#### Yale

# Digital Repository Development at Yale University Library

**Michael Dula** 

CTO, Yale University Library

December 8, 2014



Four groups comprise Library IT:

- Workstation and Technology Services
- User Experience
- Enterprise Systems and Architecture
- Digital Library & Programming Services

Metadata standards

- Discoverability
- Managing access and permissions
- Preservation



• Metadata standards: inconsistent



# Historically, how are we doing?

• Metadata standards: inconsistent

Discoverability: spotty



# Historically, how are we doing?

- Metadata standards: inconsistent
- Discoverability: spotty
- Managing access and permissions: several methods, none very granular

# Historically, how are we doing?

- Metadata standards: inconsistent
- Discoverability: spotty
- Managing access and permissions: several methods, none very granular
- Preservation: we have duplicate copies somewhere—that's enough, right?

# How did we get into this fix?

Walled gardens: with unicorns in them





# How did we get into this fix?

# LOTS of gardens, and rooms, and attics...

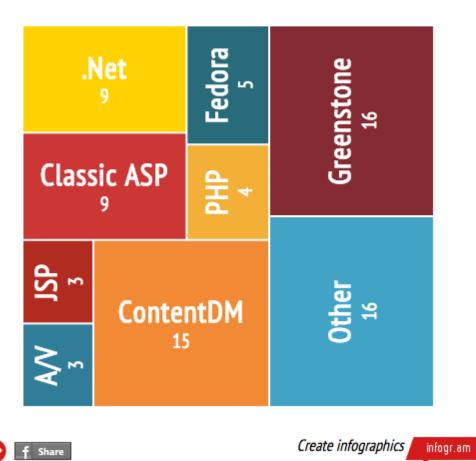








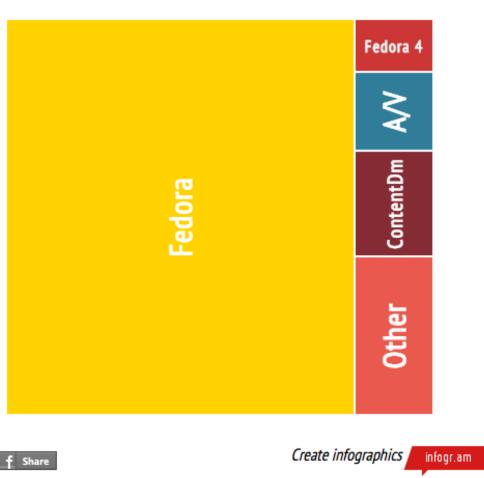






# By End of 2015...

## **Collection Technologies 2015**



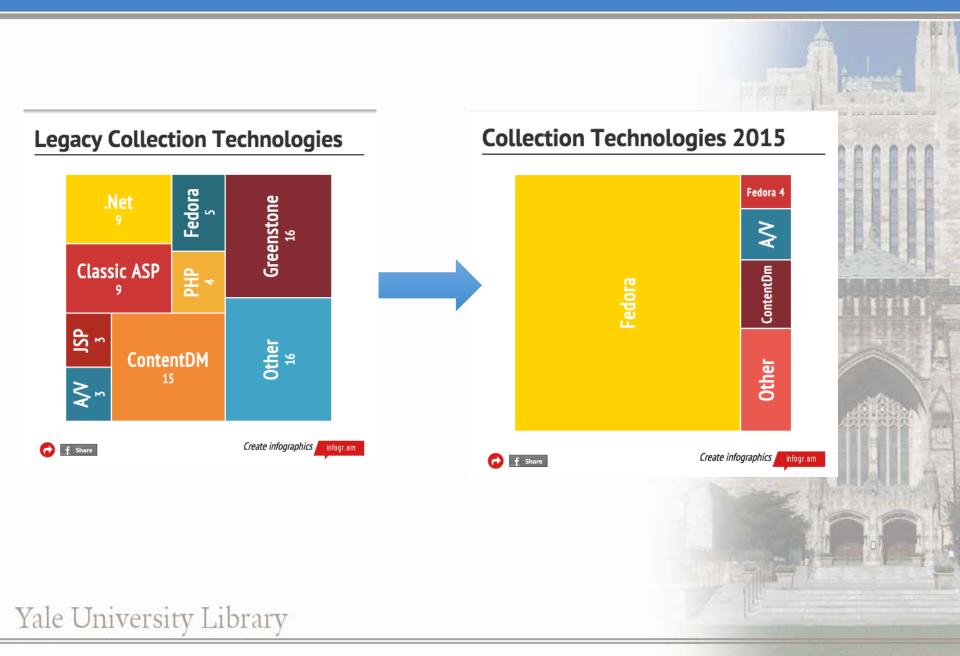






#### How do we get from here to there?

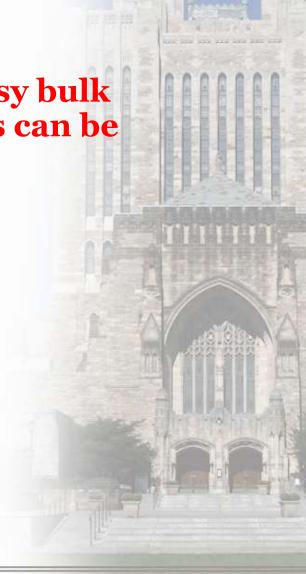
## Yale



# A Note About Legacy Migrations

#### Metadata often turns out to be:

• Imperfect. Strategy: If there is an easy bulk transformation, do it, but other fixes can be post-migration.



#### Metadata often turns out to be:

 Imperfect. Strategy: If there is an easy bulk transformation, do it, but other fixes can be postmigration.

Not meeting minimum standards for display.
 Strategy: Fix or retire collection.

#### Metadata often turns out to be:

- Imperfect. Strategy: If there is an easy bulk transformation, do it, but other fixes can be postmigration.
- Not meeting minimum standards for display.
  Strategy: Fix or retire collection.
- Highly custom. Strategy: Migrate the data, but may not have a use for it immediately.

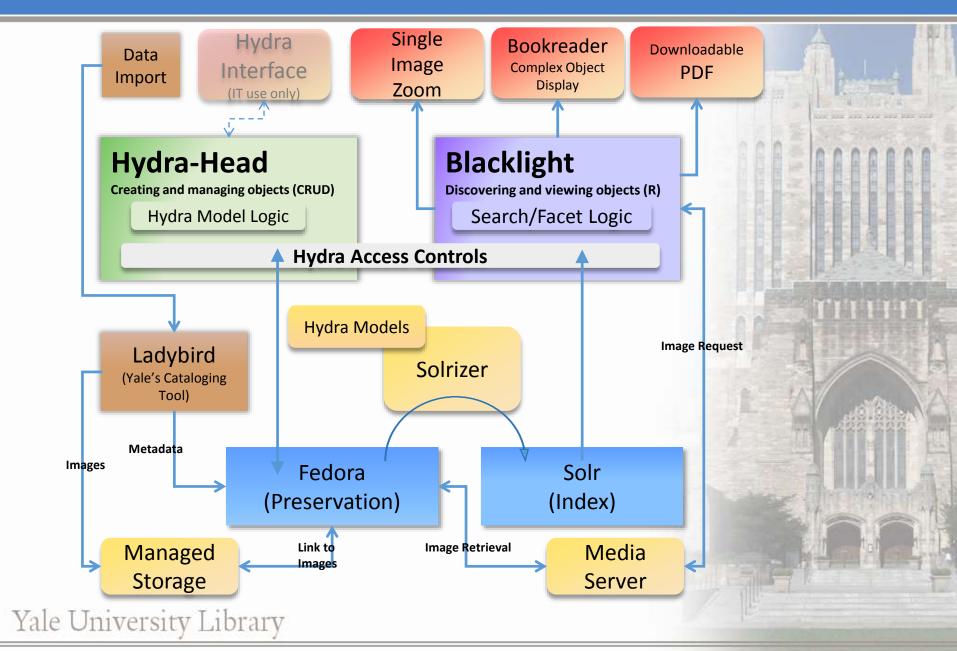
## About Hydra/Fedora

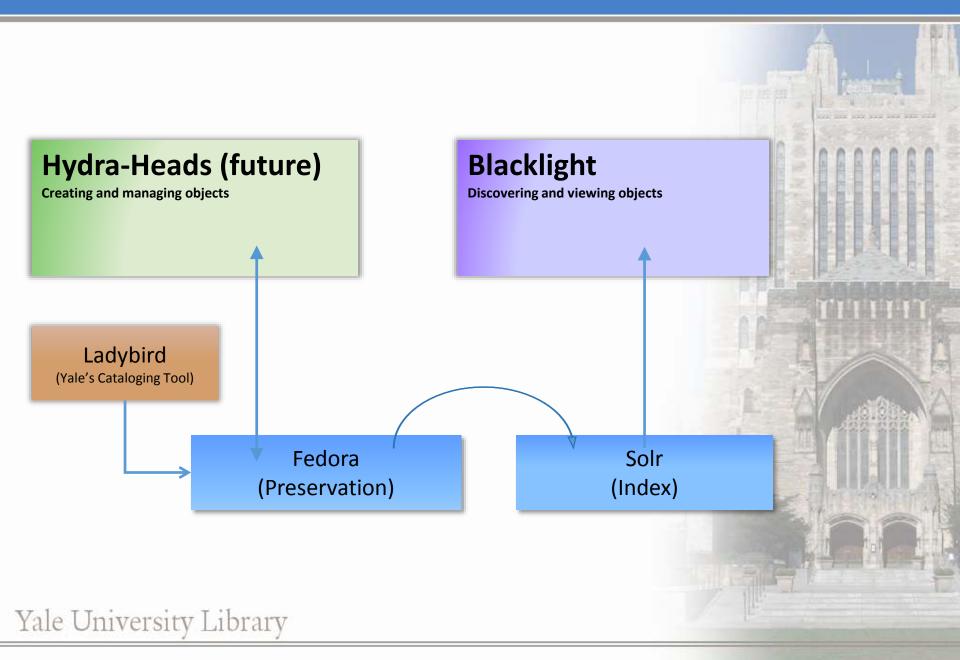


- Open-source project
- Provides a platform for digital preservation and presentation
- Over 50 Fedora Members contributing financially;
  Yale is one of these
- Yale is also a Fedora development partner, and YUL's Manager of Digital Library & Programming Services serves on the Fedora Leadership Committee
- Currently actively engaged in development of Fedora 4



- Began in 2008 as collaboration between Stanford, UVA, Univ. of Hull, and Fedora Commons
- YUL joined in 2013 as 18<sup>th</sup> member.
  Membership now up to around 25—recent additions include Princeton, Cornell, Case Western
- "If you want to go fast, go alone. If you want to go far, go together."

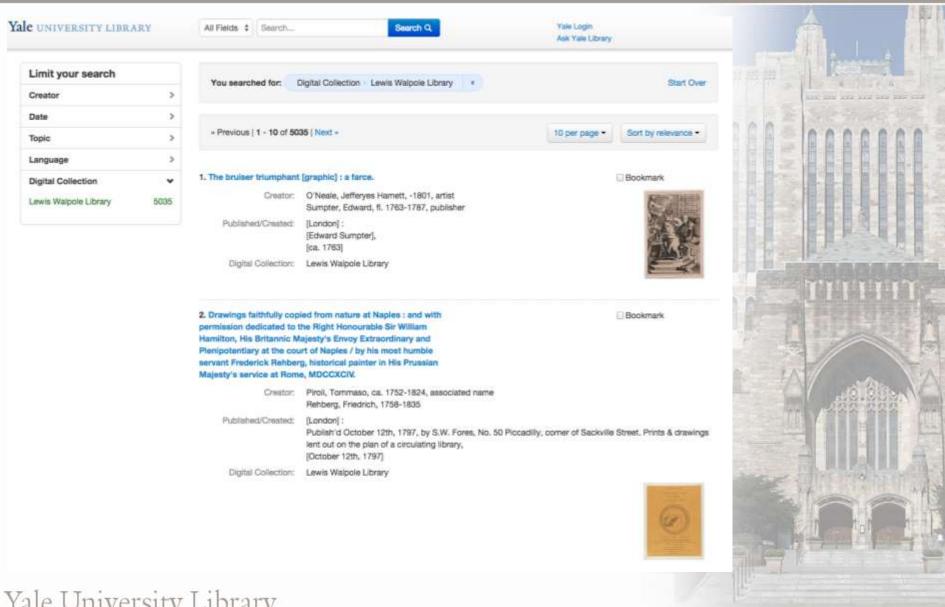




- 9 pilot collections: <a href="http://findit.library.yale.edu">http://findit.library.yale.edu</a>
- Henry Kissinger Papers
- Migrations of about 80 legacy collections
- Research Data pilot with the Institution for Social and Policy Studies

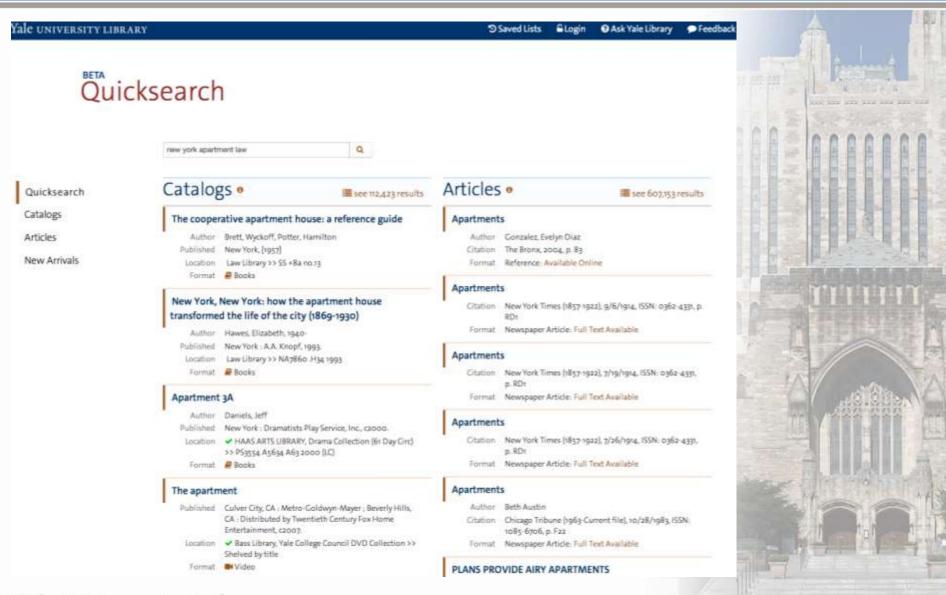
### The Findit Blacklight interface

### Yale



#### Quicksearch Interface (http://search.library.yale.edu)

#### Yale



 We have to get large—1 PB already queued up for ingest and/or migration

 We have to provide multiple levels of security, from Open Access to highly mediated

# What's next?



#### Possible Future Directions

- Curated Research Data
- Self-Archiving by Yale community
- Streaming A/V Support
- Online exhibitions
- Active Preservation tools
- Advanced Interfaces: GIS, Digital
  Humanities, Data Visualization

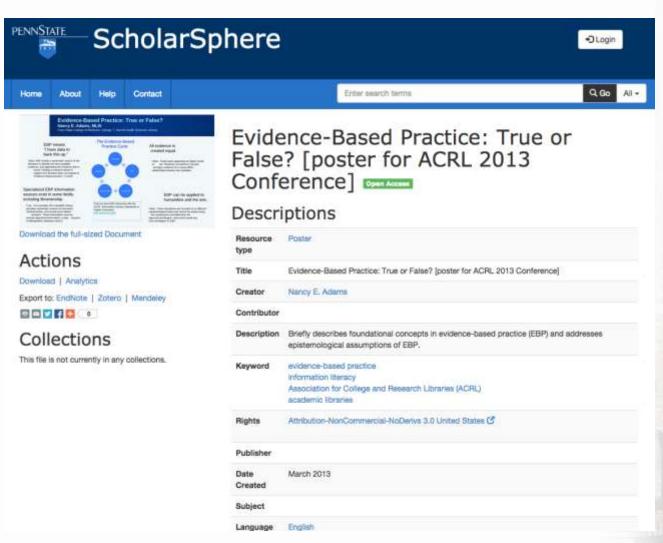


#### **Research Data**

- Pilot project underway with ISPS
  - Learn how to ingest research studies as complex objects
  - Preserve them and provide permanent links
  - How do they display in our standard Blacklight interface?
- Develop custom interfaces around this content type
- Integration with other systems



#### Yale





http://www.avalonmediasystem.org/project





#### Featured Film



The Masters of Disaster (1985)

#### Featured Film Collection



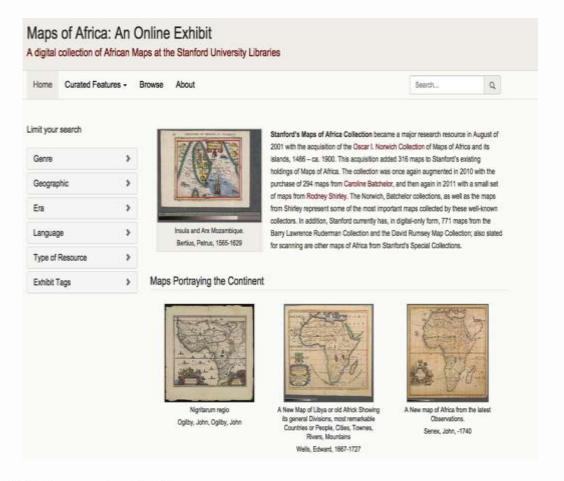
Indiana University Libraries Moving Image Archive

#### Media Collections Online houses audio and video collections.

- · Use the search box, or browse using the terms on the left side to discover content.
- · Some content is only available when you sign in with your ID and passphrase.
- Login to Oncourse or Canvas to access course materials.



https://exhibits.stanford.edu/maps-ofafrica/spotlight/default-exhibit





# **Digital Preservation Services**

- Bit Preservation
- Secure Storage with Managed Access
- Obsolescence Monitoring
- Provenance and Authenticity Assurance
- Standards Compliance
- Format migration and emulation services

#### **Possibilities for Advanced Interfaces**

GIS

Data manipulation and visualization

Digital Humanities



How can we do all this at speed and at scale?

It took 300 years for the Yale Library catalogs to grow to around 10 million items.

Just ONE large digital collection that we are currently working on will contain about 10 million items.

# Yale

# Questions?

**Michael Dula** 

