Hydra-in-a-Box Focus Group Lunch at Open Repositories 2015

On Wednesday, June 10, 2015, the HyBox team invited OR15 attendees to an informal discussion over lunch. We shared our vision for Hydra in a Box, and solicited questions and input from those who showed up. We had three full tables and it was time well spent. Transcription of the raw notes from the session is provided here.

Table 1

- migration tools: will there be any?
- ArchivesSpace
- content types:
  - Newspapers!
- IU has many ContentDM instances. Mixed collections in some, dedicated solutions at others. Would like 1 technology solution to support them all.
- DSpace migration. No tailored views, tailored functionality by content type.
- DSpace documentation a really good exemplar to follow. High quality, useful.
- ArchivesSpace functionality for digital objects is poor; can HyBox cover that?
- have really clear scope; don't bleed into other products that do their own things really well
- can you support collaborative research space?
- product consolidation would help keep the portfolio of things to update smaller (this is attractive)
- data: ability to archive from other local systems (hit the "archive" button)
- migrating images out of ContentDM and into Oregon Digital; image processing is drag on tput. Think about upfront
- bulk deposit
- proxy deposit
- how to do outreach?
  - podcasts
  - blog posts
  - screencasts on YouTube (à la Spotlight)
  - publish a project timetable, and be honest when we're behind
    - be really honest and transparent about what you're doing and when
  - for documentation, be thoughtful about who the audience is. Don't make assumptions about who is going to read it. Organize it well by task.

Table 2

- use cases on IR functionality
- use cases on traditional digital collections and special collections, especially from smaller institutions
- bulk import and management
- data sets and geospatial metadata
- desire to integrate other products in the HydraSphere into the
  - metadata remediation and creation features. Quality metadata upfront, remediation capabilities afterwards
- support for controlled vocabs
- provide sets of templates for different types of resources. No code required.
- geocoding!
  - see the community converge around a product to support maintainability. (Maintainability as a primary design objective)
  - support for locally branded / customized versions without having to have separate codebase
  - interest and concern in helping push Hydra forward, but no forking!
  - address small institutions that have longstanding needs
  - a lot of discussion to articulate what niche this product fulfills, and dealing with scoping
    - can't be both an IR and a digital collections tool?
  - a lot of interest in hosting solutions
  - a lot of discussion about workflow. Implement it well and let it be flexible, but make sure you understand it for your workflow. don't try and intercede in a digitization workflow, because that is inst. specific. make sure you don't hardwire workflows that shouldn't be hardwired.
  - a lot of talk about consortial hosting.

Table 3

Interests

- tech stack
- how will small schools engage
- can we use Hydra Powered DPLA Aggregator that we just built (Temple)
  - there’s a need for DPLA service in our state (Colorado) especially for getting content from small institutions in
    - can the hosted service option be stood up by a medium-sized institution?
    - or is there a 3rd party hosting service that could be pointed to?
- possibly being a Hydra hosted service provider
- using HyBox w/ a very small dev team. Especially for special collections
- to be able to use HyBox with GeoBlacklight
- DevOps--how will it actually deploy?
- can Hydra run as a Platform-as-a-Service?
- general curiosity: what is it?
- aggregate to Europeana instead of DPLA
- understand if HyBox could fit into a complete overhaul of institutional digital library ecosystem
- bundling for local install. This is hard. Can we share intel across HiEd projects with like needs? (SEAD, OLE, Vivo)
- architects for Cloud. This is hard. Can we share intel across HiEd projects with like needs?
- hosting 7 different Hydra instances; can we get just one version of the app to work for all? Could we even do hosting out of a single instance?
- want to replace DSpace 1.7 with HyBox; need feature parity for important functions
Questions we should thinking about for community engagement / information:

- what is it?
- what is the timeframe?
- will it work for ____ <gis, digital collections, newspapers>?
- Look at Omeka—kind of similar for making content accessible. They added a hosted service option later.
- Who is in charge of the project? Hydra or DPLA? E.g., DPLA doesn’t take newspaper or IR content. If they are in charge, does that mean HyBox won’t support these content types?
- What about modularity? Can we still do a local install and extend it with other Hydra components?
- If an institution is ready to start on Hydra now, can they start with HyBox? How?
- What’s the migration path to HyBox from institutions with a different Hydra stack?
- Should I wait for HyBox to be done before doing anything Hydra-related?
- Can I contribute to the project?
- Who is going to install it, if anyone, during its first 2 years?
- MUST BE MULTI-TENANT.
- How will updates be made? Versioning & modules will present a challenge if there is too much customization
- How will HyBox gems / components interlock with Hydra gems in the wild?
- If HyBox lets people focus on services and collections rather than tech, it will be wildly successful
- Look at modern web application architectures. Like Slack. Simple interface plugs into remote services.
- Look at WordPress – plug in architecture w/ good versioning support
- Look at Drupal Gardens. Same as above.
- Communications
  - how do you plan on collecting requirements
  - where can I eavesdrop?
  - what can I do to participate?
  - Make sure to talk to small schools
  - Use blogs, wiki, informational calls to the community, +1 votes on requirements