Sustainability of Community-owned Repository Software: A Call to Action

CNI Albuquerque, April 2017

Mike Giarlo - Stanford University
Nick Ruest - York University
Andrew Woods - DuraSpace
Melissa Anez - Islandora Foundation
The Common Mission

Long-term preservation of and access to culturally significant resources
The Common Approaches

Open source software
- Control
- Transparency
- Collaboration
- Flexibility / Extensibility
- Long term ownership (i.e. open licensing)

Standards
- Collective mindshare (shoulders of giants)
- Staying in line with broader initiatives (tooling)
- Interoperability
- Longevity
Environmental Scan

We are surrounded by discontinuities that are potential threats to our mission
The Challenge

How do we collectively sustain the work we do and resist discontinuation?
Sustainability: Scope

- Content
- Software
- Community

See also:
- Strategies for Sustaining Digital Libraries
  https://educopia.org/publications/ssdl
- Sustainable Economics for a Digital Planet
Content Sustainability

See:

- Relay-supporting Archives: Requirements and Progress

- Physical layer
- Logical layer
- Administrative layer
Software Sustainability

- Licensing;
  - Using open licenses and signing CLAs
- Documentation
  - User
  - Developer
  - Deployment
- Engagement
  - Stakeholder testing
  - Interest groups
- Institutional support
Community Sustainability

- Governance
- Collaboration
  - Code of Conduct
  - Workflow policies and procedures (communication!)
- Finances
- Hiring
  - Bringing new people on
  - Moving people around
Trustworthy Repository Software

- Risk Mitigation
- Sustainability
- Transparency
- Standards
- Durability
Case Study: Fedora
Intro to Fedora

Led by the Fedora Leadership Group and under the stewardship of the DuraSpace not-for-profit organization...

Fedora is a robust, modular, open source repository system for the management and dissemination of digital content.

It is especially suited for digital libraries and archives, both for access and preservation.

http://fedorarepository.org/
Fedora: Content Sustainability

✔ Managing external content

✔ Standard metadata ontologies

½ Migrating data models (XML to RDF)

½ Standardized import / export
Fedora: Software Sustainability

- Robustness of testing (CI, unit, integration, release)
- Apache2 Licensing & CLAs
- Sprints for feature development and documentation
- \( \frac{1}{2} \) Sprints for maintenance
- \( \frac{1}{2} \) API Specification
- \( \frac{1}{2} \) Limited Java developers
- ❌ Weekend developers
Fedora: Community Sustainability

✓ Governance structure
✓ Open processes (meetings, decisions)
✓ Full-time staff
✓ Resilient to significant evolution (F3 -> F4)
½ Onboarding leadership team
½ Accessible documentation
✗ Full-time developers
Case Study: Hydra
Intro to Hydra

Founded in 2008, Hydra is a community working together on repository solutions based on a common technical framework and open-source software.

Hydra is sustained by partner institutions, currently numbering thirty-five, and is used by several dozen cultural heritage organizations.

https://projecthydra.org/
Hydra: Software Sustainability

✔ Creation of Hyrax

✔ Portland Common Data Model (PCDM)

✗ Churn of past two years

✔ Standardized license (Apache 2.0)

✔ Contributor License Agreements

½ Robustness of testing (- release)

½ Community work cycle model
Hydra: Community Work Cycles

- Balance between junior and senior developers
- Allocation of sufficient time (multi-week at >= 50%)
- Contiguous vs. scattershot contributions
- Team size
- Right roles at the right time
- Onboarding and preparation
Hydra: Community Sustainability

- Governance structure (Partner model)
- Open processes (meetings, working/interest groups)
- Code of Conduct
- Abundance of local code
- Large and growing pool of talented, passionate people
- Roadmap is organic rather than directed
- Resilient to significant evolution (F3 -> F4 / PCDM)
Case Study: Islandora
Intro to Islandora

Islandora is an open-source software framework designed to help institutions and organizations and their audiences collaboratively manage, and discover digital assets using a best-practices framework. Islandora was originally developed by the University of Prince Edward Island's Robertson Library, but is now implemented and contributed to by an ever-growing international community.

Islandora.ca
Islandora: Software Sustainability

½ Sprints for feature development and documentation
½ Sprints for maintenance
❌ Weekend developers
✓ GPLv2/3, MIT Licensing
✓ CLAs & LSAP
✓ Robustness of testing (CI, unit, integration, release)
½ Limited developers
Islandora: Software Sustainability

✓ CONTRIBUTING.md
✓ Committers Workflow
✓ Pull Request Templates
½ Vendors
✓ Committer Policy
½ Engagement
Islandora: Community Sustainability

- Governance structure
- Open processes (meetings, decisions)
- Onboarding leadership team & contributors
- Accessible documentation (½)
- Full-time staff
- Full-time developers (×)
- Resilient to significant evolution (1.x -> CLAW) (½)
Islandora: Community Sustainability

☑ Community releases (1.x)
☑ Code of Conduct
☒ Burnout
☑ Interest Groups
☑ Islandora Awesome
☑ Conferences and Camps
What do you see?

Where can we improve sustainability?
Thank you!