January 28, 2016

Technology infrastructure and staffing discussion

Angela, Missy, Stephen, Liz, Gwen, Tom, Terry, Nathan

Terry

--Hydra option: Individual with developer skills, until it’s out of the box it’s not out of the box. Someone who is familiar with DevOps, will need to configure Solr and Fedora. Need to maintain them on their own, write code specifically around Fedora. Ruby on rails work to make changes on model. Any kind of data augmentation will need serious changes on ingest.

These folks need to be familiar with the community. Fedora is niche software, Fedora 3 is a little obscure because it’s homegrown. Finding a developer is hard. Ruby on Rails developers are easier to find however there is learning along the way probably a year doing it. There are contractors that do this work, i.e. Data Curation Experts (DCE), they will project manage the work start to finish, recommend components to the stack. OSU has people who do this. OSU worked with Hydra community. Case Western has used DCE to migrated Fedora to Hydra. They seemed pleased with the work. Their assumption was that DCE would do migration. Case and had a staff member that would learn Hydra from DCE and would do the maintenance and now the person left, so now they’re a bit concerned.

Hydra community is university based, with 30 institutions that support the Hydra. Cincinnati is a member and OSU is likely to join. The group is fairly grass roots; they manage and support the project. When things go wrong, it is difficult getting answers, getting solutions takes a lot of work. Hydra has a larger community. Hydra in a box will eventually simplify. Not sure how much Hydra in a box will deal with DPLA stuff. Anyone who is going to implement Hydra it’s an 8-12 month project and that’s an aggressive timeline because the technology stack takes a long time to learn Fedora, interact Fedora through Ruby, etc.

Repox is a Lampstack. Terry indicated that it’s relatively easy to run. There’s work to be done, creating profiles. Fedora is challenging for folks in libraries, it doesn’t come out of the box, it’s what the current Hydra is using. Repox/Lampstack uses Java script the same that works with a DSpace instance. Europeana has used Repox, two contributors are doing most of the work with it at Europeana. A metadata person who handles building the multiple XSLT is needed for metadata ingest. Repox is lowest level of pain. New York folks are using Repox, it works for them, but they’re actively looking to get away from it due to sustainability and scalability. In the long term we’ll need to migrate.

Terry noted that the application to DPLA, if we’re accepted they currently expect the state to be ready to go. We need to know that what we have will be up and running, that we have the proven ability to send metadata to DPLA. What can we do this within 3 months after being accepted at DPLA. That has implications in technical solutions and staffing. We need to determine if we need to comply with the 3 month implementation. Liz suggested that Ohio and other states that don’t have an implemented shared repository is different that the organizations that DPLA has previously worked with. We need to contact DPLA and indicated that we aren’t really going to move forward with an implementation until we know that we’re accepted. Liz will email Emily to find out what Ohio’s options are.

Technology working group will do an analysis of both option including what staffing required and pros and cons of Repox and Hydra. Each approach is designed to work in organizations that look like the organizations that created them. Assumption that you’ll get the staff to look like them.

Technology working group will be meeting to review options. The plan is to review the options and then make recommendations. Terry is going to write up options that are viable.

Where does the technologist reside? We need to have a discussion with DPLA on a timeline. If we want a Hydra solution, we build it into the timeline. The person needs to be able to talk to people doing the same thing. Idea of outsourcing to a technology firm may be off the table.

Missy—have only so much money, have to go through state processes and requirements. Can’t hire IT staff, only way to hire a programmer and do the technology piece as a software piece. Terry recommended hiring DCE if going with Hydra.

Angela—remember we’re talking about a 3 year pilot, technology will change.

Group discussed a strategy of going with the easiest possible technology path and build technology change into the sustainability. This strategy will be to use a Repox lampstack. Terry indicated that he will talk with the folks at North Carolina Digital Heritage. See how much can we borrow.

Missy: Hiring an outside contractor; if it’s a contract for over $50K then go to the state controlling board. Independent company to get the stack up and running. UPDATE: Missy reported that the Library doesn’t need to go to the controlling board for this project.

Does it include hosting services? Missing indicated that this would be separate.

If it’s Repox—the project would include setting up the software, working out bugs, then the Library would step in and use it. It would be doable. Might want to keep them on retainer for disaster. Terry asked if there would be both a development and production environment? We should consider need a back-up system in place.

Who will be doing the work of profiling and feeds? There will be onboarding process where profiles are created, i.e. this is how this institution’s metadata looks.

Who will be managing the system?

Terry will load the Repox to test harvest some sites and get an idea to do normalization. Could try to do something within next couple of weeks. Initial set up looks easy. Harvesting OAI content will take awhile. Will talk with NC about their Repox installation and their long term support.

Missy—Timeframe, would like to see position descriptions posted by end of April if at all possible.

Do we have funds for an Amazon server if Terry can’t get space? Angela indicate it would be possible if Missy allows it. Missy indicated all we need to do is a budget transfer.

Next steps:

* Have options researched, challenges identify what else needs to be done regarding testing etc.
* Terry will prototype Repox instance.