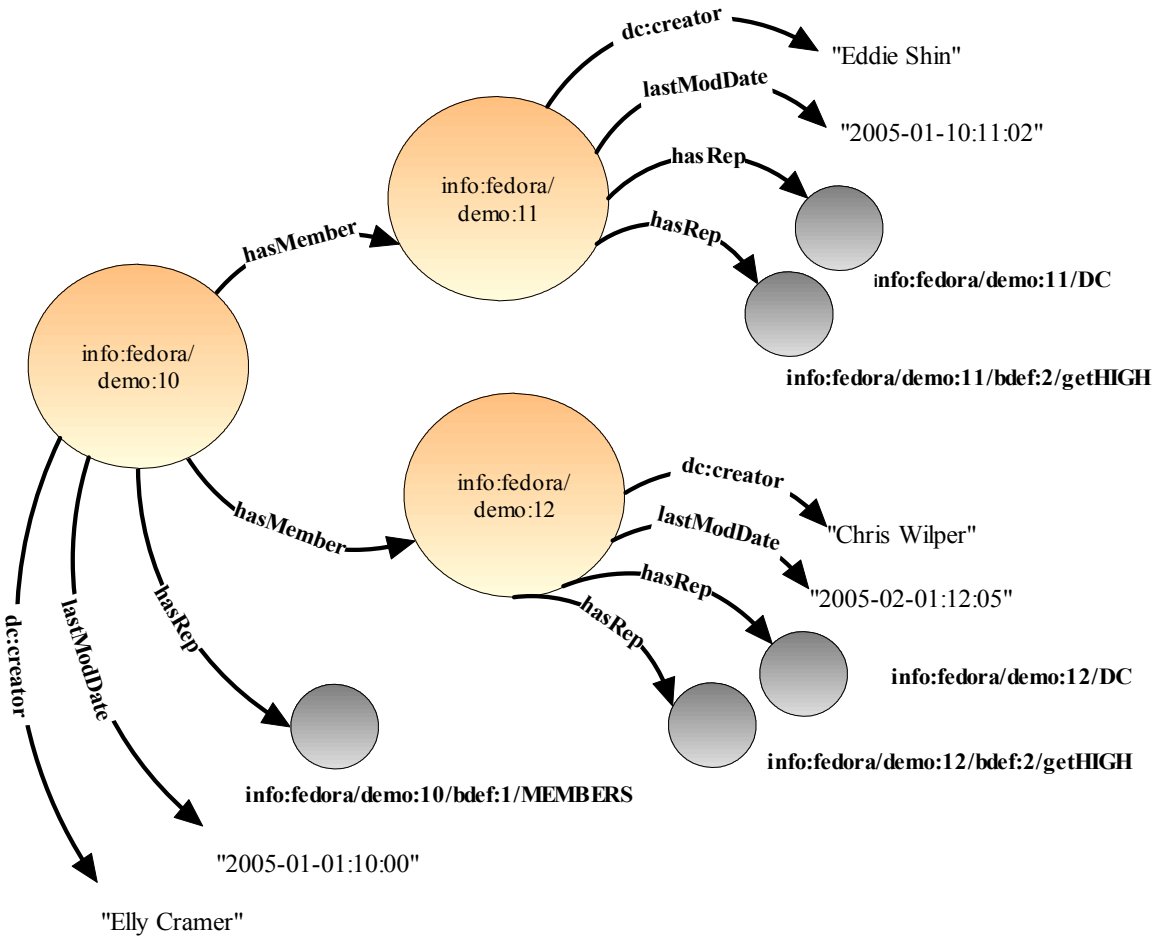
CS 431 – April 19, 2006

Carl Lagoze – Cornell University

Fedora Resource Index:
Using RDF and ontologies

Fedora Digital Objects

Resource Index View



Fedora 2.0 and RDF

- **Object-to-object and object-to-literal Relationships**
  - Ontology of common relationships (RDF schema)
  - Relationships stored in special datastream (RELS-EXT)

- **Resource Index (RI)**
  - RDF-based index of repository (Kowari triple-store)
  - Graph-based index includes:
    - **Object properties and Dublin Core**
    - **Object Relationships**
    - **Object Disseminations**

- **RI Search**
  - Powerful querying of graph of inter-related objects
  - REST-based query interface (using RDQL or ITQL)
  - Results in different formats (triples, tuples, sparql)

Uses of Object Relationships

- Define collections (e.g., collection objects)
- Assert critical relationships among object for management purposes
- Enable network overlay
  - Surrogate objects referring to external entities
  - Assert relationships among them
  - Assert other relationships (e.g., annotations)
- Enable navigation of repository (as tree or graph)

Fedora Relationship Ontology (RDFS)

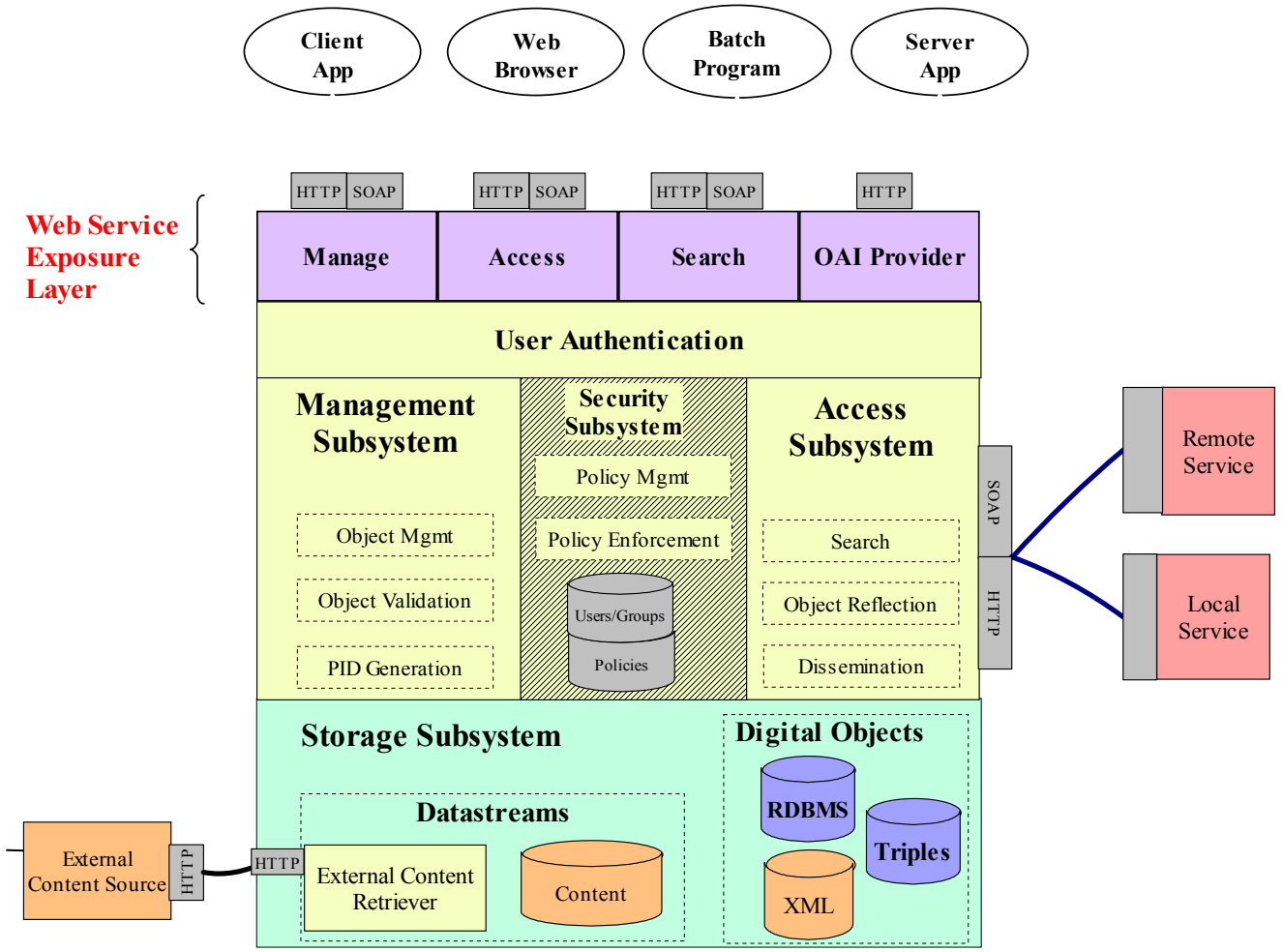
- isPartOf / hasPart
- isMemberOf / hasMember
- isDescriptionOf / hasDescription
- hasEquivalent
- … others

Demo:
Collection – Member Relationships

- Collection Object [smiley]
  - Datastream containing a query to Resource Index for all members of collection
- Image Objects [brush]
  - Use RELS-EXT datastream to assert relationship to collection object

Fedora Repository Service

Fedora Repository Service



Fedora Web Service APIs in a Nutshell

- **Management Service (API-M)**
  - Ingest Object
  - Export Object
  - Get Object XML
  - Purge Object
  - Modify Object
  - Get Next PID
  - Get Datastream(s)
  - Get DatastreamHistory
  - **Get DisseminatorHistory**
  - Get Disseminator(s)
  - Add/modify/purge Datastream
  - Add/modify/purge Disseminator
  - Set State

Fedora Web Service APIs in a Nutshell

- **Access Service (API-A and API-A-LITE)**
  - Describe Repository
  - Get Object Profile
  - Get Object History
  - Get Datastream
  - Get Dissemination
  - Find Objects
  - Resume Find Objects

Fedora Software Distribution

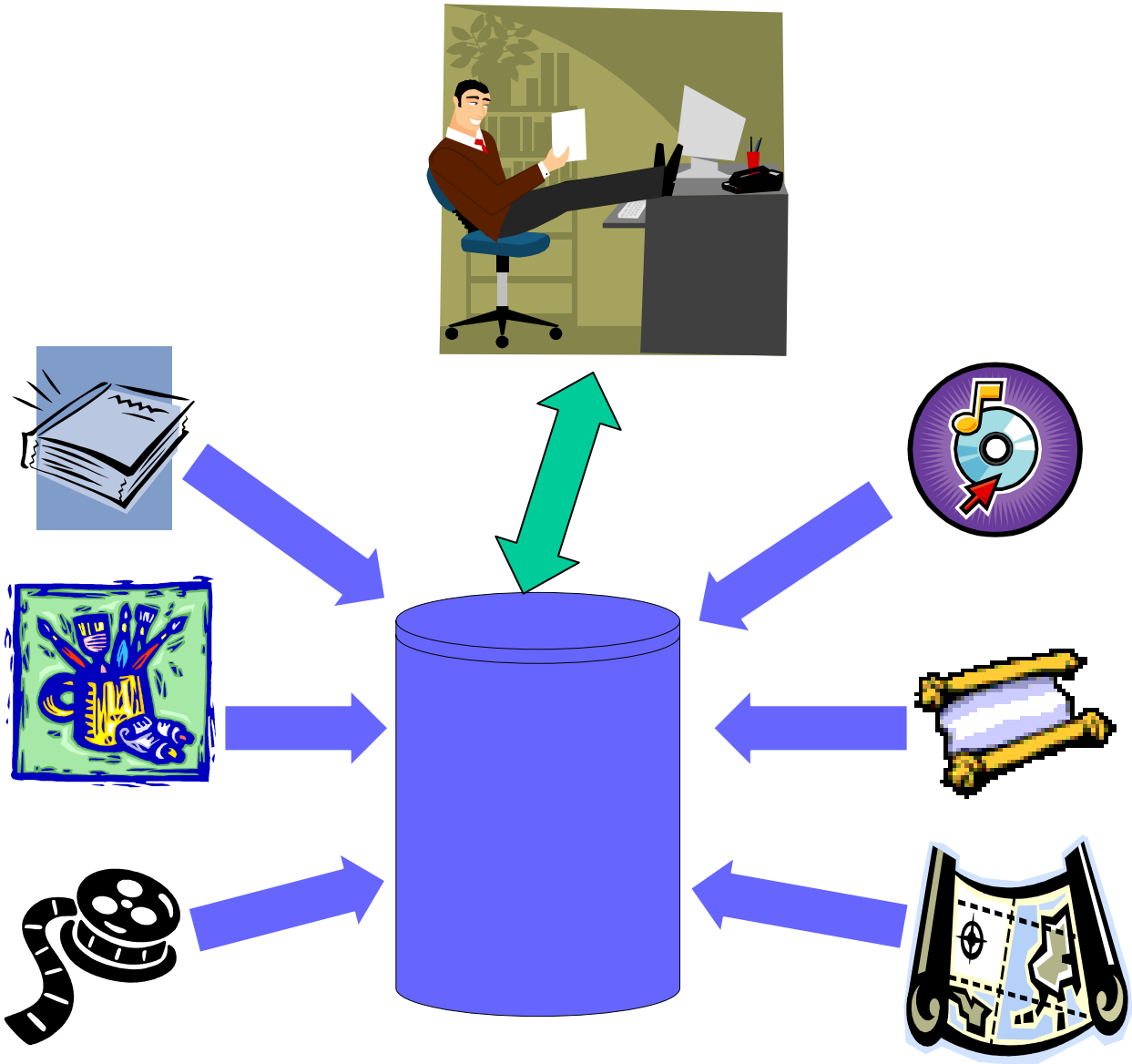
- **Open Source (Mozilla Public License)**
- **100% Java (Sun Java J2SDK1.4)**
- **Supporting Technologies**
  - Apache Tomcat and Apache Axis (SOAP)
  - Xerces for XML parsing and validation
  - Saxon for XSLT transformation
  - Schematron for validation
  - MySQL and Mckoi relational database
  - Oracle 9i support
  - Kowari for triple-store
- **Deployment Platforms**
  - Windows 2000, NT, XP
  - Solaris
  - Linux
  - Mac OSX

What is a digital library anymore, anyway
Information Network Overlays

Executive Summary

- Move beyond the one way information flow
  - catalog->index->search->access
- Enhance primary data with knowledge activity
  - "People as part of the information infrastructure"
    (Dolores Iorizzo)

Digital Libraries – Ingest Focus



Digital Libraries – Federation Phase

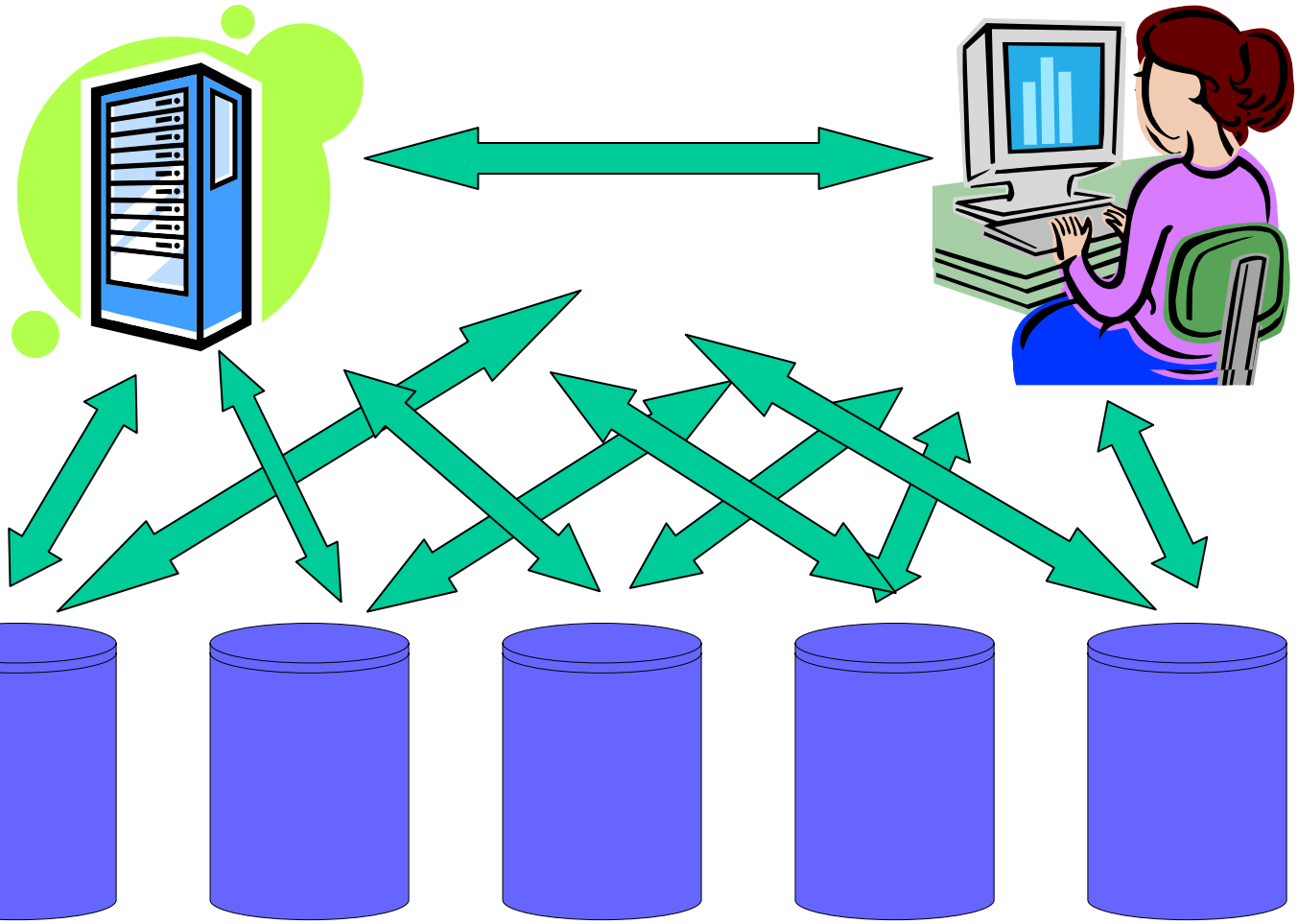
Z39.50

Dienst

SDLIP

OAI-PMH

SRW/SRU



We have been very successful!

So, are we done?

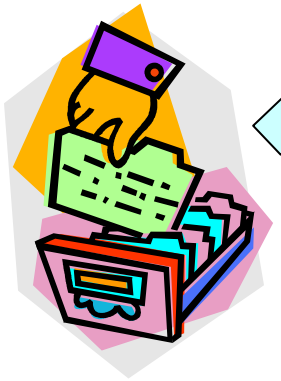
The primary goal of digital libraries has been often been misconstrued as providing accessibility to a massive volume of resources (e.g., Google, institutional repositories). The real opportunity is to:

- Make the library more inclusive of different types of information like data, products, computational services
- Establish the library as a knowledge environment where people organize around information, contribute new information, and collaboratively create new knowledge (i.e., the “wisdom of crowds”).

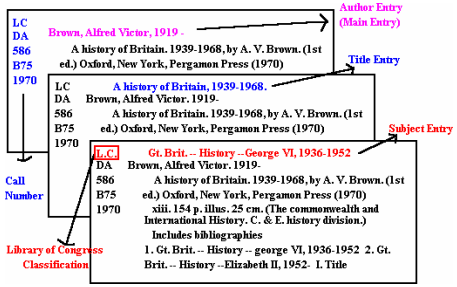
Information Flow in Traditional Library

In-Band

Out-of-Band



Knowledge



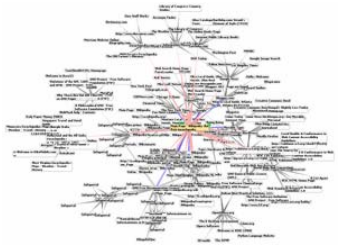
Information Flow in the Digital Library

In-Band

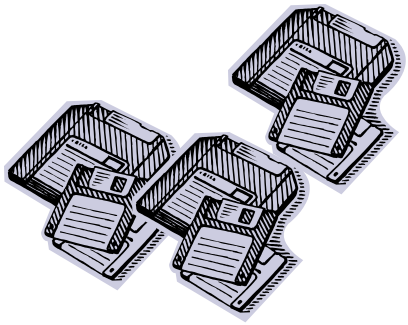
Out-of-Band

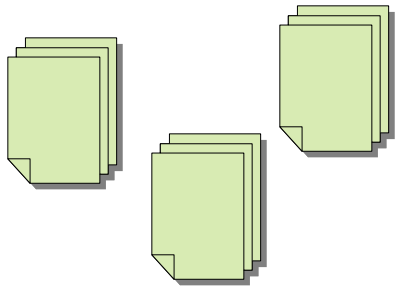
Knowledge

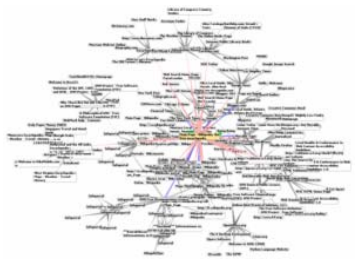
Dublin Core Used Here



Creating a Collaborative Knowledge Network



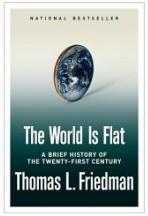




Knowledge Network

Carl, this item is recommended for you. Why?

Join Amazon Prime and ship Two-Day for free and Overnight for $3.99.



Share your own customer images
Search inside this book

The World Is Flat: A Brief History of the Twenty-first Century (Hardcover)

by Thomas L. Friedman "No one ever gave me directions like this on a golf course before: "Aim at either Microsoft or IBM..." (more)

(288 customer reviews)

List Price: ~~$27.50~~
Price: **$16.00** and eligible for **FREE Super Saver Shipping** on orders over $25. See details
You Save: $11.50 (42%)

Availability: Usually ships within 24 hours. Ships from and sold by Amazon.com.

Want it delivered Monday, September 12? Order it in the next 8 hours and 25 minutes, and choose **One-Day Shipping** at checkout. See details

78 used & new available from $13.75

Other Editions	List Price	Price	Other Offers
Paperback	~~$48.00~~	$10.88	
Hardcover (Large Print)	$32.95	$32.95	10 used & new from $37.95
Audio Cassette (Unabridged)	~~$44.95~~	$29.67	41 used & new from $28.21

See all 8 editions...

Add to Shopping Cart
or
Sign in to turn on 1-Click ordering.
Instant Reward Off. Learn how to reactivate.

78 used & new from $13.75
Available for in-store pickup now from $19.25
Price may vary based on availability
Enter your ZIP Code:
Choose a store
Have one to sell? Sell Yours Here

Add to Wish List
Add to Wedding Registry
Don't have one? We'll set one up for you.

Best Value

Buy **The World Is Flat: A Brief History of the Twenty-f...** and get **History: Fiction or Science?** at an **additional 5% off** Amazon.com's everyday low price.

Buy Together Today: $49.20
You Save: $13.25

Buy both and save

Click here for promotion rules

Customers who bought this item also bought

Freakonomics : A Rogue Economist Explores the Hidden Side of Everything by Steven D. Levitt
Blink : The Power of Thinking Without Thinking by Malcolm Gladwell
Collapse: How Societies Choose to Fail or Succeed by Jared Diamond
1776 by David McCullough
The Lexus and the Olive Tree: Understanding Globalization by Thomas L. Friedman
The Tipping Point: How Little Things Can Make a Big Difference by Malcolm Gladwell

Explore Similar Items: in Books, in Music, and in DVD

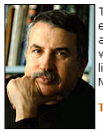
Editorial Reviews

Amazon.com

Thomas L. Friedman is not so much a futurist, which he is sometimes called, as a presentist. His aim, in his new book, *The World Is Flat*, as in his earlier, influential *Lexus and the Olive Tree*, is not to give you a speculative preview of the wonders that are sure to come in your lifetime, but rather to get you caught up on the wonders that are already here. The world isn't going to be flat, it *is* flat, which gives Friedman's breathless narrative much of its urgency, and which also saves it from the Epcot-style polyester sheen that futurists--the optimistic ones at least--are inevitably prey to.

What Friedman means by "flat" is "connected": the lowering of trade and political barriers and the exponential technical advances of the digital revolution have made it possible to do business, or almost anything else, instantaneously with billions of other people across the planet. This in itself should not be news to anyone. But the news that Friedman has to deliver is that just when we stopped paying attention to these developments--when the dot-com bust turned interest away from the business and technology pages and when 9/11 and the Iraq War turned all eyes toward the Middle East--is when they actually began to accelerate. Globalization 3.0, as he calls it, is driven not by major corporations or giant trade organizations like the World Bank, but by individuals: desktop freelancers and innovative startups all over the world (but especially in India and China) who can compete--and win--not just for low-wage manufacturing and information labor but, increasingly, for the highest-end research and design work as well. (He doesn't forget the "mutant supply chains" like Al-Qaeda that let the small act big in more destructive ways.) Friedman tells his eye-opening story with the catchy slogans and globe-hopping anecdotes that readers of his earlier books and his *New York Times* columns will know well, and also with a stern sort of optimism. He wants to tell you how exciting this new world is, but he also wants you to know you're going to be trampled if you don't keep up with it. His book is an excellent place to begin. --*Tom Nissley*

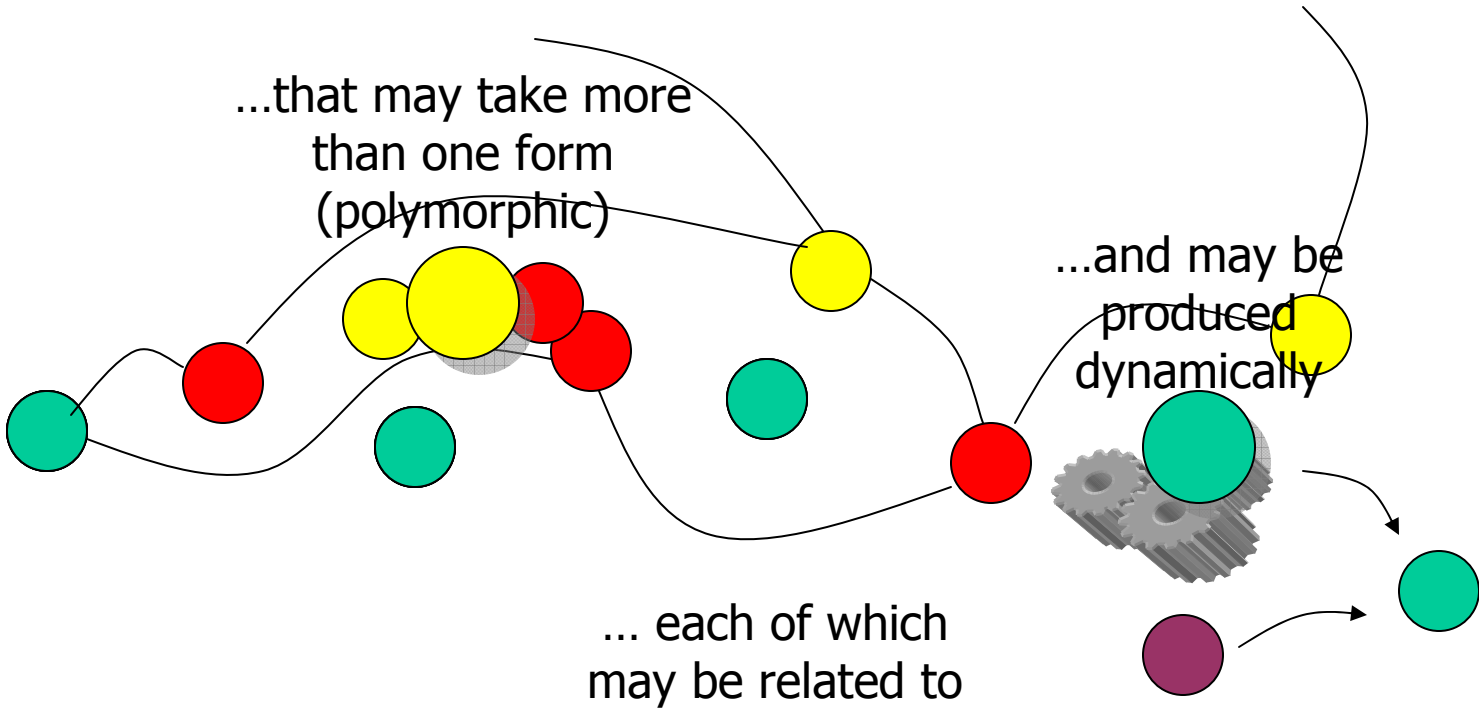
Where Were You When the World Went Flat?



Thomas L. Friedman's reporter's curiosity and his ability to recognize the patterns behind the most complex global developments have made him one of the most entertaining and authoritative sources for information about the wider world we live in, both as the foreign affairs columnist for the *New York Times* and as the author of landmark books like *From Beirut to Jerusalem* and *The Lexus and the Olive Tree*. They also make him an endlessly fascinating conversation partner, and we'd happily have peppered him with questions about *The World Is Flat* for hours. Read our interview to learn why there's almost no one from Washington, D.C., listed in the index of a book about the global economy, and what his one-plank platform for president would be. (Hint: his bumper stickers would say, "Can You Hear Me Now?")

The Essential Tom Friedman

Dealing with the "Item Problem"



NSDL

Login

Home | Search | Collections | Of Interest | AskNSDL | About | Contact | Community

NATIONAL SCIENCE DIGITAL LIBRARY

Search

NSDL PORTALS

Middle School Portal
Middle School Math, Science, and Technology presented by Eisenhower National Clearinghouse

What are NSDL Portals?

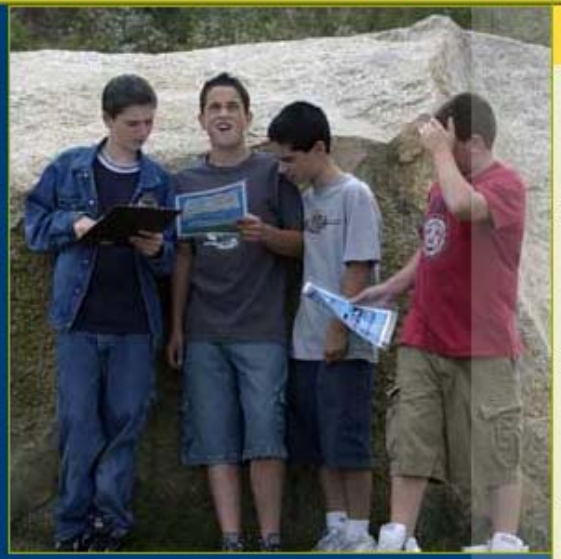
FUTURE PORTALS

Teachers' Domain Pathways to Science
Rich media for K-12 teaching presented by Teachers' Domain at WGBH

The Math Gateway
Undergraduate Mathematics presented by the Mathematical Association of America

The Computational Science Education Reference Desk
Computational Science presented by the Shodor Education Foundation

The Applied Mathematics and Science Education Repository
Community Colleges presented by Internet Scout at the University of Wisconsin



The Nation's online library of resources for science, technology, engineering, and mathematics education and research.

EXPLORE

AskNSDL
How? What? Why?... Try the NSDL ask-an-expert service.

New In The Library
Australian Virtual Engineering Library

Resource Of Interest
Game Technology and Children

HEADLINES

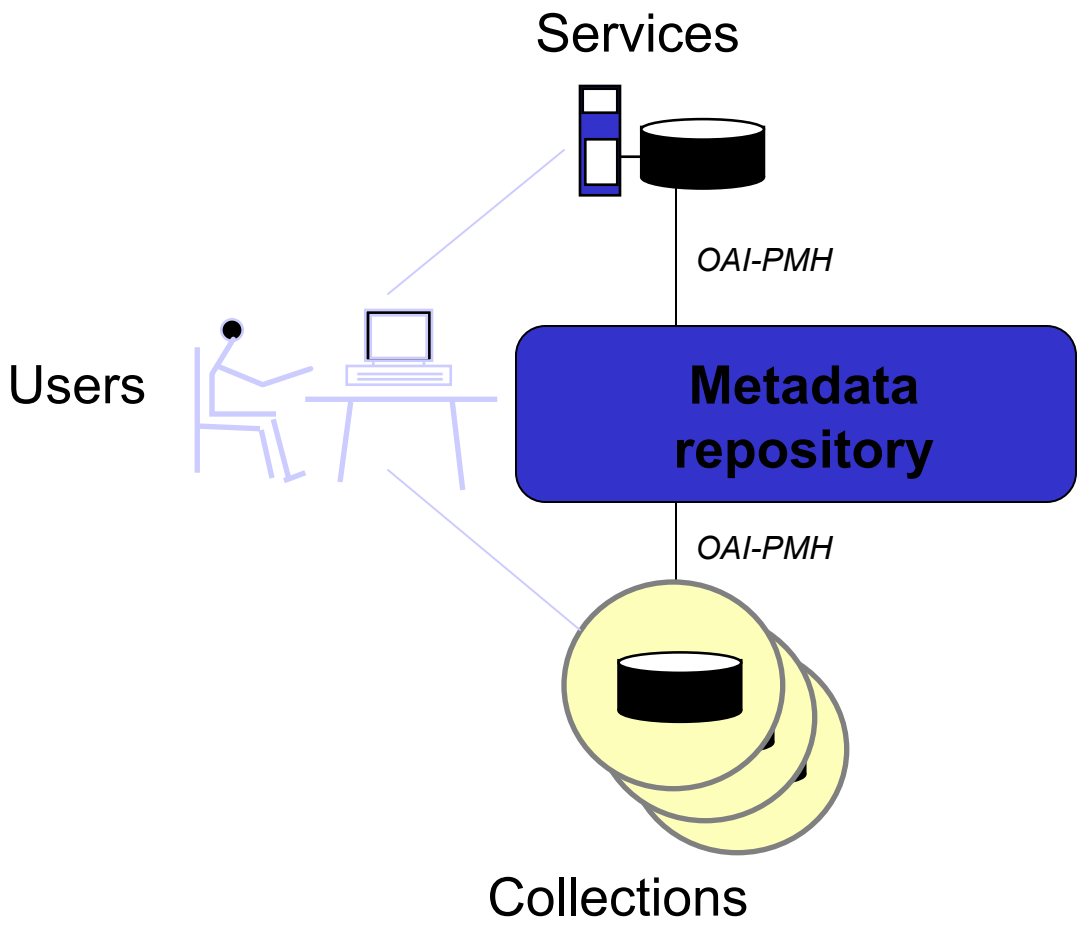
NSDL Partners-at-Sea: Scripps Institution of Oceanography (SIO)

NEW NSDL Search Service: Focusing in on the NSDL Collection

A bit of NSDL background

- Mission: "Improve Science, Math, Engineering education through digital libraries"
- Original NSDL solicitation in 1999
- Over 180 projects funded
- Core integration (Columbia, Cornell, UCAR) charged with providing organizational, technical infrastructure
- CI (Cornell) funding through 2006
- http://www.nsdl.org

Phase I Metadata-Centric Approach



The metadata repository is a resource for service providers.

It holds information about every collection and item known to the NSDL.

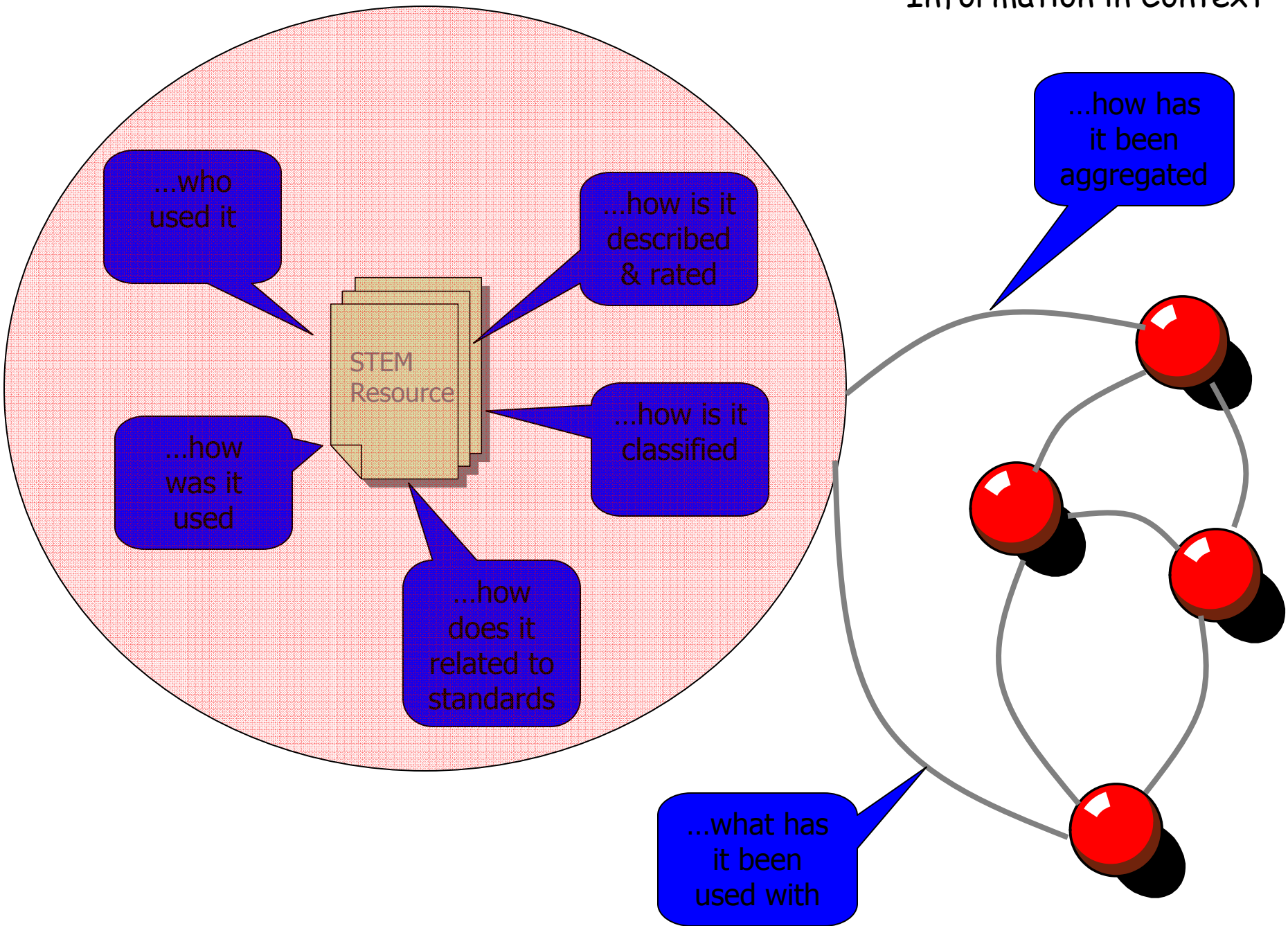
Characteristics of the Metadata Repository

- Oracle database
- Qualified Dublin Core
- Item records with collection association
- OAI-PMH ingest and exposure
- Current collection ~ 1,000,000
- Metadata quality issues

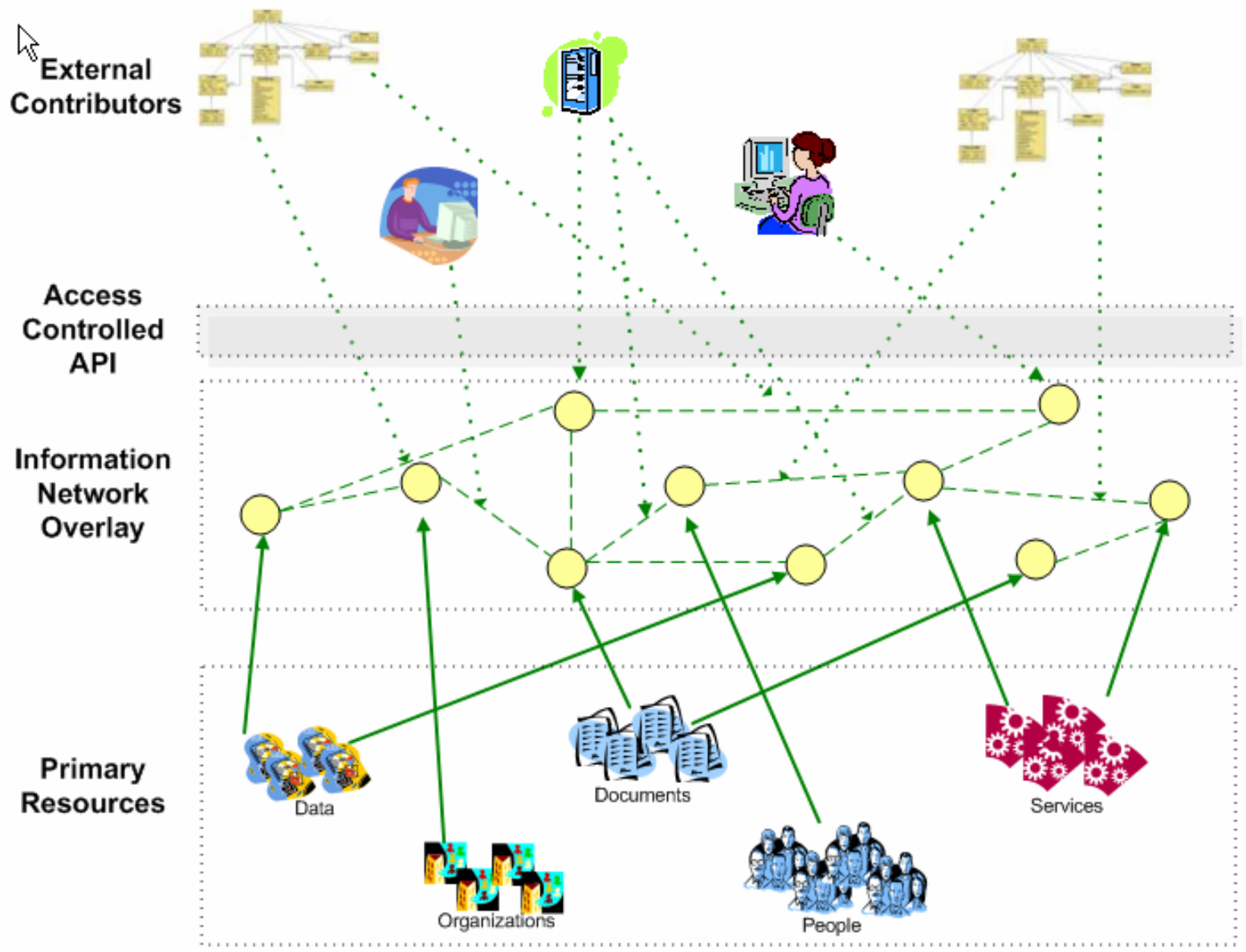
Broader Problems in this approach

- Access alone does not equate to educational value
  - Reeves *Impact of Media and Technology in Schools*
- Static metadata records don't capture changing and multiple contexts of use and applicability
  - Recker and Wiley *Designing Instruction with Learning Objects*
- Patterns of use, informal opinions, descriptions often more useful than taxonomic classification.
  - Collis and Strijker *Technology and Human Issues in Reusing Learning*

Information in Context



Information Network Overlay



Characteristics of the Network Overlay

- Integrate local and distributed entities
- Entities are polymorphic (operational semantics)
- Ontology-based relationships
- Web service integration

Digital Libraries: Beyond Search and Access

- Build on foundation of near universal access
  - "Google *"
- Provide context for:
  - Content aggregation: combining information entities in novel ways
  - Knowledge integration: capturing semantic relationships between information entities
  - Information reuse: allowing secondary, tertiary products
  - Information transformation: combining information entities with computational services
  - collaboration and contribution: blurring the line between authors, publishers, users, experts…

Translate to Technical Requirements

- Rich information objects
  - Integration of local and remote mixed-genre content
- Dynamic information objects
  - Integration with local and distributed services
- Graph-based information model
  - Nodes are information objects
  - Edges are relationships among those objects
- Access and management API
  - exposing full functionality for programmatic access
- Fine granularity access management

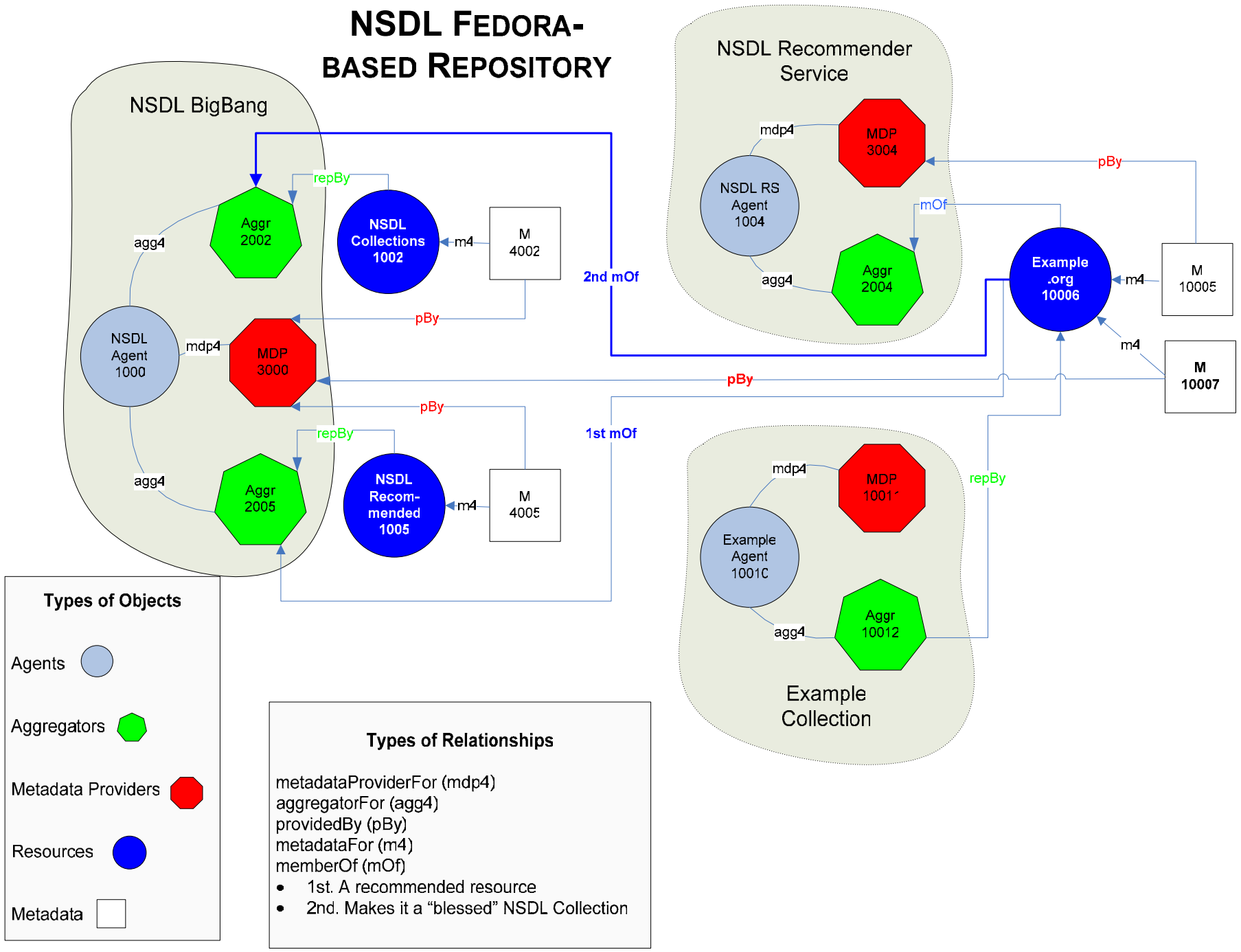
NSDL Data Repository (NDR)

- Fedora-based implementation of information network overlay
- Content model to represent NSDL information entities and relationships
- Extensive use of resource index and new oai service

Fedora NDR

- Objects: agents, metadata items, resources, services (metadata providers), aggregations
- Relationships: metadataFor, providedBy, memberOf, representedBy + ontology-specific
- Disseminations: metadata transformations
- OAI harvesting: both static and generated metadata formats
- Authentication/Authorization: Collections and services manage their own repository content, contribution of annotations, new content

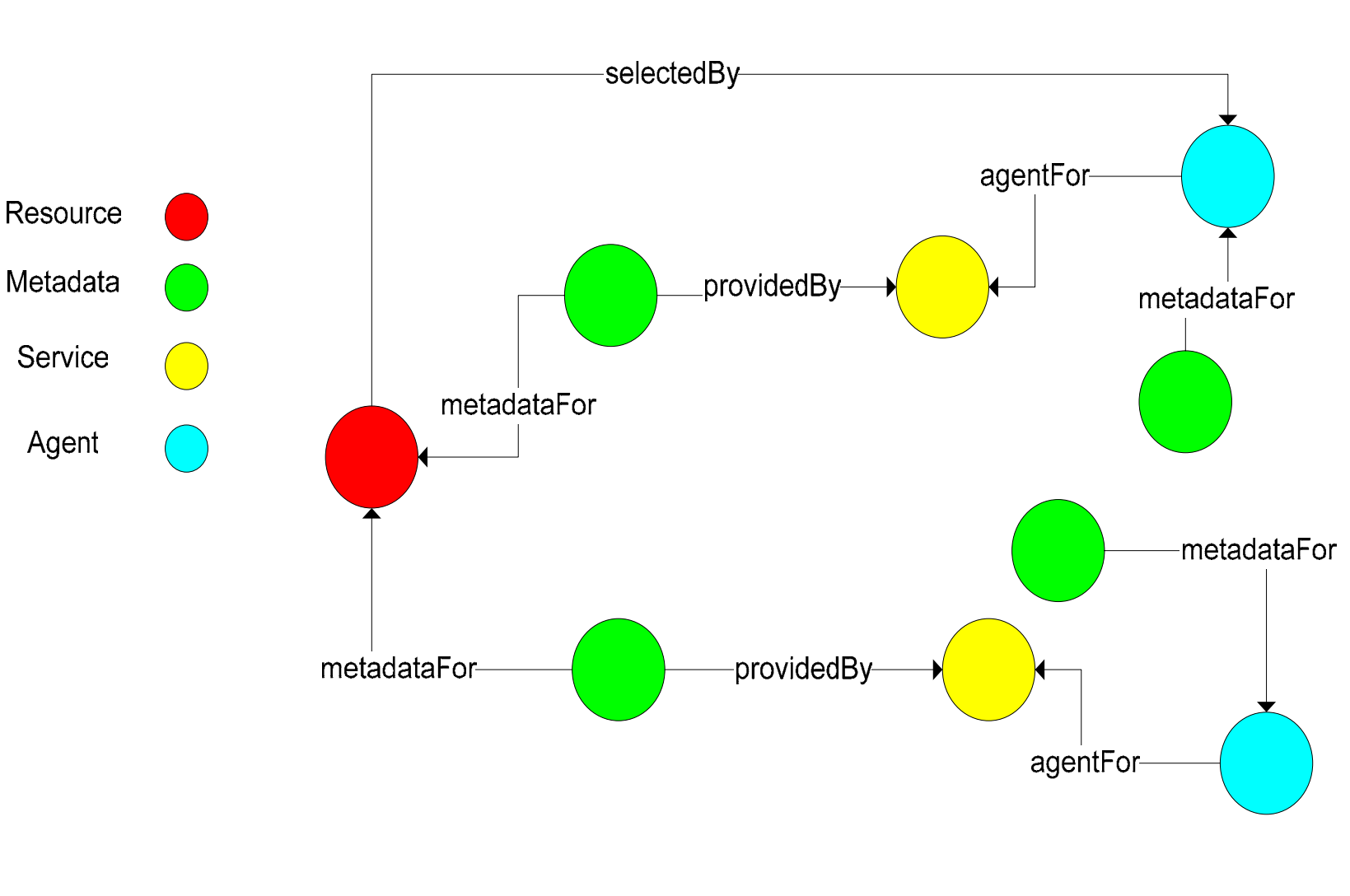
NSDL Fedora-based Repository



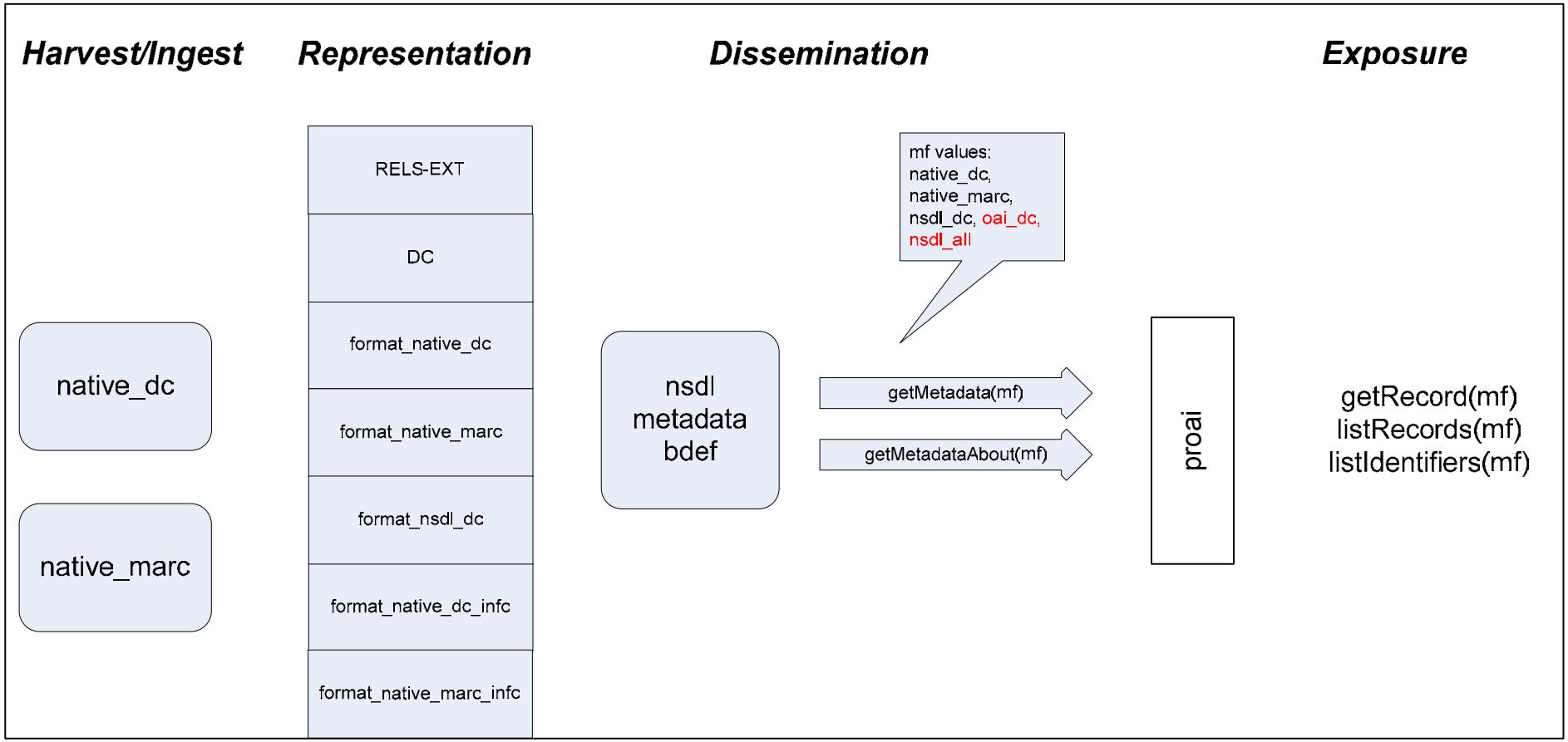
Metadata in the NDR

- Multiple formats
  - static (ingested from provider)
  - generated/crosswalked
- Multi-sourced
  - de-dupped
  - Retain branding of metadata
- OAI-PMH harvesting

Resources, Metadata, Metadata Providers



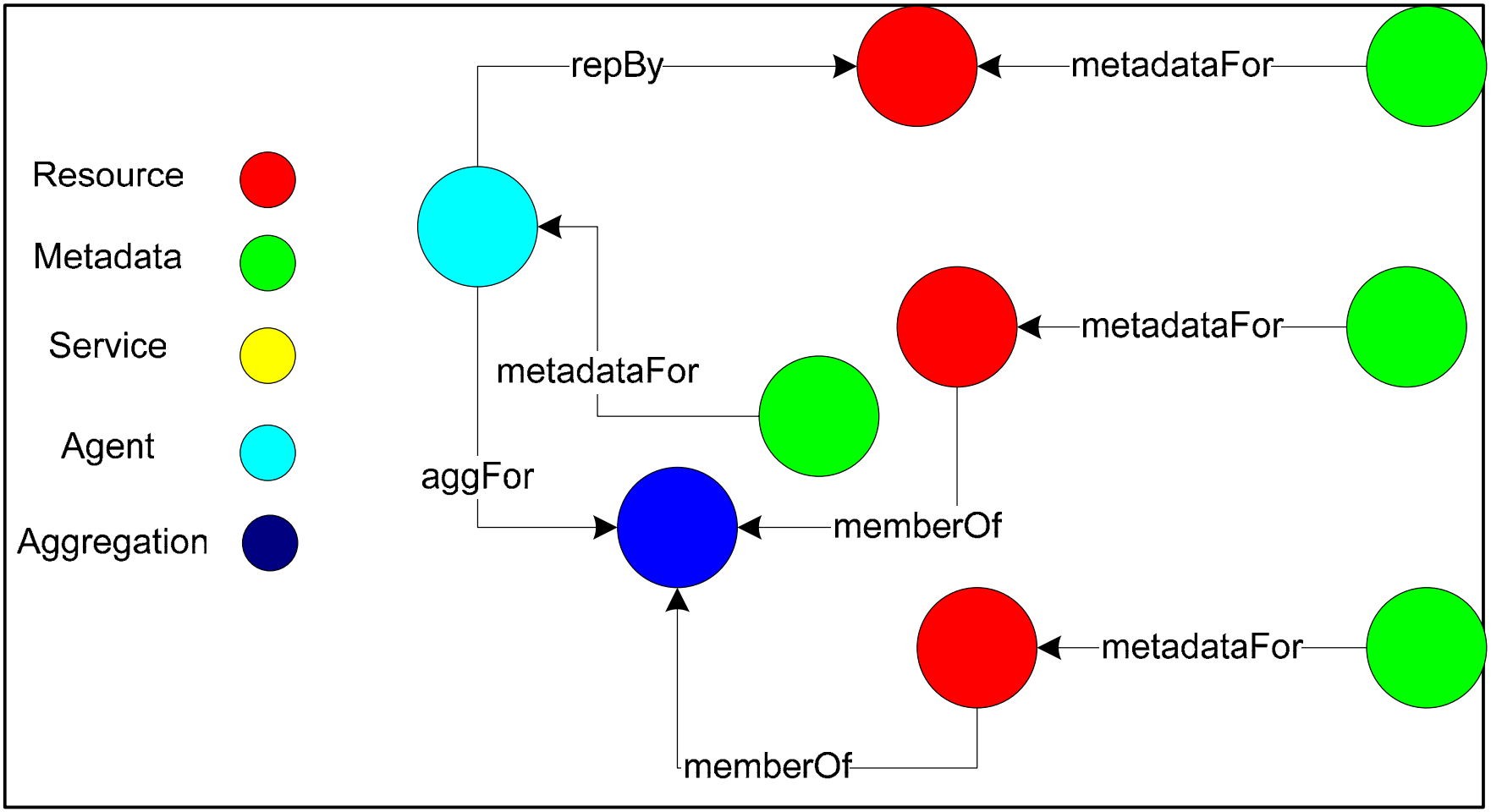
Metadata Content Model



Collections and Aggregations

- Set basis
- Semantic basis
- Agent associated

Aggregation Model



Annotation/Reviews

- Unstructured metadata about a resource
- Exists as resource and annotation
- Separate agent provenance from annotated resource

Annotation Model