Hydra and the future

A roadmap of sorts…

Richard Green

UNIVERSITY OF Hull

Overview

- Hydra's strategic plan
- Priorities for 2015
- Community cohesion
- Product cohesion
- The DPLA and the IMLS grant
- Other major initiatives
- The other strategies?

Hydra's Strategic Plan

- Originally seven, now eight elements:
  - Strategy 1: Develop a full suite of Solution Bundles
  - Strategy 2: Develop turnkey / hosted solutions
  - Strategy 3: Grow the vendor ecosystem & define the vendor community structure
  - Strategy 4: Training framework
  - Strategy 5: Documentation framework
  - Strategy 6: Technical framework
  - Strategy 7: Adaptive Community framework
  - Strategy 8: Marketing and Community expansion

Priorities for 2015

- ***Community Cohesion*:**
  - How can we become more intentional and effective in structuring community growth and distributed activity leading to collective benefit?
  - How can we best foster Partner and Community acculturation "onboarding", alignment, and contributions?

- ***Product Cohesion*:**
  - How can we become more intentional and effective in
    1. producing feature-complete solutions;
    2. developing an easy-to-adopt solution;
    3. developing an easy-to-maintain solution;
    4. facilitating the sharing of code and innovations across the project?

Community cohesion #1

- Hydra's Community growth is only going to continue. We need to explore Community needs and opportunities...

- We should formally organize Interest Groups and Working Groups to pursue already identified interests and needs, and to pursue the most critical of the project's eight "strategic planks", based on importance, urgency and opportunism/feasibility. The most important of these are...

    - Web Presence: to ensure Hydra's web presence (web site, wiki, Github) are up to date and appropriately reflective of the project's professionalism and quality
    - Marketing: to proactively promote the project in relevant forums, and elevate its reputation and recruit new adopters, Partners, vendors, funders and allies
    - Membership: establishing and activating mechanisms to review the level of engagement and value each Partner is giving to and deriving from the project, helping onboard new Partners and new staff at existing Partners, and helping knit the Partners closer together as collaborators overall

Sidebar: Current Interest and Working Groups

Reminder: WGs have deliverables and a timeframe; IGs are discussion. A WG can morph into an IG and vice-versa. IGs are open to all adopters, WGs require an IP agreement from participants.

- Archivists Interest Group
- Digital Preservation Interest Group
- Geospatial Interest Group
- Marketing Interest Group
- Metadata Working Group
- Page Turner Interest Group
- Service Management Interest Group
- User Experience Interest Group
- Web Presence Working Group
- (Portland Common Data Model group)
  - …now bigger than just Hydra folks

All meetings are advertised, with reminders on the lists, and each group publishes minutes of its calls.

Community cohesion #2

- Identify, engage and firm up an organisational home for Hydra that can provide specific banking and legal administrative services

  - *This task was completed early in 2015. Hydra has signed a Memorandum of Understanding with DuraSpace describing how they will provide us with specific financial and legal services. The MoU commits Hydra to the exploration of this second bullet point:*

- Determine whether the Project's needs would be well served by a closer collaboration with such an organisation, allowing us potentially to leverage further services (such as web hosting, marketing, fund raising, etc)

The "product"

- Currently we have the Hydra Framework consisting of a considerable number of gems, including
    - Active Fedora
    - Hydra Head
    - Hydra Jetty (Solr and Fedora)
    - Opinionated Metadata (OM)
    - Hydra (*one gem to rule them all and in the darkness bind them?*)

- Sufia, a Rails engine for creating a self-deposit institutional repository

- The Avalon Media System solution bundle. This is a turnkey-esque Hydra for media management and delivery

- Etc...   - but implementations tend to be driven by local needs; data modelling, technical design, and code development are often done in idiosyncratic ways that inhibit diffusion and sharing

The architectural "big picture"
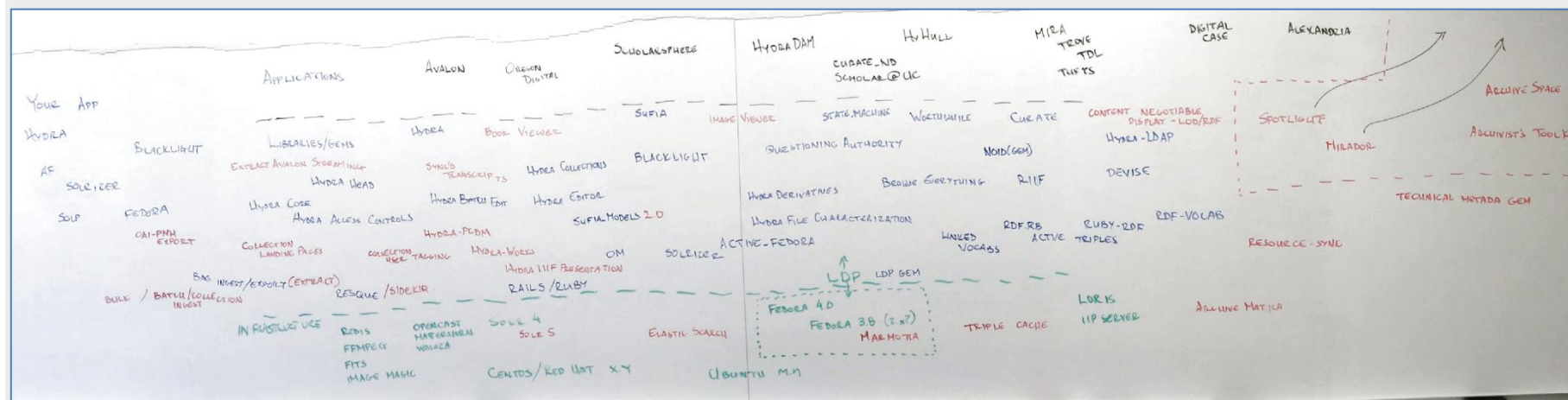
As at 27 March 2015

You won't be able to read the whiteboard, but this might give you some idea of the breadth (!) of our developers' vision…   Needs a big whiteboard.

Product cohesion

- Hydra has been successful in producing a robust technical platform but we can/should do better at:

    - producing feature-complete solutions - "solution bundles"
    - producing consolidated content models
    - facilitating code sharing and reuse

- The advent of Fedora 4 gives us a possibly unique opportunity:

    - Existing Fedora 3.x users will, at some point, need to migrate. Consolidation could be part of the migration process if there were compelling reasons

    - To bring this about we would need, at least, to produce and document a common approach to content modelling in Fedora 4 with migration tools and good documentation – including exemplars

Product cohesion #2

- We should perhaps aim towards 'Hydra in a Box':

    - polished and feature rich IR solution
    - easy to install
    - easy to maintain
    - cloud-ready
    - widely adoptable
    - based on Fedora 4, Hydra 9 and the Portland Common Data Model

- Something / some things to 'cohere' around

- A common technology base that could be built on

- Potentially *very* expensive to develop but "if we don't do it, someone will do it for us possibly outside of our control"

Enter the DPLA DIGITAL PUBLIC LIBRARY OF AMERICA

- The Digital Public Library of America aggregates metadata records from libraries, archives, museums and other cultural heritage institutions to provide a single portal

- Collection of the metadata records would be much simplified if its providers used a common tool to steward and submit assets
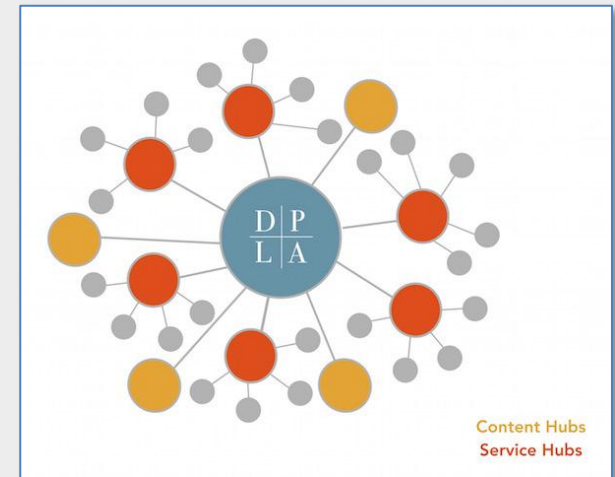
Image: DPLA website (CC-BY)

IMLS grant

- In January, the DPLA with Stanford University and DuraSpace, submitted a grant proposal to the US Institute of Museum and Library Services (IMLS) for $2m matched funding over 2½ years to build a turnkey, cloud-ready, Hydra-based solution that could be easily adopted by DPLA hubs (and others)

- Late March 2015 proposal accepted

- "They will collaboratively extend the existing Hydra project codebase to build, bundle, and promote a feature-complete, robust digital repository that is easy to install, configure, and maintain—in short, a next-generation digital repository that will work for institutions large and small, and is capable of running as a hosted service." *DPLA press release*

- Watch this space…

Other major initiatives: Avalon

- The IMLS grant is not the only major initiative for 2015. Indiana University and Northwestern University have been awarded $750,000 by the Andrew W. Mellon Foundation to enhance Avalon, the Hydra based media system:

  1. develop additional features and functionality for Avalon to better meet needs of collection managers and users
  2. conduct studies of use of audio and video collections by researchers in humanities disciplines to help ensure future support for scholarly use
  3. integrate Stanford University's Spotlight exhibit tool with Avalon to allow librarians, archivists, and scholars to showcase and provide additional context for media items and collections
  4. develop and implement a community-funded business and governance model to sustain ongoing support and development for Avalon
  5. deploy Avalon in a hosted software-as-a-service model for use by institutions that need the functionality of Avalon but would prefer to utilize a cloud-based software-as-a-service option rather than support a locally hosted instance

Avalon development proposal

Other major initiatives: HydraDAM2

- Indiana University and WGBH (the Boston-based US public broadcaster) have been awarded a $399,239 grant from the NEH Preservation and Access Research and Development Program to support the two year project "HydraDAM2: Extending Fedora 4 and Hydra for Media Preservation."

- "The project, will:

  1. Extend the HydraDAM digital asset management system to operate on Fedora 4
  2. Develop Fedora 4 content models for audio and video preservation objects, including descriptive, structural, and digital provenance metadata, based on current standards and best practices and utilizing new features in Fedora 4 for storage and indexing of RDF
  3. Implement support in HydraDAM for two different storage models, appropriate to different types of institutions:
     a. direct management of media files stored on spinning disk or on tape in a hierarchical storage management (HSM) system; and
     b. indirect management and tracking of media files stored offline on LTO tapes
  4. Integrate HydraDAM into preservation workflows that feed access systems at IU (Avalon) and WGBH (OpenVault) and conduct testing of large files and high-throughput workflows
  5. Document and disseminate information about our implementation and experience to the library, archive, digital repository, and audiovisual preservation communities" *HydraDAM2 proposal*

The other strategies?

We haven't ignored the other strategies, just prioritised some...

- We are still looking to get new **vendors** involved to support further, healthy community growth

- We continue to develop **training** modules to assist new and existing adopters

- Probably no open-source project has "excellent" **documentation**, but we've made great strides and continue to improve it

- We're getting more proactive about **marketing** and have a new Marketing Interest Group (which will likely soon spawn a Working Group or two)

A closing thought...

"The future started yesterday, and we're already late."

John Legend

Thank you!

Links

- Hydra website: http://projecthydra.org

- Hydra wiki: https://wiki.duraspace.org/display/hydra
    - or http://projecthydra.org/wiki

- Hydra Github: https://github.com/projecthydra

- Avalon: http://www.avalonmediasystem.org