

# History



By Tony Webster from Portland, Oregon (Portland Oregon White Stag Sign) [CC BY 2.0 (<http://creativecommons.org/licenses/by/2.0>)], via Wikimedia Commons

# PCDM History - What was Hydra Up To?

- Penn State implements basic RDF April 2012
- UCSD complex object data modeling work published October 2012
- Oregon RDF work - linked open data and vocabularies registry 2013
- Stanford RDF meeting October 2013
- **Hydra Connect 2014**
- **Hydra Works use cases on Github - open process 2014**



- **Portland meeting, “Making Progress” November 2014**

Some slides 3-29 adapted from: Estlund, K., Matienzo, M. A., Fleming, D., & Stroop, J. (2015, December 17). Portland Common Data Model (PCDM): Creating and Sharing Complex Digital Objects. Retrieved from [osf.io/x573h](https://osf.io/x573h)

# PCDM History: Islandora

Another Community at Work



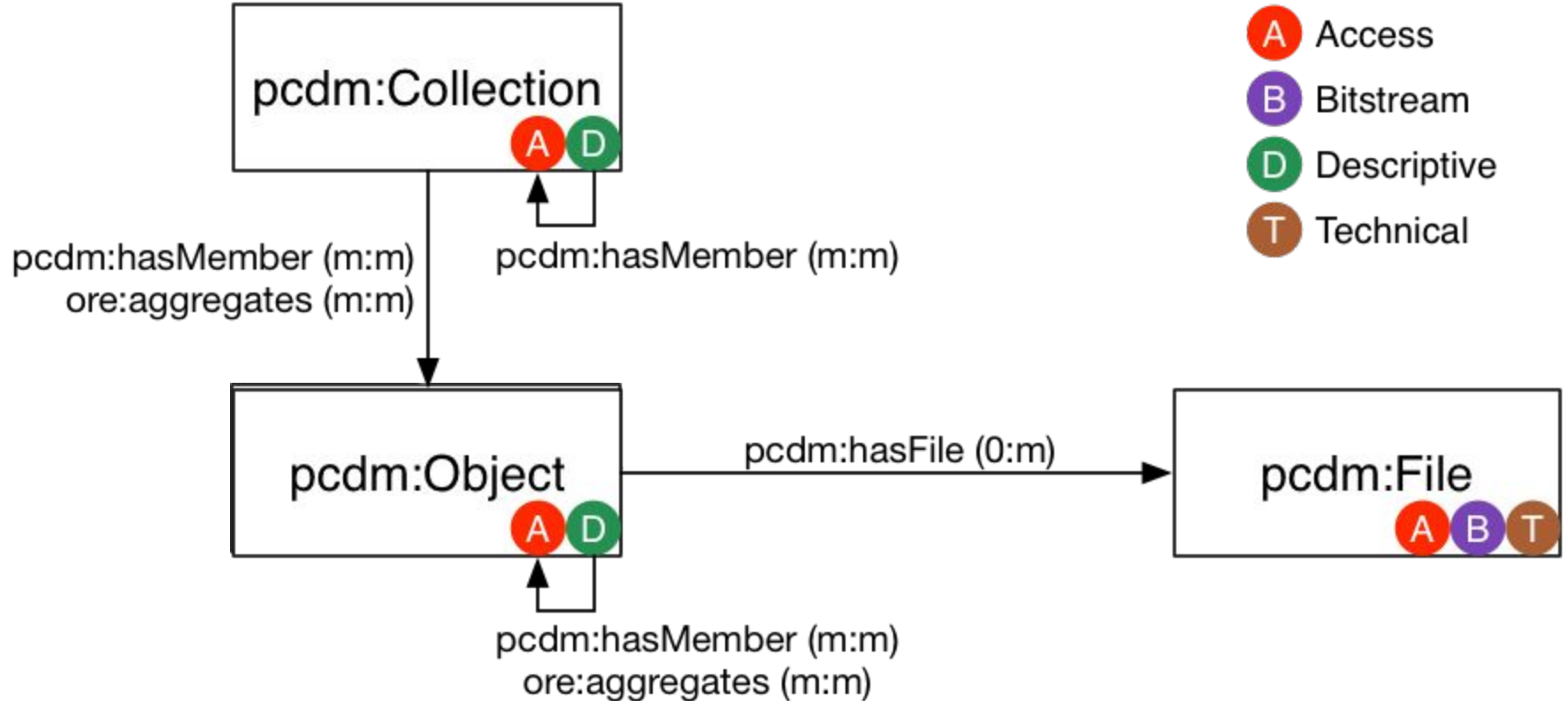
# PCDM - All Together Now!

- **Portland Code4Lib PCDM Meeting February 2015**
  - concepts like “works,” “resource,” FRBR removed from the model
  - embraced it as an abstract object model not bibliographic
- Whiteboard modeling exercises 2015
- Technical and Rights Metadata Work  
May - October 2015
- Refined Ordering - OR 2015
- Works Extension March 2016
- **Initial Implementations Spring / Summer 2016**

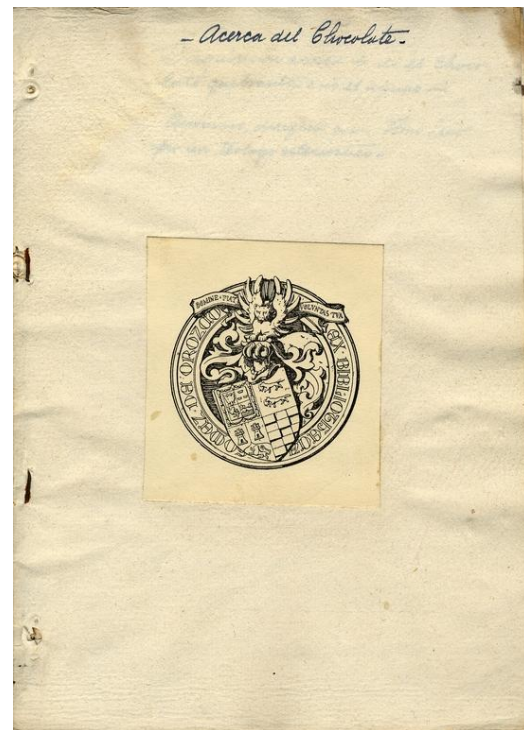
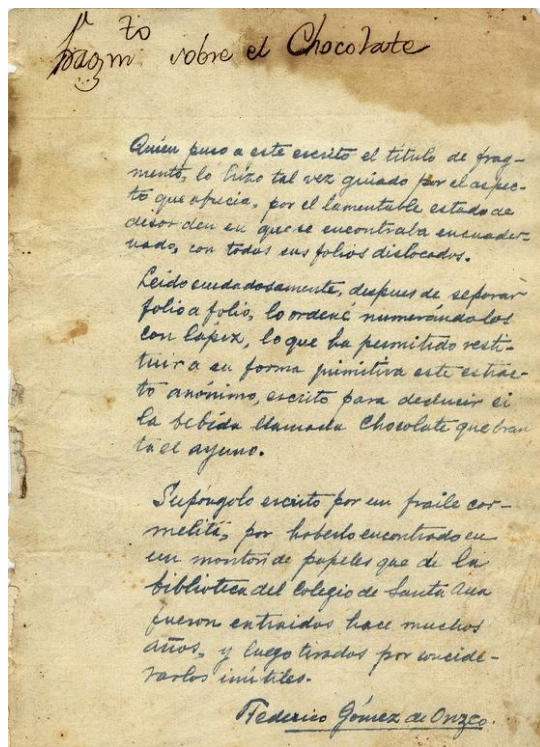
# Model



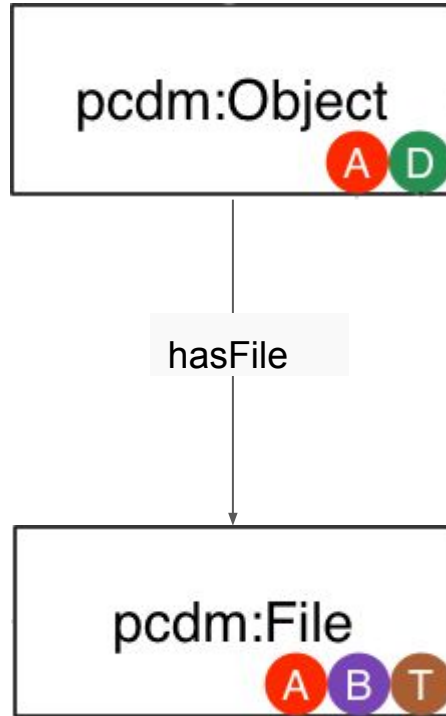
# PCDM Domain Model



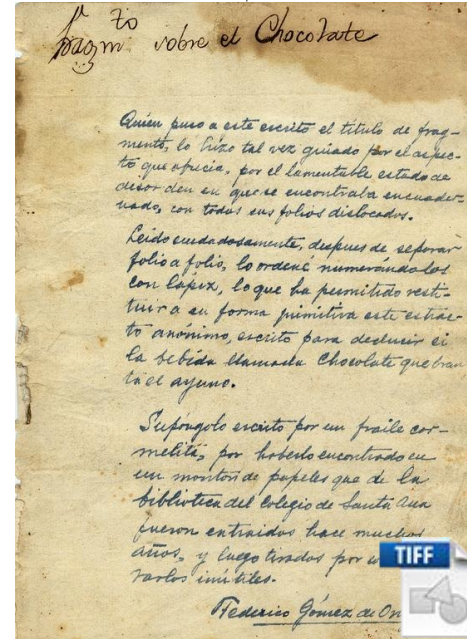




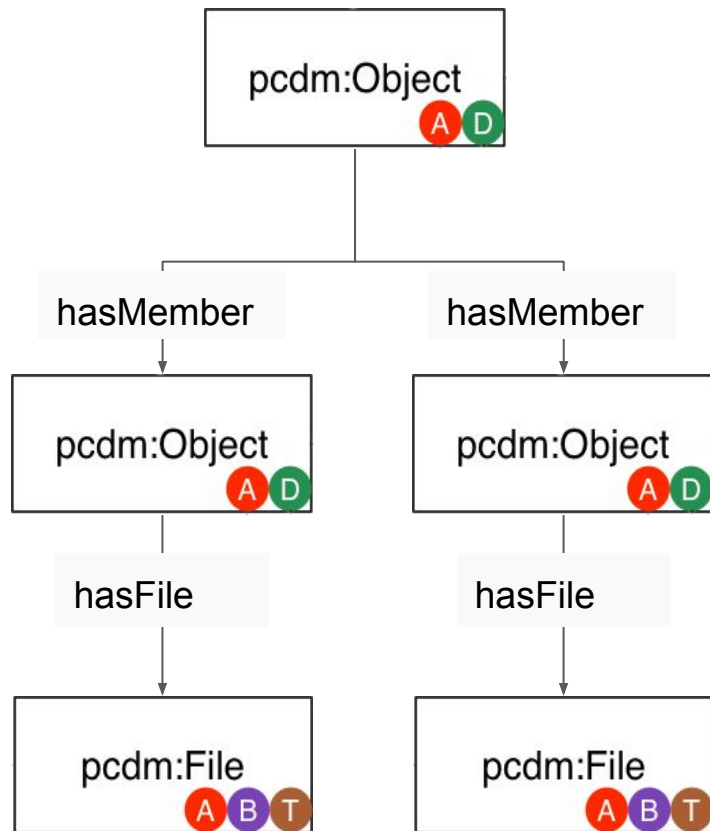
Example: Acerca del chocolate. Mandeville Special  
Collections Library, UCSD



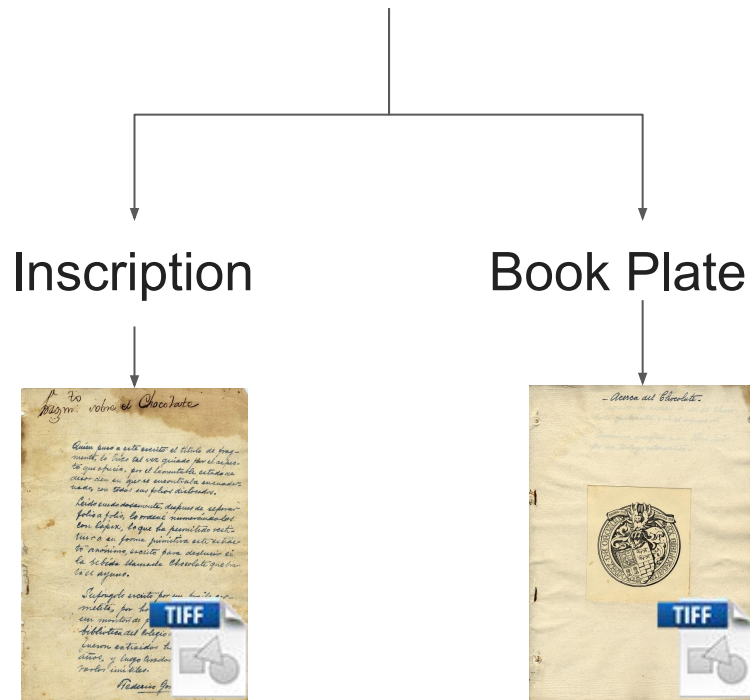
## Acerca del chocolate



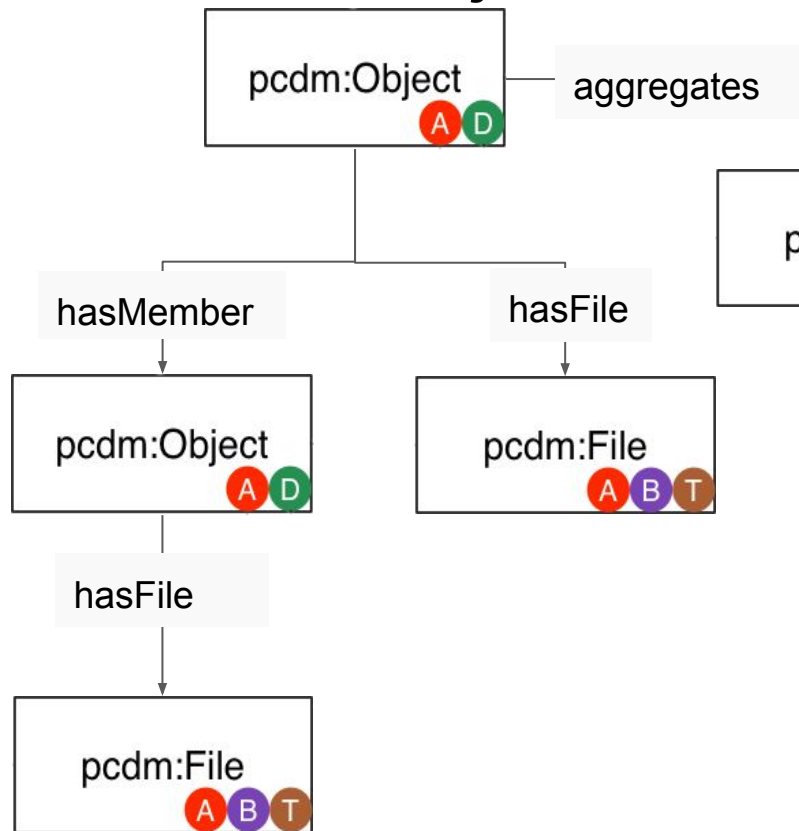
# More than one Object



## Acerca del chocolate

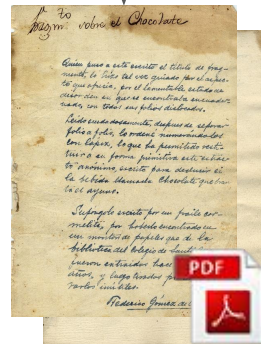
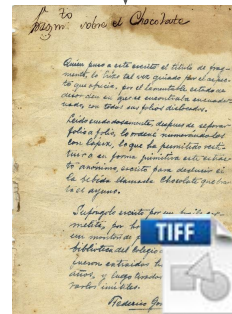


# More with Objects!

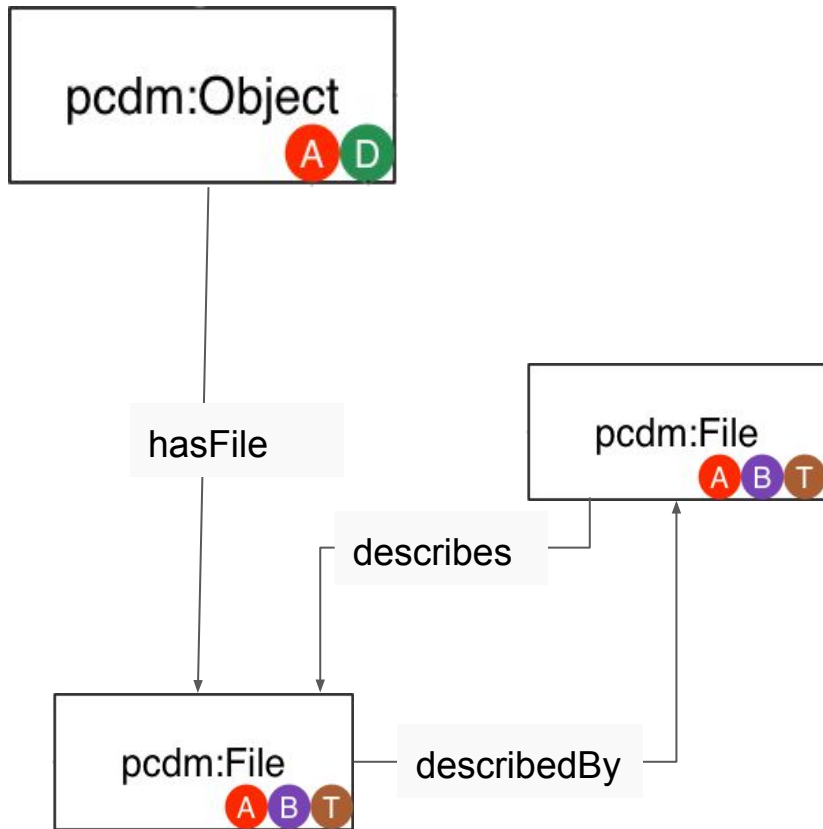


## Acerca del chocolate

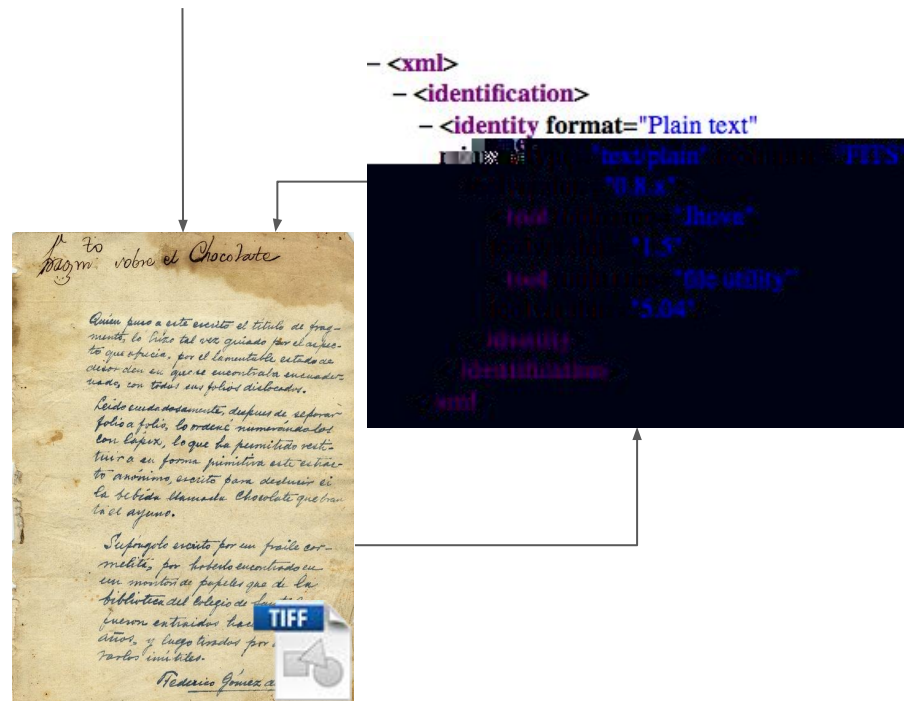
### Inscription



## Related Files



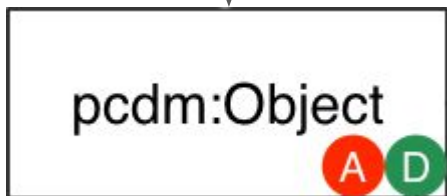
# Inscription



# Collections



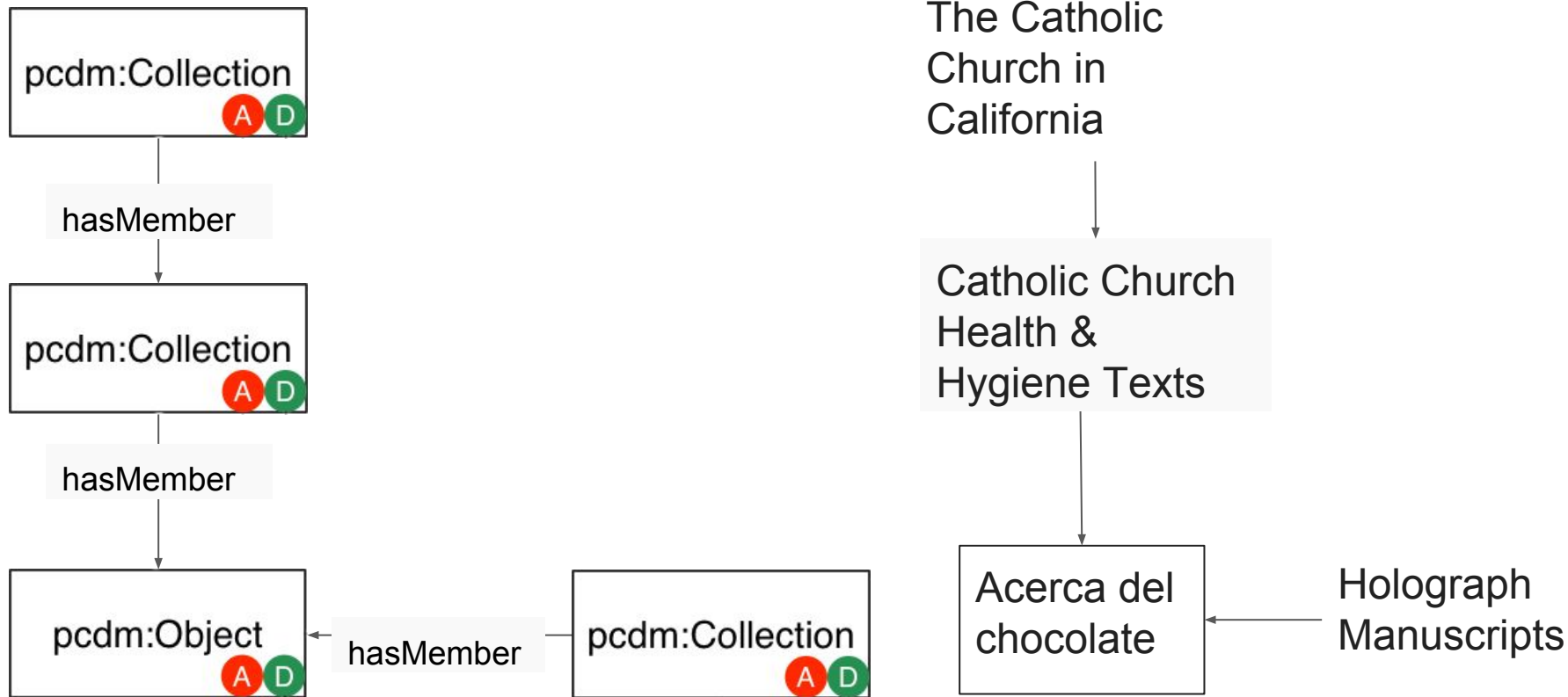
hasMember



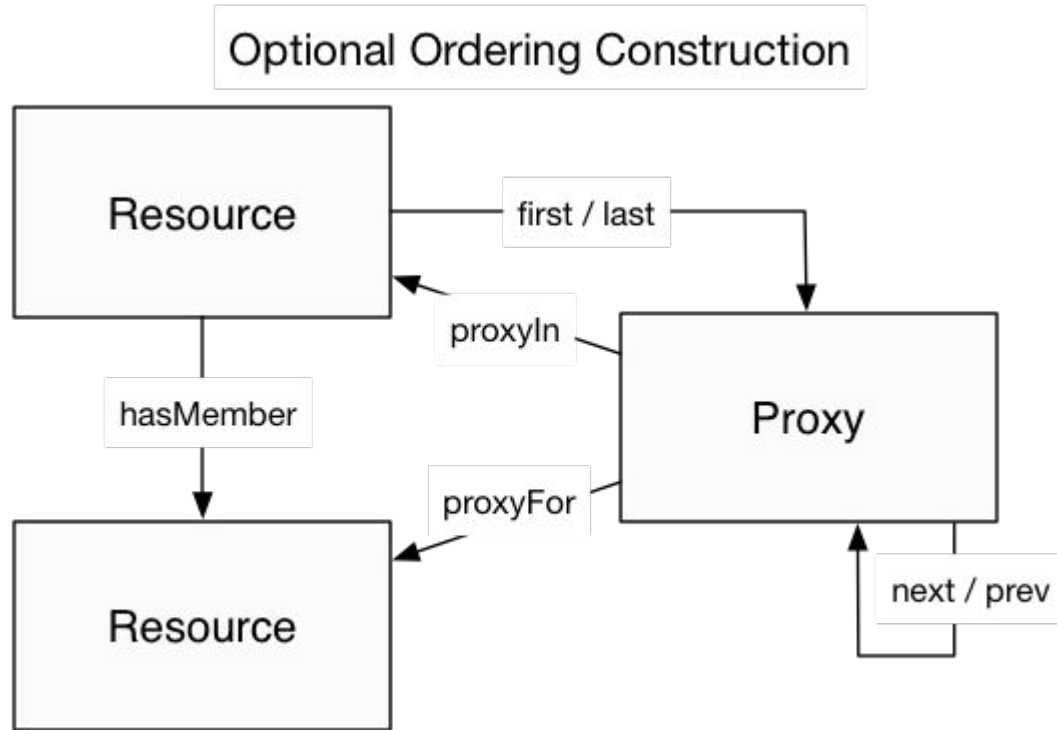
18th Century Catholic  
Church Medicinal Texts



# More Collections



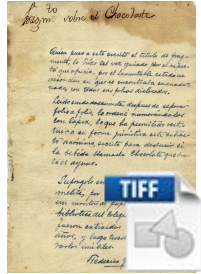
# Optional Ordering Extension



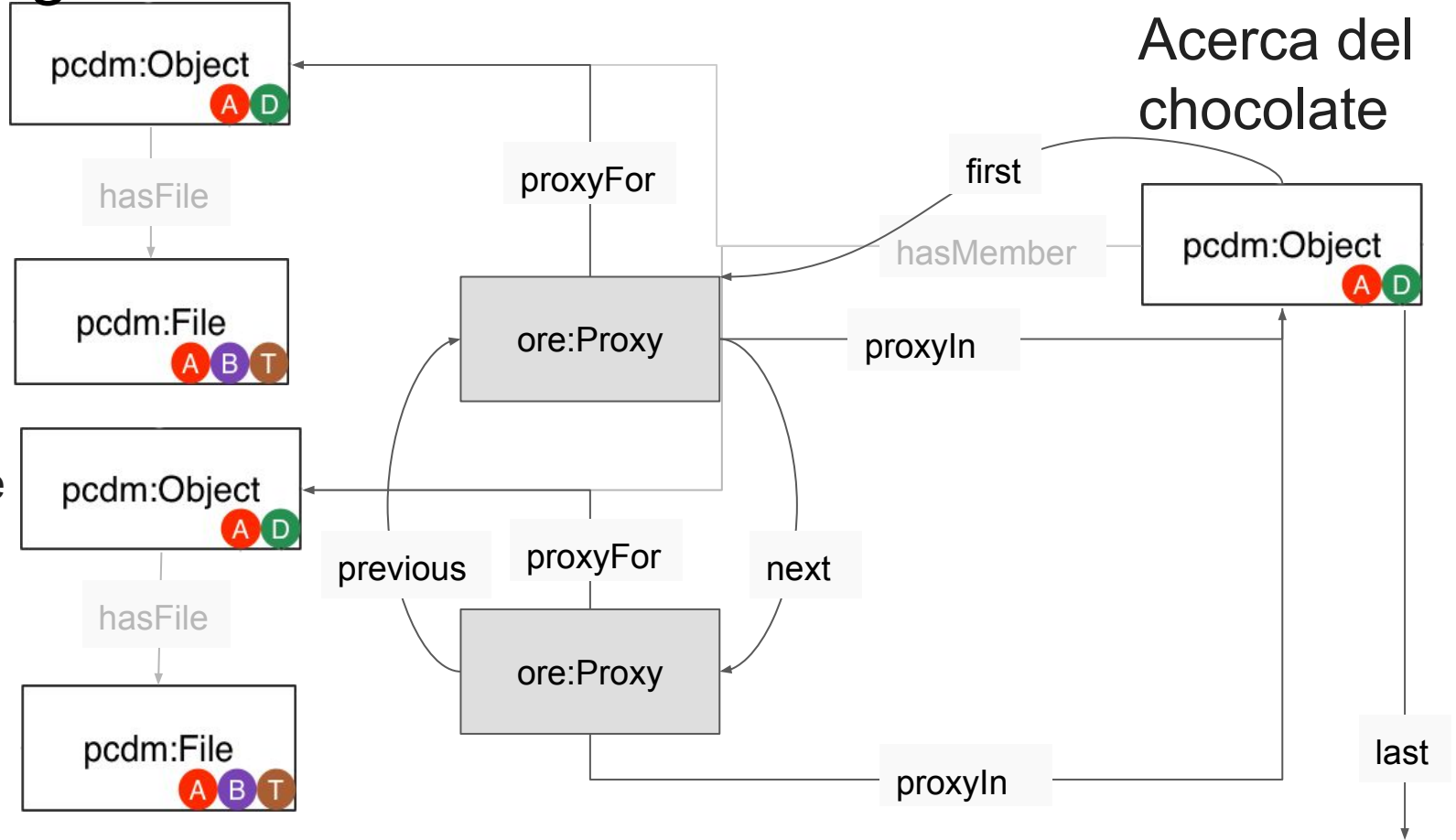
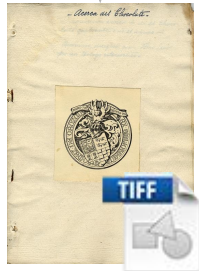


# Ordering

Inscription







Book Plate



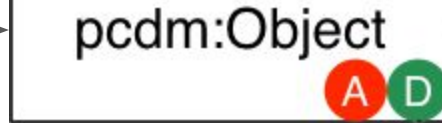
# What Can go Where?

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

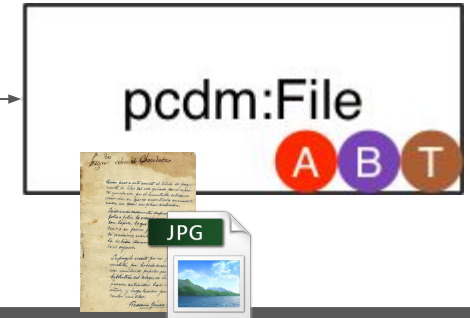
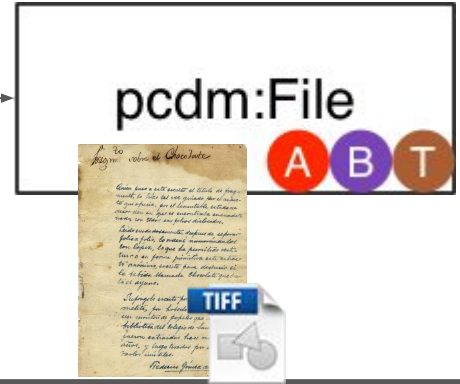
-  Access
-  Bitstream
-  Descriptive
-  Technical

# Access Control - pcdm:Collection, pcdm:Object, and pcdm:File

Group A can VIEW



Group A can VIEW



# Web ACLs

Property	Range	Usage	Obligation
Agent (acl:agent)	foaf:Agent	Individual user this ACL applies to.	min 0, max unbounded
Agent Class (acl:agentClass)	rdfs:Class	Class of users this ACL applies to.	min 0, max unbounded
Mode (acl:mode)	rdfs:Class	Actions permitted by this ACL (e.g., acl:Read, acl:Write, hydra:Discover, etc.).	min 1, max unbounded
Resource (acl:accessTo)	gen:InformationResource	Individual resource this ACL applies to.	min 0, max unbounded
Resource Class (acl:accessToClass)	rdfs:Class	Class of resources this ACL applies to.	min 0, max unbounded

# Descriptive Metadata - pcdm:Collection and pcdm:Object

## 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

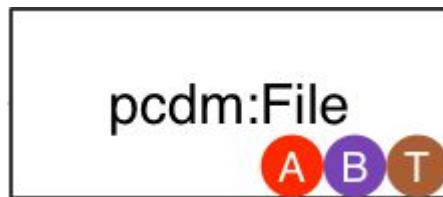
<https://libraries.ucsd.edu/ark:/20775/bb3482101x>

# Technical Metadata and Bitstream: pcdm:File

2013-11-29T13:56:46-04:00

d3b07384d113edec49eaa6238ad5ff00

Intermediate File



527 KB

20775-bb3482101x-1-1-3.jpg

0000E32F  
90990902

image/tiff

# Technical Metadata Recommendations

- ebucore:filename
- ebucore:fileSize
- rdfs:label
- ebucore:dateCreated
- premis:hasMessageDigest
- premishash:md5
- rdf:type
- ebucore:hasMimeType
- ebucore:dateModified
- pronom:puid
- sweetjpl:hasByteOrder

# File Use Vocabulary

- Original File
- Thumbnail Image
- Extracted Text
- Preservation Master File
- Intermediate File
- Service File
- Transcript



<https://wiki.duraspace.org/display/hydra/File+Use+Vocabulary>

<https://github.com/duraspace/pcdm/blob/master/pcdm-ext/use.rdf>



# Rights Metadata Recommendation

## Minimum:

- edm:rights
- dct:rightsHolder

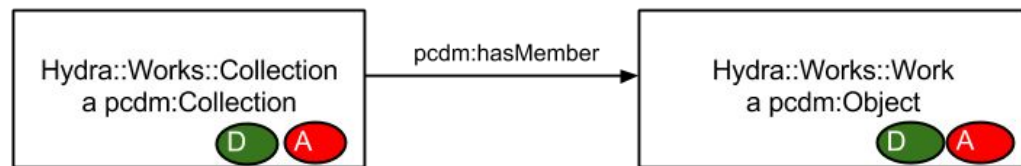


## Recommended:

- dc:rights
- marcrel:cpc (claimant)
- premisHasCopyrightStatus
- pcdmrts:rightsOverride
- pcdmrts:rightsOverrideExpiration

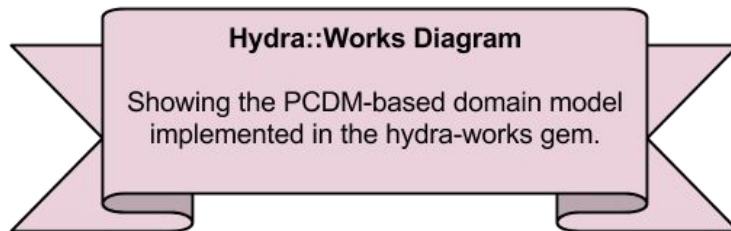
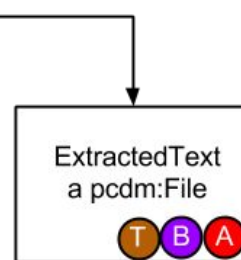
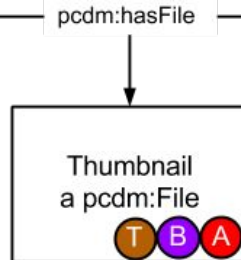
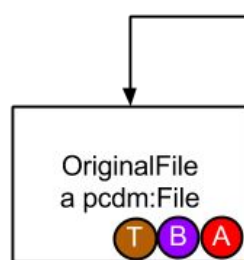
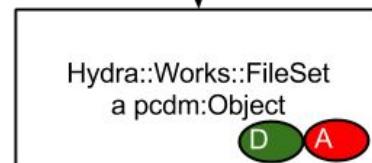
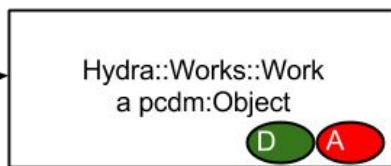
# Works Extension

- **Work:** A work or intellectual entity, such as a book, film, dissertation, etc. Works have member FileSets representing their physical structure (e.g., pages in a book), and related TopRanges representing their logical structure or structures (e.g., sections and chapters in a book).
- **FileSet:** A group of related Files, typically a single master File and its derivatives.
- **TopRange:** A logical ordering of the component parts of a Work, corresponding to a IIIF Range with the "top" viewing hint. Has member Ranges that represent the logical structure, such as chapters within a book, scenes in a film, etc.
- **Range:** A section of a Work, corresponding to a IIIF Range. Has member FileSets representing the physical parts of the Work are part of the section (e.g., which pages are in a chapter).



**Key**

- A Access
- B Bitstream
- D Descriptive
- T Technical
- Structural



# PCDM 2.0

## PCDM 2.0

Esmé Cowles edited this page 23 days ago · 2 revisions

[Edit](#)[New Page](#)

### **pcdm:Collection** < ore:Aggregation

A Collection is a group of resources, such as:

- Resources managed by an institution or unit
- Resources received from a donor
- Resources related to a topic
- Resources selected by a user
- etc.

Collections have descriptive metadata, access metadata, and may link to Objects and/or Collections. By default, member Objects and Collections are an unordered set, but can be ordered using the ORE Proxy class.

Properties:

- `pcdm:hasMember < ore:aggregates` : Links to an Object that is a member of this Collection, or a child Collection. Note on transitivity: `hasMember` is not defined as transitive, but applications may treat it as transitive as local needs dictate.
- `pcdm:hasRelatedObject < ore:aggregates` : Links to an Object that is related to the collection, but not a member of it. Typically used for documentation, thumbnails, etc. Can also be used to link to `AlternateOrder` to encode multiple orders.


### **pcdm:Object** < ore:Aggregation

▼ Pages

[Home](#)[Diagrams](#)[Diagrams with rdflib](#)[PCDM 2.0](#)[PCDM Converters](#)[PCDM Converters Process](#)[PCDM Contributors](#)[Presentations](#)[Related Files](#)[Works Extension](#)

Not a Github user?

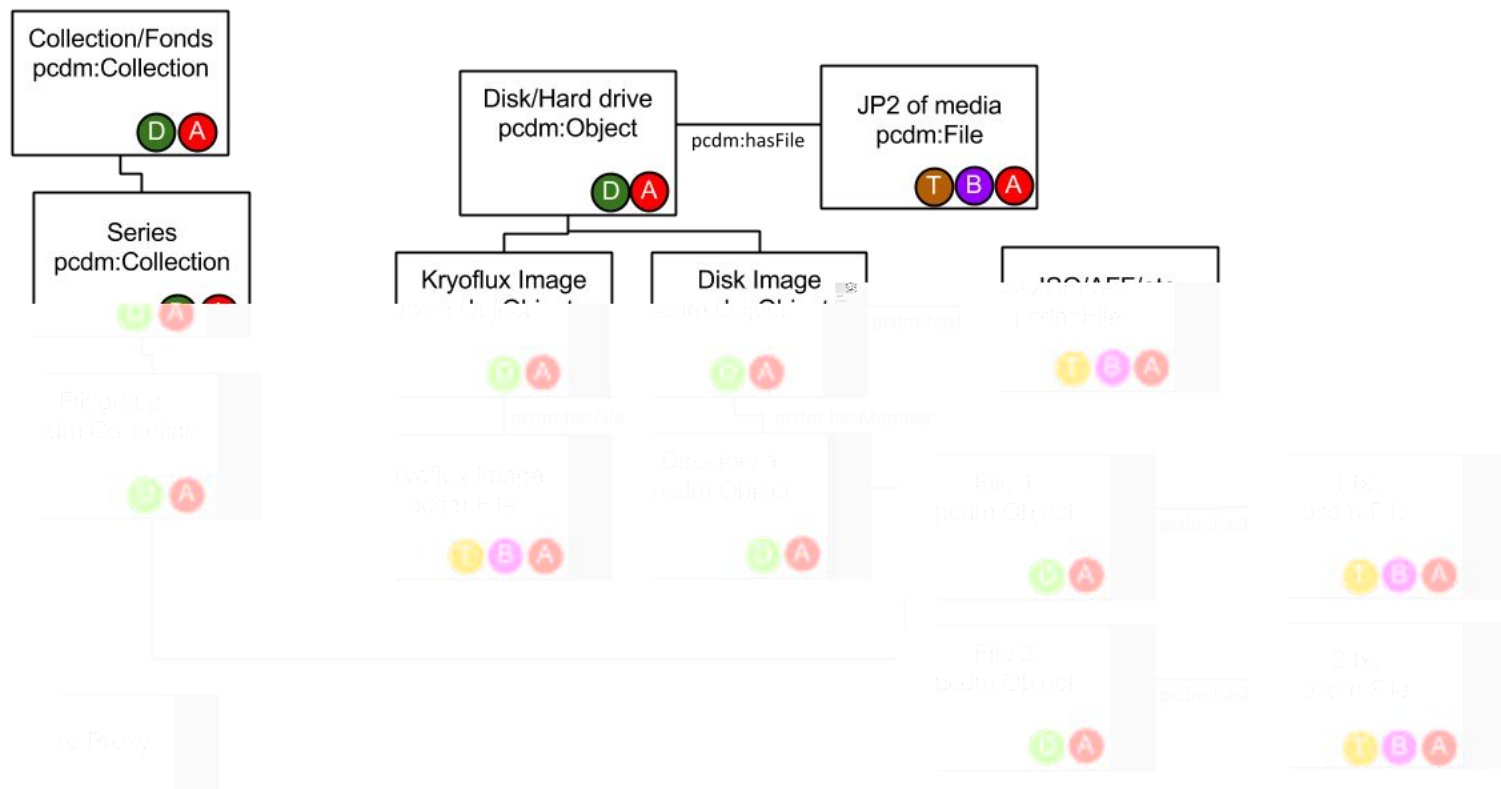
Clone this wiki locally

[How to install this wiki locally](#) 

 Clone in Desktop

# Reference Diagram from Use Case Exercise

Archival Disk Image example



# Community

pcdm.org



## **PCDM Ontologies**

- [PCDM Models Ontology](#)
- [PCDM Use Ontology](#)
- [PCDM Rights Ontology](#)
- [PCDM Works Ontology](#)
- [PCDM File Formats Ontology](#)

## **[PCDM Wiki](#)**



# GitHub



This repository

Search

[Pull requests](#) [Issues](#) [Gist](#)[duraspace](#) / [pcdm](#)[Unwatch](#) 39[★ Unstar](#) 57[🍴 Fork](#) 12[Code](#)[Issues](#) 7[Pull requests](#) 1[Wiki](#)[Pulse](#)[Graphs](#)[Settings](#)Portland Common Data Model <http://pcdm.org/models> — [Edit](#)[50](#) commits[2](#) branches[0](#) releases[8](#) contributorsBranch: [master](#)[New pull request](#)[Create new file](#)[Upload files](#)[Find file](#)[Clone or download](#)[awoods](#) Merge pull request #52 from duraspace/revert-50-adminset

Latest commit 8c3ed39 on Apr 25

[assets](#)

improve HTML display

8 months ago

[examples](#)

Add inverse props now #19 has hit; rename to context as pcdm in hostname

11 months ago

[pcdm-ext](#)

Adding Works extension ontology

3 months ago

[.gitignore](#)

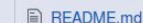
Create stylesheet for html versions of models, rights, and use. Address

9 months ago

[LICENSE](#)

Initial commit

a year ago

[README.md](#)

Add IRC channel

8 months ago

[context.json](#)

Cleanup unused namespaces

3 months ago

[models.rdf](#)

Remove AdministrativeSet. (#50)

a month ago

[pom.xml](#)

Add TransformationSet for format types and works ontologies.

2 months ago

[rdfs2html.xsl](#)

Changing file-format-types to be RDFS instead of OWL, adding support ...

8 months ago

[README.md](#)

# Ontology for the Portland Common Data Model

Namespace: <http://pcdm.org/models#>

Description: This model is intended to underlie a wide array of repository and DAMS applications

[github.com/duraspace/pcdm/wiki](https://github.com/duraspace/pcdm/wiki)

Documentation, Governance, Contributing, Process

[github.com/ruebot/pcdm.org](https://github.com/ruebot/pcdm.org)

The Portland Common Data Model (PCDM) is a flexible, extensible model that is intended to underlie a wide array of repository applications. The primary objective of this model is to establish a common data model for the development of tools [e.g., [Hydra](#)-based engines, such as [Sufia](#) and [Islandora](#)].

## Ontology and Extensions

- [PCDM Models Ontology](#)
- [PCDM Use Ontology](#)
- [PCDM Rights Ontology](#)

## More Information

The [PCDM Wiki](#) provides a fuller description of the core data model, diagrams, project documentation, and [recent presentations](#).

## How to get involved?

All contributions are welcome: use-cases, documentation, code, patches, bug reports, feature requests, etc. You do not need to be a programmer to speak up!

The RDF vocabularies are in the [GitHub repository for PCDM](#), which also includes an [issue queue](#) that you're welcome to contribute to. Anyone can open an issue to discuss changes to the model or documentation.

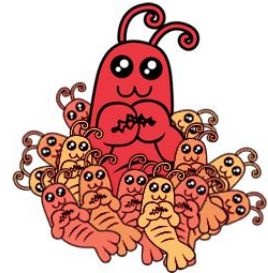
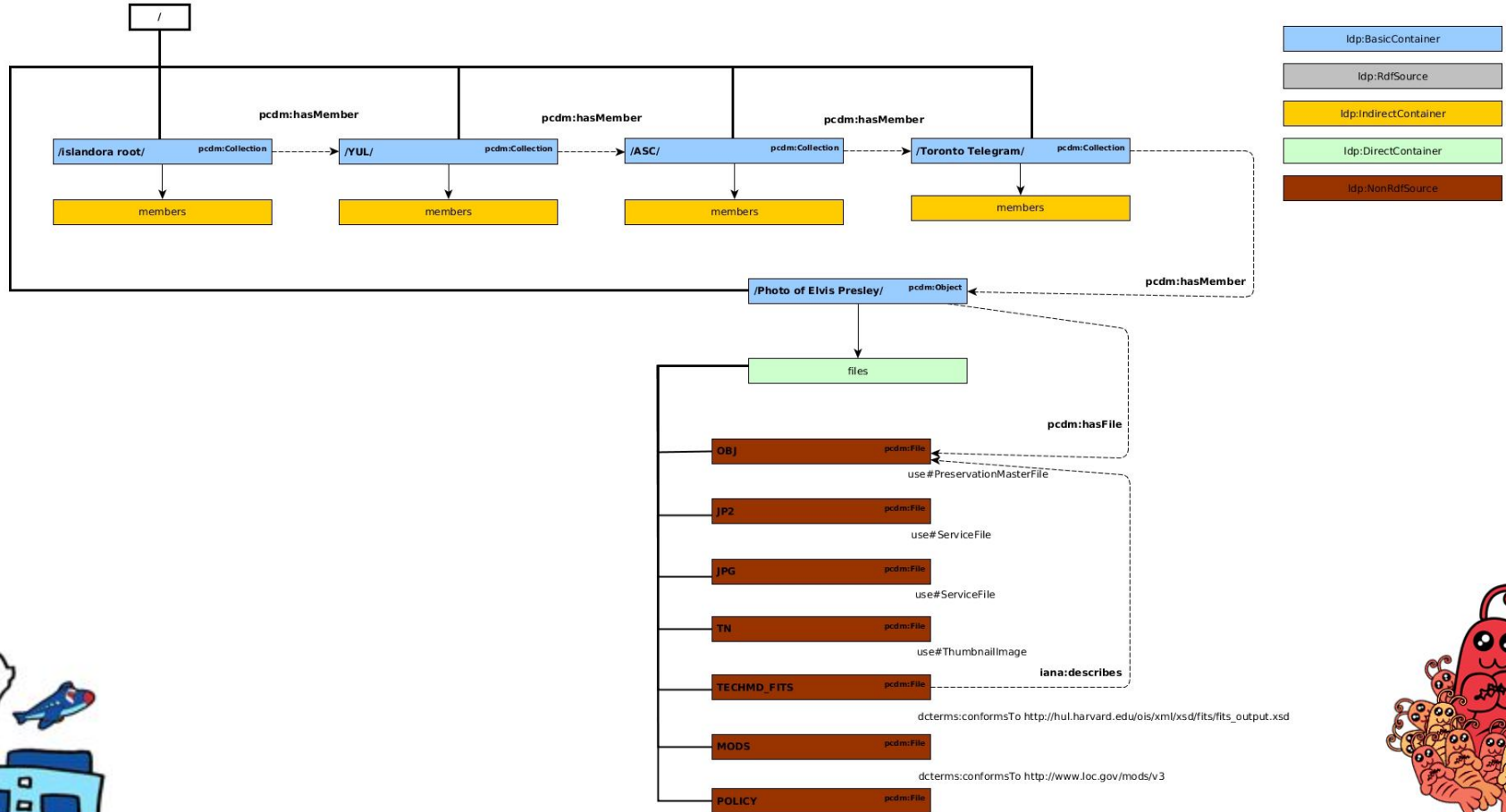
In addition to the GitHub repository, you can also interact with the community via our [Google Group](#).

# Islandora

CLAW!

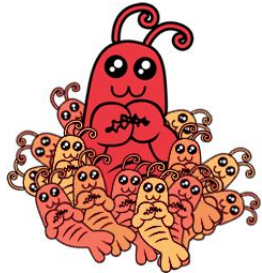
# Islandora, LDP, PCDM, and Fedora 4 (Islandora Large Image Object)

Last modified: December 14, 2015



# The Islandora-CLAW approach

PHP Microservices dealing with semantic layers





# Nested Micro Services Layers

**Chullo:** PHP package that deals with Fedora 4 connection/Sparql. A thin client

**Crayfish:** Silex PHP Microservice (exposes CRUD HTTP Endpoints) provides UUID resolution against fedora 4 resource Paths and deals with RDF resource creation implementing Chullo.

**PDX:** Silex PHP Microservice (exposes CRUD HTTP Endpoints), provides PDCM resource creation. pcdm:Collection, pcdm:Object, pcdm:File. Uses Crayfish internally to build pcdm:Objects and LPD containers

**Drupal:** Model based on related drupal Nodes (entities) that map against fedora 4 resources (bi directional-sync) via Micro Services in JSON

Internally CLAW microservices handle resources as JSON-LD!

# Abstracting Ontologies

## **Future and WIP discussion:**

Allow an extra, flexible layer (microservice) that allows CLAW to implement PCDM and other Ontologies, not as pre-defined in-code models, but as a flexible pre-parsing service based on simple(simplistic) reasoning over OWL 2.

The Idea: give our Users the power to decide which of the n-possible valid models for a given ontology they want to use (or create), and based on this generate narrowed down and better restricted ontologies that assert more specific use cases and knowledge domains.



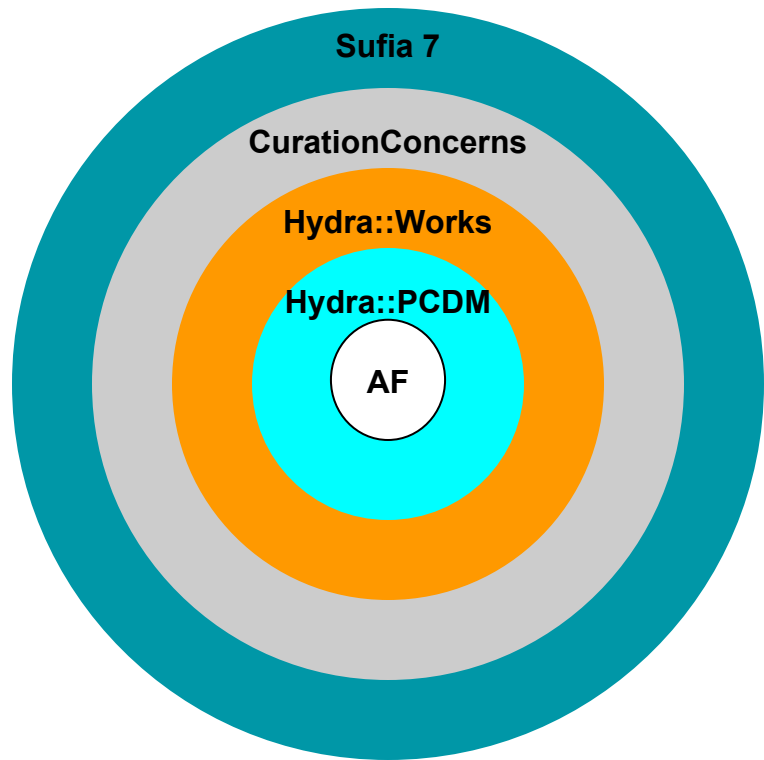
Esmé Cowles



Princeton University Library

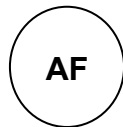
# Building PCDM into the Hydra stack

- Building a new Hydra stack for PCDM
- Community realignment around collaborative development
- Doing things the Hydra Way™
- Current status:
  - Curation Concerns 1.0 beta
  - Sufia 7.0 beta



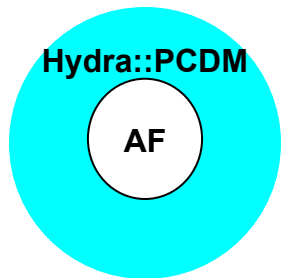
# ActiveFedora

- (Still) the foundation of the Hydra stack
- Makes Fedora resources look like ActiveRecord
- Added support for ORE Proxy ordering



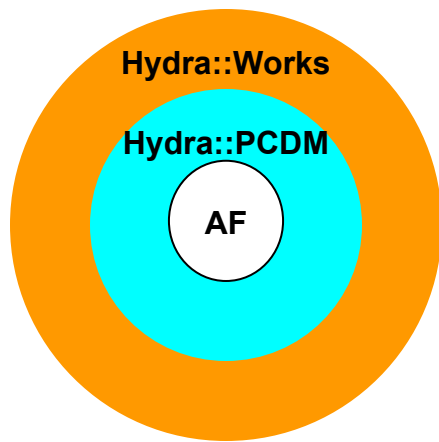
# Hydra::PCDM

- Implementation of core PCDM classes and relationships
- Intentionally simple and unopinionated to allow other Ruby apps to use PCDM without having to buy into the rest of the stack



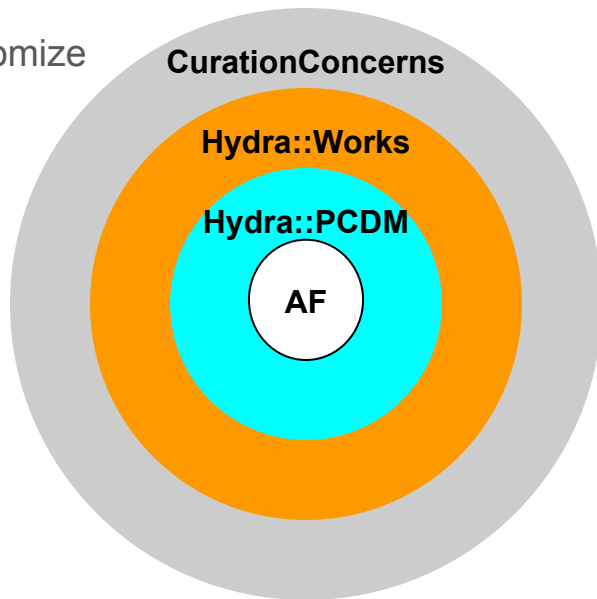
# Hydra::Works

- Repository toolkit
  - Characterization, derivatives, collections, virus scanning
  - Services for uploading files, versioning, etc.
- Extends PCDM with a few more classes
  - Work: pcdm:Object representing whole or part of digital object
  - FileSet: pcdm:Object representing a File and its derivatives
  - see also: PCDM 2.0
- Good place to start for
  - gems and tools
  - radically different apps



# CurationConcerns

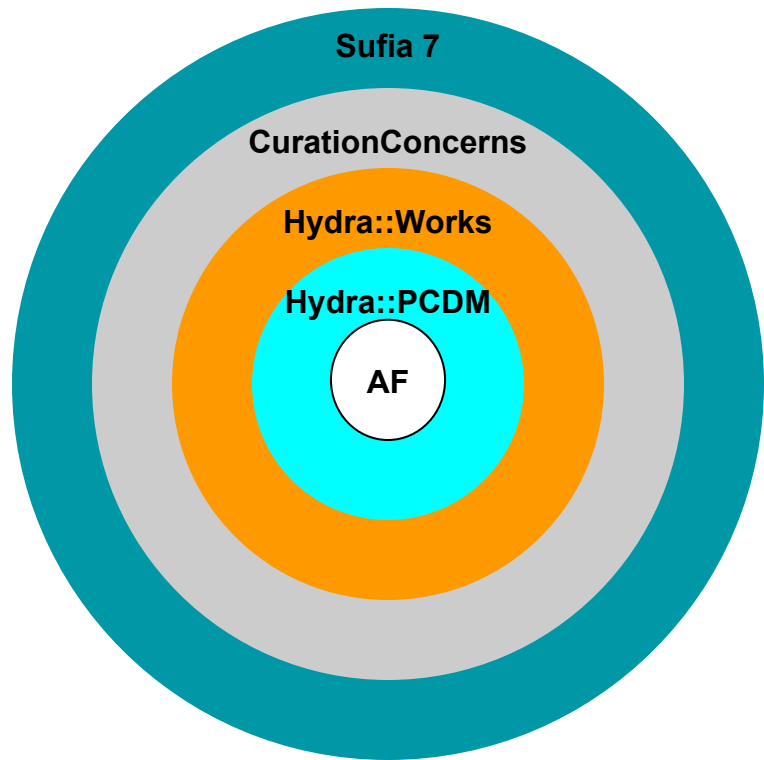
- Rails engine with controllers and views
  - Provides a usable application out of the box
  - Generators for creating work models
  - Functionality in modules to make it easier to override/customize
- Default functionality
  - Works with member FileSets
  - Works with member Works
  - UI and workflow for uploading files, FITS characterization, generating derivatives, etc.
  - Drag-and-drop ordering
- Good place to start for digital collections apps





# Sufia

- Adds institutional repository functionality
  - Self-deposit workflow, proxy deposit, social features
  - Dashboards, statistics, reporting
  - Now extending CurationConcerns
- Default functionality
  - GenericWork with member FileSets
- Good place to start for insitutional repositories



# Evolving Data Model

- Do we need a Work class?
  - Hydra::Works has a Work class and uses FileSets for parts
  - Moving towards using pcdm:Object for works and their parts
- FileSet: logically-grouped sets of Files
  - e.g., a source File and its derivatives — not a part/component/page
- Physical Structure vs. Logical Structure
  - (Top)Range to represent overlays like tables of contents
  - Also aligns with IIIF

# Questions!