

Hypatia

Hydra Platform for Access to Information in Archives

DLF Forum * Baltimore * October 31, 2011

Tom Cramer
Michael Olson
Naomi Dushay

 Stanford University

Bradley Daigle
Julie Meloni
 University of Virginia

Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



Introduction	Tom Cramer
AIMS	Bradley Daigle
Born Digital Materials (& Forensics)	Michael Olson
Hypatia Functional Requirements	Michael Olson
Data Models & Loading	Naomi Dushay
Demonstration	Julie Meloni & Michael Olson
Q&A	<i>Discussion</i>
In Sum & Looking Forward	Tom Cramer

What is Hypatia?

- **Hydra Platform for Access to Information in Archives**
- Repository-powered solution for **digital archival materials** management, preservation and access
- One component in a larger (eco)system for archivists
- Open source software based on Hydra & Fedora
- The potential nucleus of a larger, sustained, collaborative effort

Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



Origins

- Outgrowth of AIMS project
- Leveraging the Hydra project
- Functional requirements and content from the AIMS partners (Virginia, Hull, Stanford, Yale)
- Technical Development by Stanford, Virginia & MediaShelf (contract)

Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



What is Hydra?

Partners

- DuraSpace
 - Northwestern University
 - Notre Dame
 - Rock & Roll Hall of Fame
 - Stanford University
 - University of Hull
 - University of Virginia
- + half dozen more in ramp-up mode

Technology

OSS stack featuring Fedora, solr, Ruby on Rails, Blacklight

“Solution Bundles”

- IR
- ETD's
- Research Data
- Video
- Images
- Archives → Hypatia
- Open Access Articles
- Digitization Workflow
- Digital Monograph Acquisitions
- Exhibits
- Digital Preservation

Hypatia Partners

Stanford University

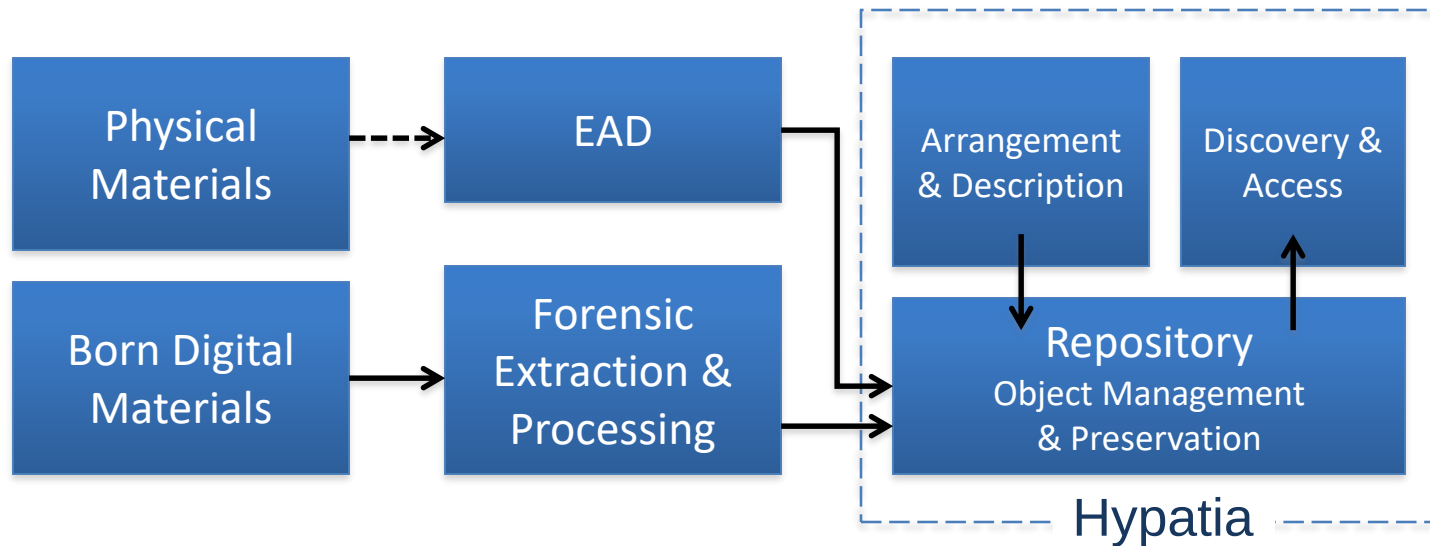
University of Hull

Yale University

University of Virginia



The Workflow



Hypatia Partners

Stanford University

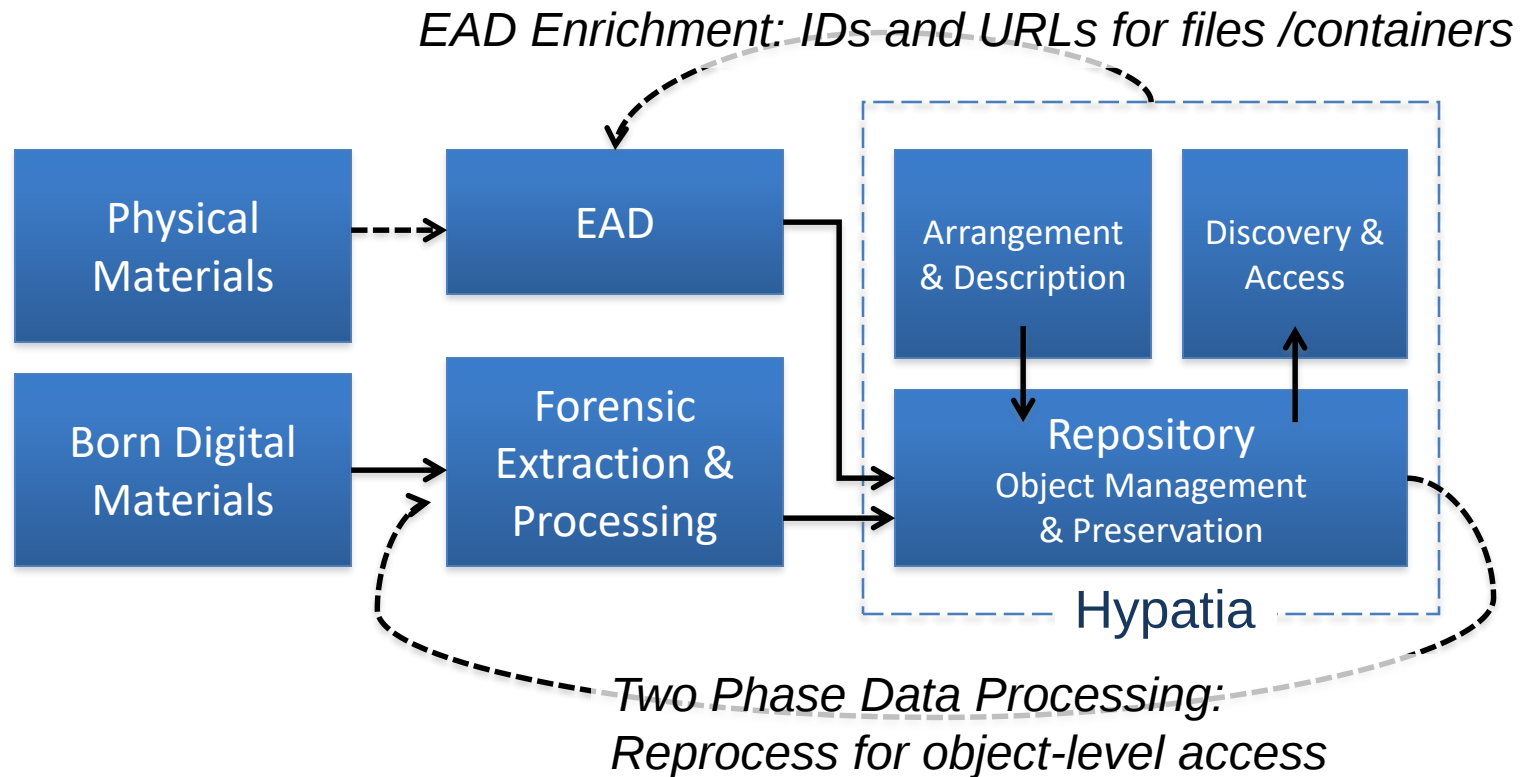
University of Hull

Yale University

University of Virginia



Iterations & Enrichment



Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



AIMS: BORN-DIGITAL COLLECTIONS

An Inter-Institutional Model for Stewardship (AIMS)

October 2009-September 2011

Funded by The Andrew W. Mellon Foundation

Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia

The Purpose

- To bring four top research libraries together to develop a common framework for the collecting, organizing, preserving and delivering of born-digital archival collections

Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



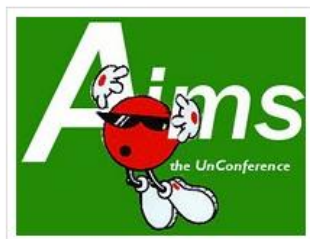
born digital archives

The AIMS project, funded by The Andrew W. Mellon Foundation, represents a co-operative of four partner institutions, to energize collection development in the area of born-digital paper, empower librarians and archivists in the management of born-digital assets. The four partner project led by the [University of Virginia](#) are [Stanford University](#), [University of Hull](#) and [Yale University](#).

[The Collections](#) [Documentation](#) [Links etc](#) [About Us](#)

Wednesday, 18 May 2011

AIMS: the UnConference



Not two full weeks into my new job as Digital Archivist at UVA on the AIMS grant, I rolled up my sleeves to facilitate and host an [unconference](#) with my fellow Digital Archivists. Our unconference would be two full days of discussions, demonstrations,

lightning talks, and networking with digital archivists from around the globe. At first the thought was a little terrifying – I'm not even fully sure I know what this job is yet, how could I actually lead discussions on the salient topics? But my fears were baseless: all the unconference attendees were thoughtful, articulate, and lively participants. I learned much more from them than they probably did from me.

The unconference was held on the 13th and 14th of May at the Omni Hotel in Charlottesville. The [27 participants](#) represented libraries, archives, museums, and digital humanities centers across the US, Canada, and the United Kingdom. Despite the differences in our institutions, backgrounds, and training, we learned that we not only shared similar challenges, but also the same hopes for collaboration and innovation.

The first day started off with a round of lightning talks. Each participant had 5 minutes to present a topic, project, problem or idea that they were interested in talking about. The variety in the talks was remarkable to me, traversing the breadth and depth of all that can be thought of as "born-digital" and the many processes involved in managing it. The lightning talks were also great way to get an introduction to each participant as well as their [perspective](#) or the particular issues they were dealing with in their institution. A brief outline of each of the talks is available on the [AIMS Unconference Wiki](#).

Search this blog

Labels

[AIMS](#) [conference](#) [Dig](#) [archivemata](#) [archives](#) [boy](#) [ingest](#) [tools](#) [survey](#) [BEAM](#) [Digital Archiving](#) [ECA 2010](#) [METS](#) [Pecha Kucha](#) [Rare B](#) [accession](#) [arrangement](#) [cu](#) [digital forensics](#) [digital live](#) [digitisation](#) [donors](#) [files](#) [microservices](#) [preservation](#)

Subscribe To

Posts
All Comments

Our Favourite Blo

[Digital Collections Archives \(Tufts\)](#)
An Affectionate Look
Affiliates
[Digital Forensics @ University Libraries](#)
Workshop on "Using I
and AccessData FTK
and Process Born Dig



University of Virginia



Yale University



Hull University



Stanford University

Funded by the Andrew W. Mellon Foundation

Outcomes

- Energized collection development & delivery of born-digital archives
- Portable tool set for the management of born-digital collections that other institutions could adapt
- A foundation on which to build a digital archivist community
- Published guide of methods used and lessons learned

Key to Success

- AIMS is based on stewardship of materials-- both analogue and digital
- Informs development of tools and solutions that are informed by archival practice
- Identify the various stages in workflows and suggest methodological solutions

Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



Follow us

- Blog: Born Digital Archives
<http://born-digital-archives.blogspot.com>
- Website: AIMS – Born Digital Collections
<http://www2.lib.virginia.edu/aims>

Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



Functional Requirements Gathering

- Created by AIMS Digital Archivists'
- January – March 2011
- Initial Focus on Arrangement and Description – a tool for Archivists'
- Second focus on Discovery and Access

Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



Functional Requirements for AIMS Hydra Head (“Hypatia”)

This work is licensed under a [Creative Commons Attribution-ShareAlike 3.0 Unported License](#).

[Functional Requirements for Arrangement and Description](#)

[A&D 00: Fundamentals \(MM\)](#)

[A&D 01: Graphical User Interface \(MM\)](#)

[Overview](#)

[Questions](#)

[A&D 01.01: Representation and manipulation of hierarchy](#)

[A&D 01.02: Drag and drop functionality](#)

[A&D 01.03: Sort records](#)

[A&D 01.04: Copy and paste of hierarchical structure](#)

[A&D 02: Technical Metadata \(PC\)](#)

[Overview](#)

[A&D 02.01: File-level technical metadata](#)

[A&D 02.02: Directory-level technical metadata](#)

[A&D 02.03: Presentation of technical metadata](#)

[A&D 03: Descriptive Metadata \(SW\)](#)

[Overview](#)

[A&D 03.01 Importing existing EAD](#)

[A&D 03.02 Viewing/editing descriptive metadata](#)

[A&D 03.03 Creating new description and intellectual arrangement](#)

[A&D 03.04 Exporting EAD data](#)

[A&D 03.05: Controlled vocabularies](#)

[A&D 04: Rights/Restrictions \(MM\)](#)

[Overview](#)

[A&D 04.01: Date-based access restrictions with automatic removal](#)

[A&D 04.02: Access restrictions to be removed manually at a later date](#)

Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



Functional Requirements for Discovery and Access in AIMS Hypatia ☆

File Edit View Insert Format Tools Table Help

Last edit was made 241 days ago by glynnedw



Normal text ▾

Trebuchet ... ▾

14pt ▾

B

I

U

A ▾

A ▾

1 2 3 4 5 6 7

Functional Requirements for Discovery and Access in AIMS Hypatia (Hydra Platform for Access To Information in Archives)

Will need a publication method that moves files to D&A environment

[D&A_01: Filter Content](#)

[D&A_02: Descriptive Metadata](#) (import)

[D&A_03: Administrative view](#)

[D&A_04: Search](#)

[D&A_05: Display / View Files / Representation](#) (text, image, audio, video, web, etc.)

[D&A_06: Registration & Terms & Conditions](#)

[D&A_07: Sort on field \(NEW\)](#)

[D&A_07B: Select](#) (individual or multiple, check or highlight)

[D&A_08: Reporting \(MM\)](#)

[D&A_09: Edit / Annotate](#)

[D&A_10: Request](#) (mediated v easy download)

[D&A_11: Upload / Edit](#)

[D&A_12: Browse](#)

D&A_01: Filter Content (ge)

Hypatia Partners

[Stanford University](#)

[University of Hull](#)

[Yale University](#)

[University of Virginia](#)



How do we get to Hypatia 1

- Archival Description – Encoded Archival Description (EAD)
 - Repository data
 - Collection data (title, extent,)
 - Physical and Intellectual arrangement

Challenges with EAD

- Encoding Standards institutional specific
- Doesn't scale for born digital archives (100,000 of files)

Hypatia Partners

Stanford University

University of Hull

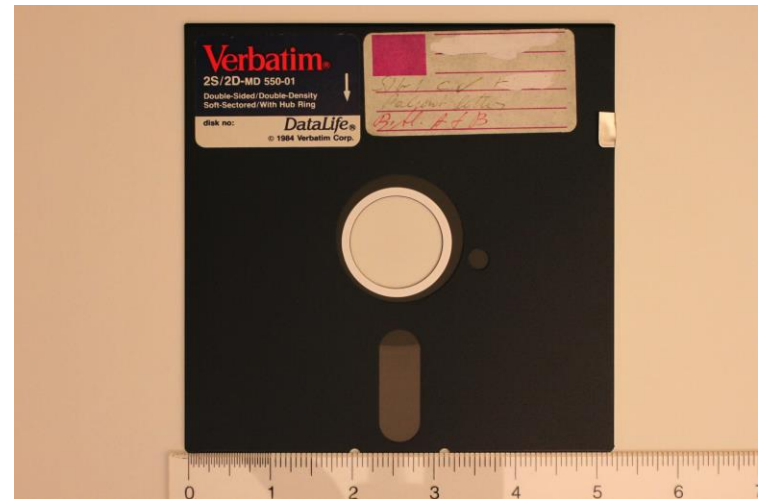
Yale University

University of Virginia



How do we to Hypatia 2

- Archival payload – disk images / files
 - Typically stored on obsolete media
 - Minimal descriptive metadata



Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



The POWER of Digital Forensics

- Specialized software to help archivists' preserve
- provenance by:
 - Migrating data off of legacy at risk media
 - Captures create, modify date, last accessed date
 - Preserves original media file paths, OS and low level formatting?
 - Original applications including fonts that created the data

Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



Digital Forensic Processing

- Archivists use commercial or open source software to tag large quantities of born digital archival materials
 - Keyword, pattern search to find files that have sensitive information (Health records, Credit Card data, etc.)
- Bulk edit tagging for restricted files, subject, source media (what disk did the file come from?)

Hypatia Partners

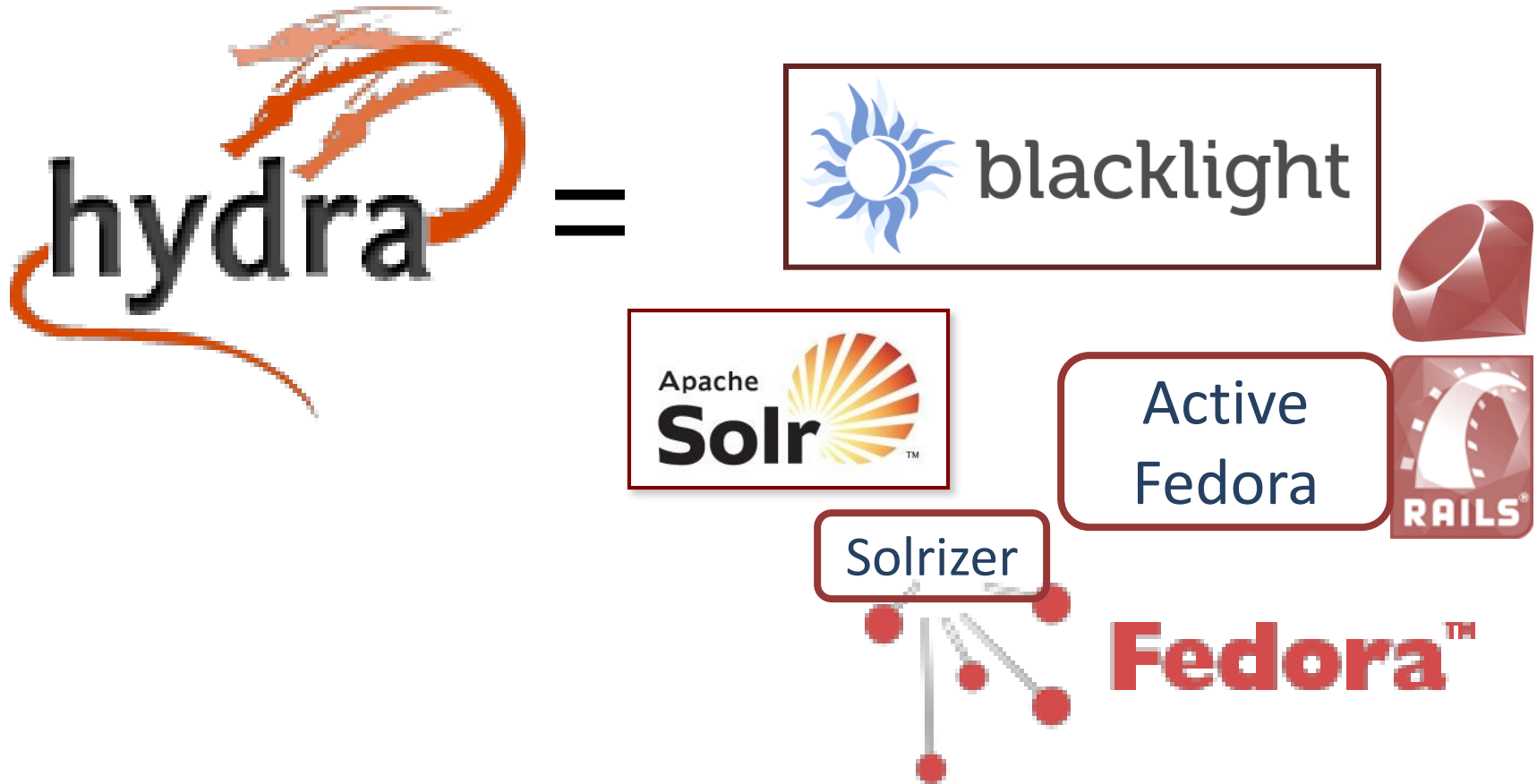
Stanford University

University of Hull

Yale University

University of Virginia





Hypatia Partners

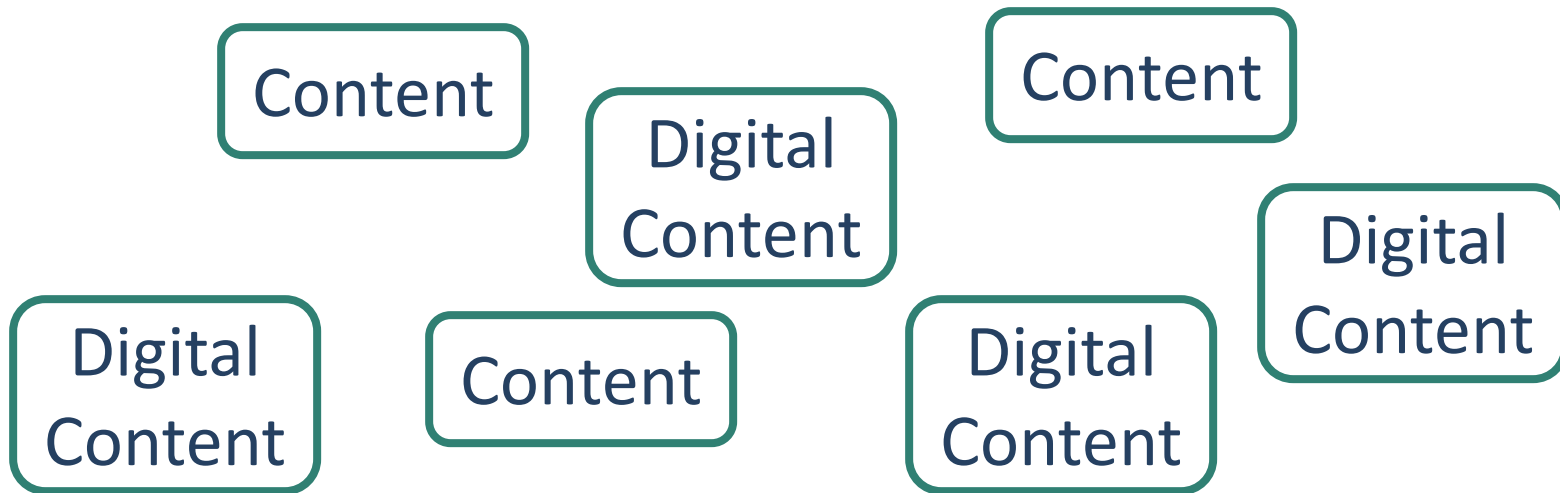
Stanford University

University of Hull

Yale University

University of Virginia





Hypatia Partners

Stanford University

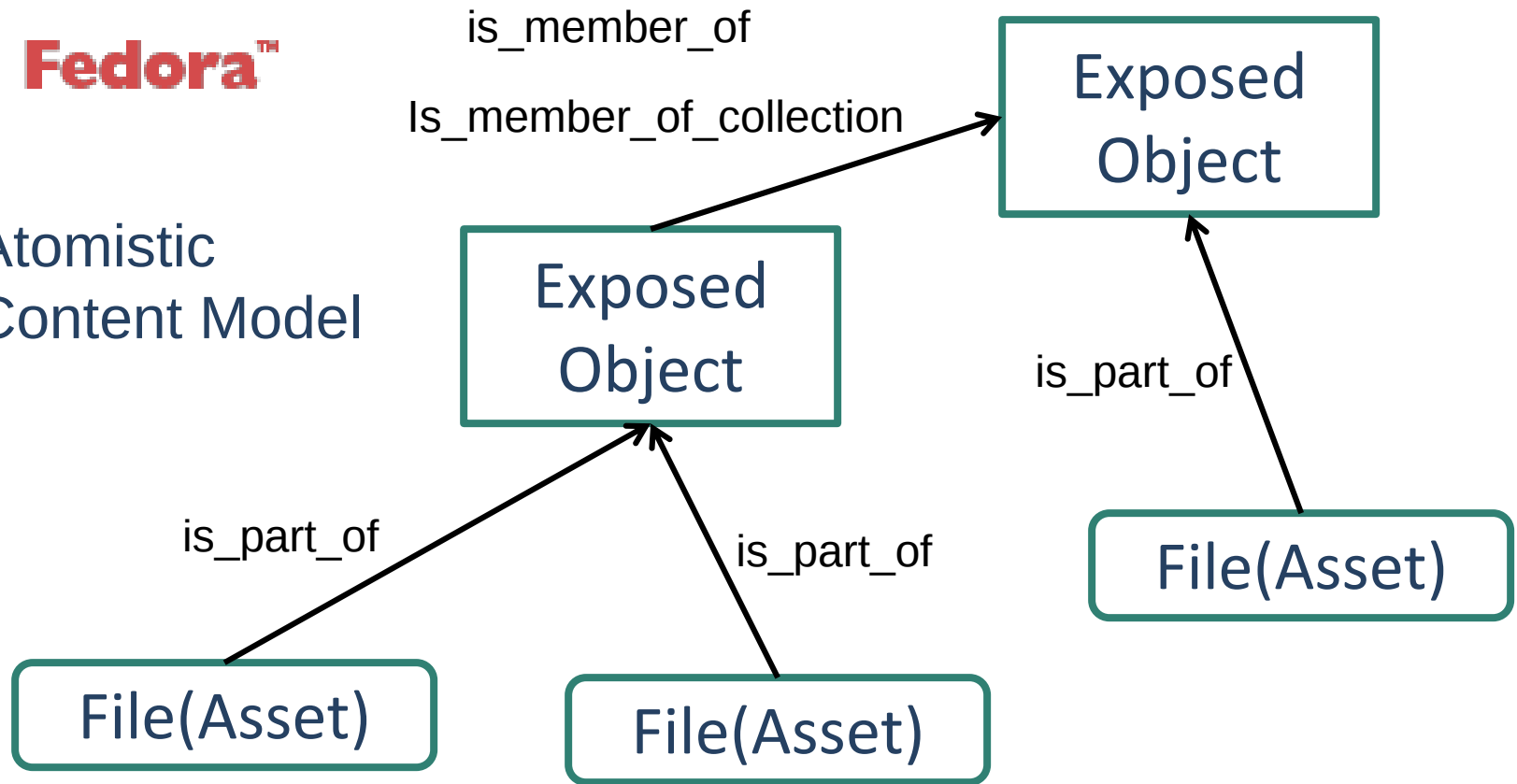
University of Hull

Yale University

University of Virginia



Atomistic
Content Model



Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



Descriptive Metadata (MODS)

Content Metadata (Stanford)

Rights Metadata (hydra)

RELS-EXT (model, parent, ...) (Fedora)

DC (Fedora)



Hypatia Partners

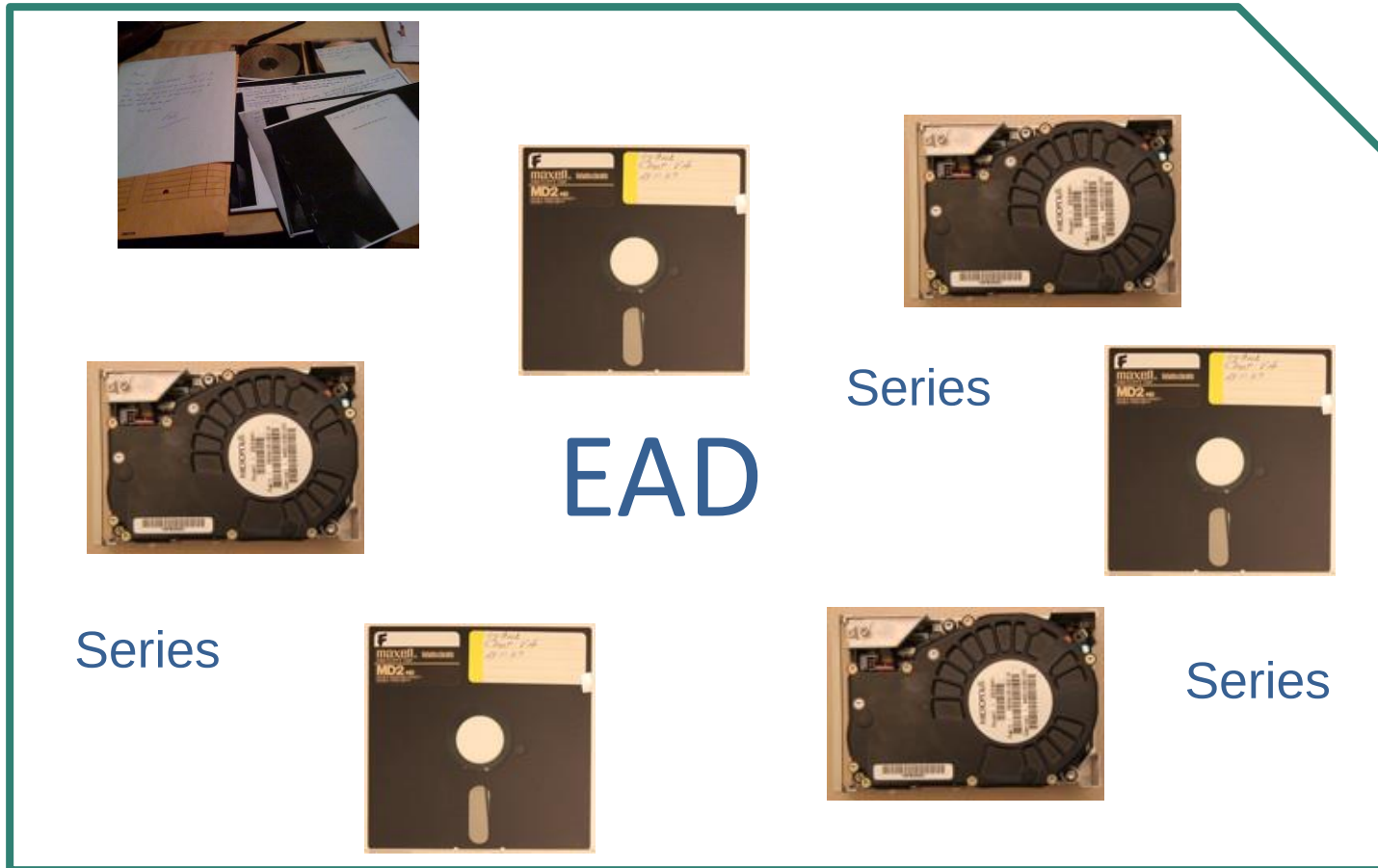
Stanford University

University of Hull

Yale University

University of Virginia





Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia

EAD



Digital
Content



Digital
Content



Digital
Content



Digital
Content



Digital
Content



Digital
Content

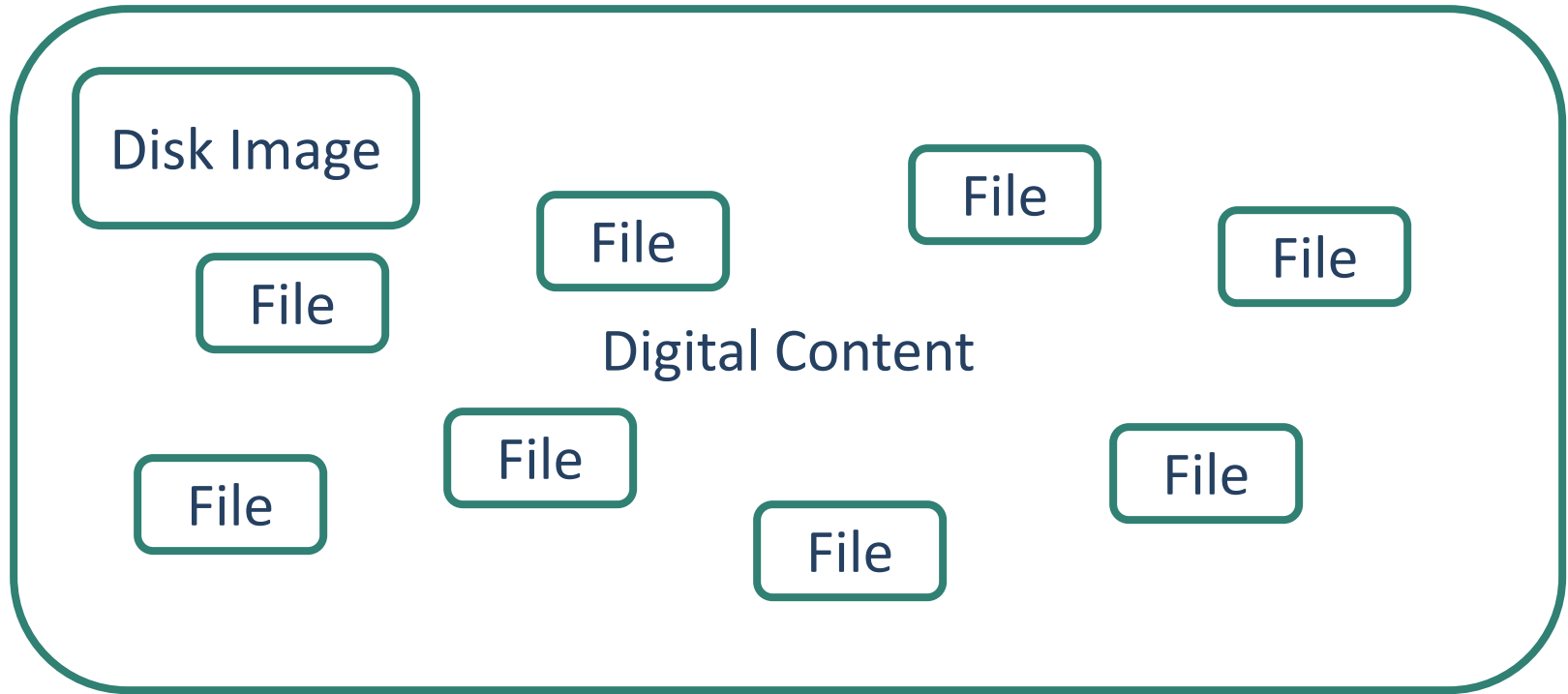
Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



Hypatia Partners

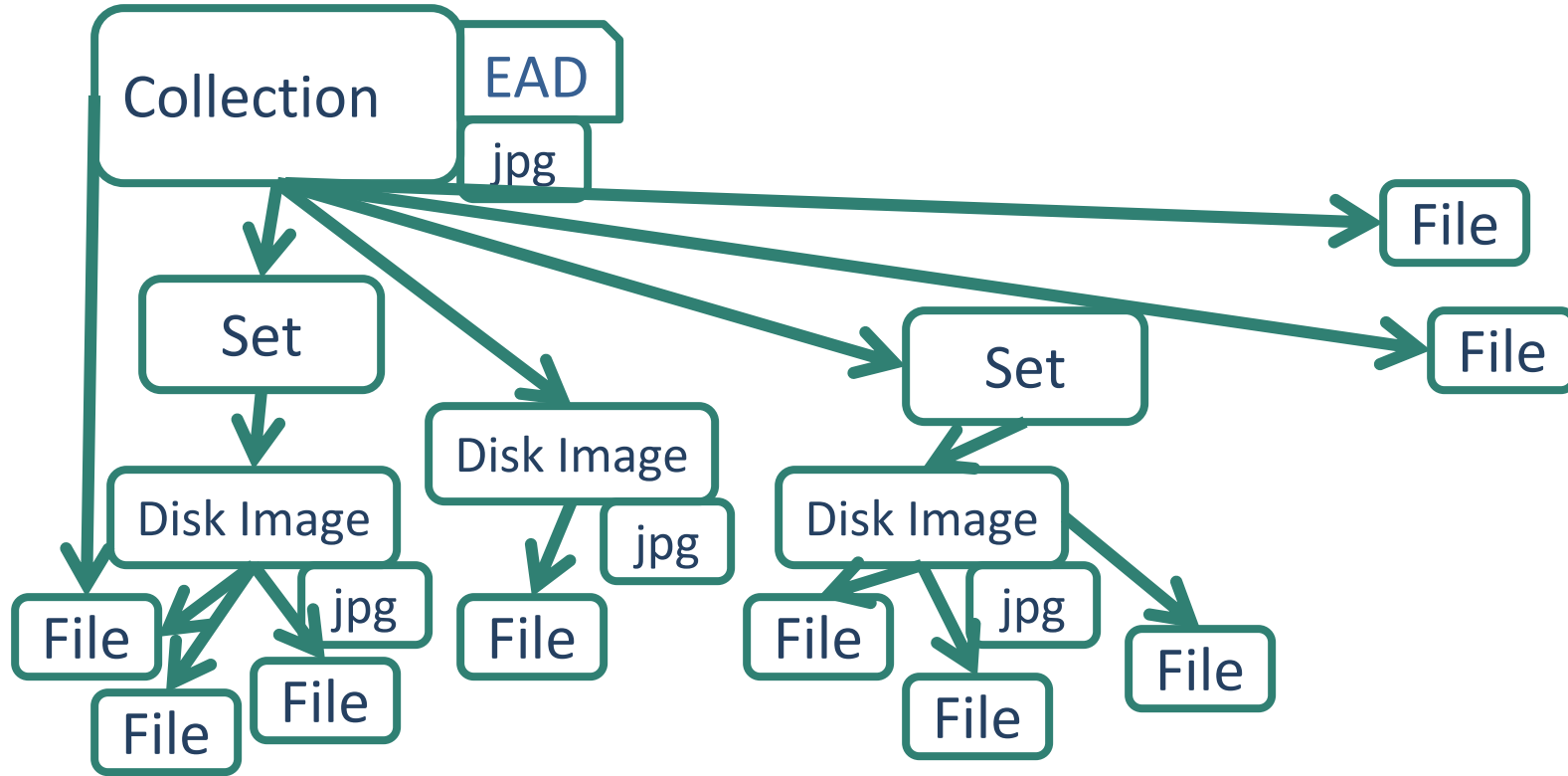
Stanford University

University of Hull

Yale University

University of Virginia

Hypatia



Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia





FTK
processing



Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia

FTK Output not designed for this:

```
//fo:page-sequence[2][fo:flow/fo:block[text()='Case Information']]/fo:flow/fo:table[1]/  
fo:table-body/fo:table-row[3]/fo:table-cell[2]/fo:block/text()
```

Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



FTK practices vary

Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia

Whole (digital) archival management

- A Case Study: Feigenbaum Papers at Stanford, prior to SALT
- Previously:
 - Files on a separate file store
 - Permissions management & preservation challenges
 - A distinct index with its own faceted browser (Flamenco)
 - A separate Drupal site for collection landing page
 - A separate Mysql db for tags
 - A separate Finding Aid, stored in Archivists Toolkit
- = a Nightmare to synchronize, update and migrate

Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



Hypatia Benefits

- Integrated solution for archival digital objects management & access
- Granular permissions management for discover, read, edit, and administrate
- Support for multiple arrangements: physical, logical, archival
- Enables ongoing processing as resources become available
- Integrated approach for digital preservation

Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



WIIFM?

- Open source code base for digital archives management
- Nucleus for further, community development
- Forensic toolkit patterns, best practices, Fedora loading scripts
- Data models for EAD and digital archival objects in Fedora
- Functional requirements for arrangement, description, discovery and access for digital archives

Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



Next Steps

- Pilot usage in a archives processing digital materials
- Another round of development
- Development of bulk permissions management, arrangement & description
- Experiment with UI and tools for archivists and for end users

Hypatia Partners

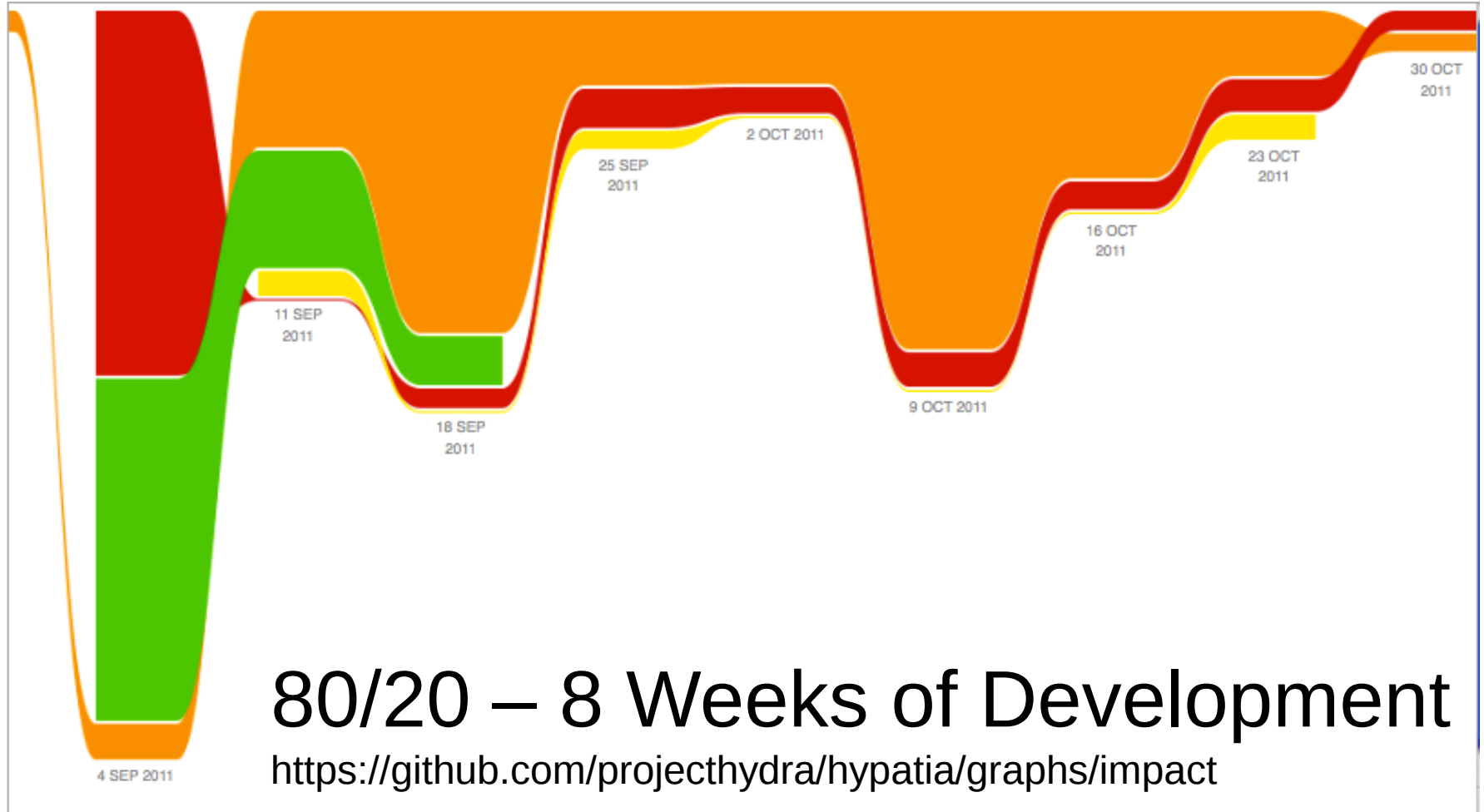
Stanford University

University of Hull

Yale University

University of Virginia





Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia



Connect

Demo: <http://hypatia-demo.stanford.edu>

Wiki: <https://wiki.duraspace.org/display/HYPAT/Home>

List: <http://hypatia-tech.googlegroups.com>

Code: <https://github.com/projecthydra/hypatia>

Hydra: <http://projecthydra.org>

AIMS: <http://www2.lib.virginia.edu/aims/>

Hypatia Partners

Stanford University

University of Hull

Yale University

University of Virginia

