Portland Common Data Model (PCDM):
Creating and Sharing Complex Digital Objects

Karen Estlund, Penn State University
Declan Fleming, University of California, San Diego
Mark A. Matienzo, Digital Public Library of America
Jon Stroop, Princeton University Library
Agenda

Purpose and History
Technical Details
Current Work
Future
Q & A
Why a common data model?

- Common language and concepts for:
  - simple to complex object structure
  - rights assertions and access control
  - technical metadata
- Allows for shared objects and coding efforts rather than reinventing for every instance
- Aggregates many minds, projects, and products around difficult concepts
- RDF approach allows for extensibility for the future and local customization
PCDM History

- Hydra implements basic RDF 4/2012
- UCSD complex object data modeling work published 10/2012
- Oregon RDF work - linked open data and vocabularies registry 2013
- Stanford RDF meeting 10/2013
- Hydra Connect 2014

- Use cases on Github - open process 2014 - (more than developers)
- Portland meeting 11/2014
- Portland Code4Lib PCDM Meeting 2/2015
  - concepts like “works,” “resource,” FRBR removed from the model
  - embraced it as an abstract object model not bibliographic
PCDM - More history

- Whiteboard modeling exercises 2015
- OR 2015
  - Open solicitation of input
  - Refined ordering - still in progress but a working approach exists
- Still evolving

- “Portland” - where we met
- “Common Data Model”
  - Because it’s not just Hydra
    - Islandora, Europeana have same data needs
  - History of the naming
    - Avoiding baggage
    - Neutral model
Technical Details - Karen

**Things**
- Collection
- Object
- File

**Relationships**
- Has Member
- Has File
- Aggregates
- Has Related File

[https://github.com/duraspace/pcdm/](https://github.com/duraspace/pcdm/)
PCDM Domain Model

pcdm:Collection

pcdm:hasMember (m:m)
ore:aggregates (m:m)

pcdm:Object

pcdm:hasMember (m:m)

pcdm:hasFile (0:m)

pcdm:File

Access
Bitstream
Descriptive
Technical
Acerca del chocolate.

Quiero que este escrito el título de fragm. El poco del viejo cuad. por el pericia, que aplicó, por el inmenso estado en descripción de cuál se encomienda enaudiador hubo en todos sus frutos dedicados.

Dio su cuidado, después de super..

Firma: Federico Grimza de Onga.
What Can go Where?

...and what can we say about different types of things

A: Access
B: Bitstream
D: Descriptive
T: Technical
Access Control - pcdm:Collection, pcdm:Object, and pcdm:File

Group A can VIEW
Acerca del chocolate: Manuscript on chocolate

Catholic Church--Mexico--History--18th century
Cocoa--Therapeutic use--Early works to 1800

ca. 1730

Mandeville Special Collections Library, University of California, San Diego, La Jolla, 92093-0175

https://libraries.ucsd.edu/ark:/20775/bb3482101x
Acerca del chocolate
Acerca del chocolate

Inscripción

Libro Plate

More than one Object hasFile hasMember
Acerca del chocolate

More with Objects!
Acerca del chocolate
Inscription

<xml>
  <identification>
    <identity format="Plain text"
      mimETYPE="text/plain" toolname="FITS"
      toolversion="0.8.x">
      <tool toolname="Jhove"
        toolversion="1.5"/>
      <tool toolname="file utility"
        toolversion="5.04"/>
    </identity>
  </identification>
</xml>
Collections

pcdm:Collection

hasMember

pcdm:Object

18th Century Catholic Church Medicinal Texts

Acerca del chocolate
Archival Disk Image example

Reference Diagram from Use Case Exercise

https://wiki.duraspace.org/display/FF/PCDM+Mappings++Reference+Diagrams+for+Comment
Work with Fedora in a way that feels like Ruby on Rails
Build models based on PCDM objects and relationships
Hydra::Works

This also allows for shared code around:
- Characterization (JHOVE, DROID, etc.)
- Make Derivatives
- Virus Checking
- Text Extraction

PCDM in Hydra

- Work Objects
- FileSet Objects
- Group FileSets to create Works
Acerca del chocolate

“Page 1”
PCDM in Hydra

- Custom Models
- Routing for CRUD
- File Auditing / Versioning
- Single-Use links
- Upload File Sets
- Access Controls
- Leases and Embargoes
- Proxy Deposit
- User Dashboard & Profiles
- Featured Works & Researchers
- Contact Form
- Activity Streams
- Upload via Cloud Providers
PCDM in Islandora

https://github.com/duraspace/pcdm/wiki/Diagrams#islandora
The Future of PCDM

- Within Hydra, Islandora, and Fedora communities
- Interoperability across other communities
- Governance, engagement, and maintenance
The future of PCDM: Within Hydra/Islandora/Fedora communities

- Hydra in a Box
- Use and development of PCDM recommendations & extensions
  - Rights metadata
  - Technical metadata
    - File format genre ontology
    - File use vocabulary
The future of PCDM: Interoperability across communities

- Alignment with/mappings to additional community models
  - International Image Interoperability Framework
  - Europeana Data Model
  - DPLA Metadata Application Profile
  - RightsStatements.org

- Expansion into W3C standards development processes
  - Linked Data Platform
  - WebAccessControl
The future of PCDM: Engagement, governance, and maintenance

- Community engagement beyond the Hydra, Islandora, and Fedora 4 communities
- Determining best venues for sustained community engagement
- Vocabulary management infrastructure
- Identifying the best long-term home for PCDM
How to learn more and get involved

- Visit the PCDM Wiki: https://github.com/duraspace/pcdm/wiki
- Join the PCDM Google Group: https://groups.google.com/d/forum/pcdm
- Participate in working groups:
  - Hydra Metadata Interest Group
  - Islandora Metadata Interest Group

Questions for attendees
- Where should we go next?
- What do your communities need?
- How we can get you more involved?
- What won’t work for you?
Thank you!

Karen Estlund
Penn State University

Declan Fleming
University of California, San Diego

Mark A. Matienzo
Digital Public Library of America

Jon Stroop
Princeton University Library