COAR-SPARC
CONFERENCE 2015
Organizational Models for
Research Data Management Services

Gerald Beasley

Wednesday, April 15th, 2015
Porto, Portugal
“Hydra North”

UAL’s Next Generation
Digital Asset Management System (DAMS)/
Digital Repository Project

GEOFF HARDER
Associate University Librarian, University of Alberta
February 2015
Hydranauts

Developers - Chris, Tricia, Piyapong, Weiwei, Peter, Sam, Natasha
Metadata - Sharon, John, Bob
Systems/Infrastructure - Neil, Henry, Kenton
Services - Leah, Anna, Chuck, Larry, Peggy Sue/Debbie, Umar
...and the teams at Stanford, Duke, Indiana, Virginia Tech, Penn State, Columbia, Yale, Princeton, Cornell, Notre Dame and more.
learned by doing
built where vendors dared not go
“go big or go home” so went big
blazed new ground

where we came from
where we found ourselves

many silos
many standards
many workarounds
too many close calls (even one is too many!)
bespoke software implementations tied to legacy technical and policy choices

**many** projects behind us

**MANY** more projects ahead of us
where we need to be

Sustainable
Scaleable
Stable
Secure
Supported
Successful
what we need to do

trustworthy DAMS
consolidated access, discovery, preservation
characteristics include:
scalability
extensibility
predictability
stability
open

community alignment

what is Hydra

A robust repository fronted by feature-rich, tailored applications and workflows (“heads”)

One body, many heads

Collaboratively built “solution bundles” that can be adapted and modified to suit local needs.

IR, ETDs, image, media, archives, special collections, research data, exhibits,...

A community of developers and adopters extending and enhancing the core

“If you want to go fast, go alone. If you want to go far, go together.”
who is Hydra?

Coherence at Scale in action.
where it aligns

- Fedora for common management and preservation of all objects and metadata
- Blacklight as a common, coherent discovery tool (reducing silos)
- Community standing behind core DAMS; opportunity to invest our human resources toward solving new challenges
- Consistent content models and services across collections, e.g. book/image viewers
- A sustainable, open foundation for the future
challenges

Technical
  framework, not bundle
  Ruby on Rails
  Fedora 3 or Fedora 4
  metadata and RDF

Organizational
  leaving the way it was for the way it will be
  agile project management
roadmap & timeline

April 2015 - production version of Hydra North
Phase 1 - ERA + “simple” research data

May 2015 forward - Phase II, III, IV, ...
digitization [+ OER]
thesis submission management via Alfresco
archivematica + data (Dataverse)
multimedia with streaming (Avalon)
archives/electronic records (Alfresco)
on the road

- full partner status (final steps)
- already contributing code, joining working groups, participating in Hydra events, weekly telecons, email list discussions
- campus conversations with Digital Learning Oversight Committee, DigiTAL, others
  - preservation of MOOC assets, multimedia, OER objects
- Arts Resource Centre (ARC) - Hydra co-development
- exploring linked open data possibilities (LOD), geospatial (GeoHydra), multimedia (Avalon), and much more
more info about Hydra

Public project pages:
http://projecthydra.org/
https://github.com/ualbertalib/HydraNorth

Examples of Hydra implementations:
http://projecthydra.org/apps-demos-2-2/
http://library.ucsd.edu/dc
Scrum
Agile Development in DI/ITS

Library Council, 25 Feb. 2015
Peter Binkley
The Agile Manifesto

As developers we value ...

- **Individuals and interactions** over processes and tools
- **Working software** over comprehensive documentation
- **Customer collaboration** over contract negotiation
- **Responding to change** over following a plan

http://www.agilemanifesto.org/
Scrum Roles

- **Product owner** - Geoff
- **Stakeholders** - Leah, Sharon, Chuck, Umar, Larry, ...
- **Development Team** - Weiwei, Piyapong, Tricia, Chris, Henry, Neil, Sharon, ...
- **Scrum Master** - Peter
User Stories

"As a graduate student, I want to deposit my dissertation in the University's repository so that I can fulfill the requirements of FGSR and complete my degree."
Product Backlog

- The set of stories we're going to implement in this product
- Ordered by priority
- Requires constant grooming
- Requires constant communication between developers and users/stakeholders
Sprints

- Two weeks of work by the dev team on a selected list of stories
- Daily standup meetings (15 minutes)
- Concludes with a retrospective meeting involving the dev team, product owner and stakeholders
Cult of "Done"

Each sprint should end with a working piece of software that performs all the functions of the user stories selected for that sprint. A story once moved to "done" should not have to be revisited.
Managing Sprints on GitHub

- A story (with notes)
- A milestone (the stories for one sprint)
- A pull request (changes to the code to implement a story)
- An upstream pull request (our contribution to the community)

https://github.com/ualbertalib/HydraNorth
Why Scrum?

Scrum facilitates collaboration by framing every block of work within a conversation between the person who knows how to do the work and the person responsible for seeing that the outcome is what we need. The conversations revolve around user stories, so that we never lose sight of the final purpose, and they are structured in such a way that the history of the project can be captured for future use and reuse.