CAUL Research Repository Community Days

29 October, 2013

Jonathan Markow
Chief Strategy Officer
Who is DuraSpace?
Not for profit 501c3 formed in 2009

Bringing together of the DSpace Foundation and Fedora Commons Organization
All employees of DuraSpace
Board of Directors
Provides fiscal oversight and advises on strategic direction, do not provide oversight on projects

- James Hilton, U of Michigan
- Paul Courant, U of Michigan
- Sayeed Choudury, John Hopkins University
- Chuck Henry, CLIR
- Tom Cramer, Stanford University
- Laura Woods, Tufts University
- Tyler Walters, Virginia Tech
- Anne Jarvis, Cambridge University
- Heather Joseph, SPARC
- Dan Cohen, DPLA
We believe that preserving the world’s intellectual, cultural and scientific heritage is vital.

WE provide leadership and innovation in developing durable, open technologies. Our COMMUNITIES share solutions in more than 90 countries. Our SPONSORS are committed to preserving the scholarly record.

---

**Our Open Source Projects**

- **DSpace**: DSpace is a turnkey institutional repository application.
  - Site ➤ DSpace Users ➤ Hosted DSpace via DSpaceDirect ➤

- **Fedora**: Fedora is a framework for building digital repositories.
  - Site ➤ Fedora Users ➤ Fedora Futures ➤

- **VIVO**: VIVO is a semantic web platform enabling research discovery.
  - Site ➤ Incubated Project ➤ About VIVO ➤

---

**Our Services**

- **DuraCloud**: DuraCloud is a managed service for archiving content in the cloud.
  - Site ➤ Glacier ➤ Try DuraCloud Now ➤

- **DSpaceDirect**: DSpaceDirect is a hosted turnkey repository solution.
  - Site ➤ I’m Interested ➤
Our mission

DuraSpace is committed to providing leadership and innovation for open technologies that promote durable, persistent access to digital data.

We collaborate with academic, scientific, cultural, and technology communities by supporting projects and creating services to help ensure that current and future generations have access to our collective digital heritage.
Initial goals

• Bring two organizations together to work collaboratively vs. competitively
• Develop a business model that would sustain the OS projects and organization
In 2009 purely grant funded

Our strategy was to diversity our funding by launching a community sponsorship program to fund the OS projects.
Diversification strategy

2010 ($1.9 MM)
- Grants: 84%
- Sponsorship: 14%
- Services: 2%

2012 ($1.36 MM)
- Grants: 41%
- Sponsorship: 38%
- Services: 21%

2013 ($1.67 MM)
- Grants: 25%
- Sponsorship: 57%
- Services: 18%
Adoption of platforms over time

175 new instances in 2012
1,700+ instances to date, but the rate is slowing
Number of New Instances

- 2009: 250
- 2010: 300
- 2011: 200
- 2012: 350
- 2013: 200
Our community is global

Continued adoption around the world, particularly developing Countries
Current Status

Growth in Sponsorship

- 2010: $299,000
- 2011: $358,000 (20% growth)
- 2012: $417,750 (17% growth)
- 2013: $658,870 (58% growth)
Sponsorship – By Geography
(excluding VIVO)

- USA: 76%
- Canada: 6%
- UK: 9%
- Western Europe: 3%
- Aust/NZ: 6%
This year

• Level of funding almost matches current level of resources for the project
• Fully supported by the community and not subsidized by grants or DuraSpace reserves
Launch of services
Motivations

• Increase our value proposition within the community by adding services
• Diversify our funding portfolio
• Easier access to grant funding
Confusion in the community
...some myths

- DuraSpace has teams of developers working on DSpace and Fedora
- DuraSpace is funded by the Moore foundation and others
- DuraSpace is DSpace on Fedora
- Sponsorship dollars are funding DuraCloud
- Bringing on more projects reduces effort on existing projects
Increase Transparency

• Establish two business units- Projects and Services
• Establish governance for the projects
• Allow funds to be directed to specific projects by sponsors
• Set budgets for each project and service and share with the community based on funding raised
Overall Objectives

• Continue to build a sustainable business model and revenue stream for all projects and services
• Provide vital services to our projects
• Bring solutions to our community for managing and preserving digital content
How is DuraSpace organized?

- **Projects**
  - DSpace
  - Fedora
  - Vivo

- **Services**
  - DuraCloud
  - Education
  - Dspacedirect

Funded through sponsorship, RSP program

Funded through service revenue and grants
Our Open Source Projects

Fedora Commons

VIVO: connect • share • discover
Who Owns the Projects?

DuraSpace provides leadership, support and a home for the OS projects with funding from sponsorship dollars and developer contributions.

The communities are the owners of the OS projects.
What do Open Source Projects Need to Thrive?

- Community contributions and engagement
- Great developers
- Effective governance
- Community support, partnerships, and growth
- Strategic direction and product roadmap
- Sustainability planning
How does DuraSpace support the projects?

• Executive leadership
• Technical leadership
• Community outreach
• Communications/promotion
• Business administration
How does the community support the projects?

- Set strategic direction
- Participate in governance
- Contribute to decision-making
- Contribute financial support
- Develop the software (in-kind developer resources)
- Help get things done
- Be an advocate
Project Objectives

- Have effective governance in place for each project
- Allow funds to be directed to specific projects by sponsors
- Set budgets for each project and service and share with the community based on funding raised
Project Objectives (cont.)

• Provide highest level of project support in line with the funding raised

• Identify other candidate projects that are complimentary, serve our mission and have the potential to be self sustaining

• Form strategic partnerships (DPN, SHARE, APTrust, OpenAIRE)
# Financials

<table>
<thead>
<tr>
<th>Item</th>
<th>CY 2013</th>
<th>CY2013 budget</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revenue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grants</td>
<td>414,000</td>
<td>414,000</td>
</tr>
<tr>
<td>Services</td>
<td>266,000</td>
<td>400,000</td>
</tr>
<tr>
<td>Sponsorship</td>
<td>1,089,685</td>
<td>860,000</td>
</tr>
<tr>
<td><strong>Total revenue</strong></td>
<td><strong>1,769,685</strong></td>
<td><strong>1,674,000</strong></td>
</tr>
<tr>
<td>Direct costs</td>
<td>1,555,469</td>
<td>1,435,892</td>
</tr>
<tr>
<td>Indirect costs</td>
<td>83,049</td>
<td>103,000</td>
</tr>
<tr>
<td><strong>Total expenses</strong></td>
<td><strong>1,638,518</strong></td>
<td><strong>1,538,892</strong></td>
</tr>
<tr>
<td>Net income</td>
<td>131,167</td>
<td>135,108</td>
</tr>
</tbody>
</table>
Sponsorship – As Directed
(excluding VIVO)

- $363,833
- $155,733
- $139,304

Legend:
- DSpace
- Fedora
- General
Sponsorship—By Software
(excluding VIVO)

- DSpace: $230,870
- Fedora: $318,500
- Both: $88,500
- Neither: $21,000

Total: $659,940
DSpace RoadMap & Vision

Tim Donohue
DSpace Tech Lead
tdonohue@duraspace.org

Australia Oct 2013
Overview

• Intros
• 3.x Releases (3.1 in Jan 2013)
• 4.0 is coming (Late 2013)
• DSpace 3 to 5 Year Vision
• Next Steps
Meet the Committers!

- Andrea Bollini
- Ben Bosman
- Peter Dietz
- Mark Diggory
- Tim Donohue
- Sands Fish
- Keith Gilbertson
- Richard Jones
- Claudia Jürgen
- Bram Luyten (2013)
- Ivan Masár (2012)

- João Melo (2012)
- Hardy Pottinger (2011)
- Richard Rodgers
- Andrea Schweer (2011)
- Kim Shepherd
- Kostas Stamatis (2013)
- Keiji Suzuki (2013)
- Robin Taylor
- Graham Triggs
- Jeff Trimble
- Kevin Van de Velde (2011)
- Mark Wood

Key: 5+ Years, 3-5 Years
Meet DCAT!

- Claire Bundy
- Bharat M. Chaudhari
- James Evans
- Elena Feinstein
- Augustine Gitonga
- Yan Han
- Leonie Hayes
- Valorie Hollister
- Dibyendra Hyoju
- Iryna Kuchma
- Sue Kunda
- Amy Lana
- Bram Luyten
- Sarah Molloy
- Jim Ottaviani
- Sarah Potvin
- Sarah Shreeves
- Elin Stangeland
- Ciarán Walsh
- Maureen Walsh

http://tinyurl.com/dspace-dcat
3.x New Features

- Rewritten & improved OAI-PMH
- Advanced Embargo
- Item Versioning (beta)
- Mobile interface (beta)
- Type-based submissions
- Many major improvements: Statistics, Authentication, Discovery, etc

http://tinyurl.com/dspace-3-0
3.x Releases

• First release managed by a Release Team!
• 3.0 Release Team:
  – Sands Fish (MIT) [Team Leader]
  – Ivan Masár (aka helix84)
  – Hardy Pottinger (U of Missouri)
  – Robin Taylor (U of Edinburgh)

Want more 3.x?
Hardy Pottinger’s talk @ 4:15pm
Coming Soon: 4.0

• Late 2013 (Nov or Dec)
  – Likely DSpace 4.0 Features:
    - REST API (read-only, beta release)
    - SWORDv2 enhancements & bug fixes
    - SHERPA/RoMEO deposit integration ([http://www.sherpa.ac.uk/romeo/](http://www.sherpa.ac.uk/romeo/))
    - Faceted/Filtered Search by default (based on Discovery)
    - Basic DOI (Digital Object Identifier) support

[http://tinyurl.com/dspace-4-0](http://tinyurl.com/dspace-4-0)
Coming Soon: 4.0

– Likely DSpace 4.0 Features (cont.)
  - Ability to use an external Handle Server with DSpace
  - Numerous other enhancements & bug fixes

4.0 Release Team
– Mark Wood, Hardy Pottinger, Andrea Bollini

http://tinyurl.com/dspace-4-0
Edit Collection 123456789/2

Collection's Metadata

Name: This is my first collection
Short Description: With a very useful short description
Introductory text (HTML): Lorem ipsum dolor sit amet, consectetur adipiscing elit. Vivamus mollis, erat
sit amet puviner eleifend. arcu nisl lauculis lorem, ac pharetra sapien erat nec
sapen. Nam ornare luctus augue a sodales. Suspenisse at molestias lorem, a
congue purus. Maecenas nec nisl lacus, euismod turpis condimentum,
estibulum nisl. Phasellus atque mauris mi. Etiam potentencese ipsum nunc,
id venenatis urna tempora. Vestibulum at pharetra velit, in hendrerit enim

Copyright text (plain text): Copyright My Collection
Sidebar text (HTML): News from the collection managers

Submitters: [Edit] [Delete]
Accept/Reject Step: [Edit] [Delete]
Accept/Reject/Edit Metadata Step: [Edit] [Delete]
Collection Administrators: [Create]
Collection's settings
Item template: [Create]
Collection's Authorizations: [Edit]
Collection's Curations: [Curate]
DSpace will:

1. Focus on IR fundamentals, modern use cases
2. Be lean & flexible
3. Include “core IR” functionality which can be extended
4. Be designed to integrate well
5. Support low-cost, hosted solutions
1. Focus on IR

“Institutional Repository” at heart

• Core functionality = current & anticipated IR use cases
  – Not same use cases as 10 yrs ago

• Anything Non-Core = implement via (third-party) plugins
2. Lean & Flexible

Do NOT attempt to do or be everything

• Instead: lean, focused system

• But, support third-party plugins
3. Core IR Use Cases

What is “out-of-the-box”

vs

What may be a third-party plugin

- We need to determine what these core use cases are, together

- Retain “out-of-the-box” IR goal
4. Designed to Integrate

Standalone system = thing of the past

• Provide integration points for others to build against (APIs, etc)

• Build our own integrations with external systems / services
5. Support low-cost/hosted

Web-based Configuration & Setup

• Easy to setup / install whether on your local server or hosted

• Most options / configs should be customizable via a web UI
Next Steps

- Use Cases / feedback via email / wiki
- Collaborate with Tech Team on a Technical RoadMap / Plan
- Work towards establishing a community project
  - Governance (DSpace Steering Group)
  - Locate developers / funding

http://tinyurl.com/dspace-2013-vision
Current Strengths

- Architecture is Flexible & Extensible
- Support for Durability
- A decade of maturity & proven use
- Substantial community of adopters and contributors
- One foot in the linked data world
- Flexible content modeling
Opportunities

- Performance
- Fault tolerance and scalability
- Handle increased complexity
- Update aging code base
- Increase relatively small, fixed pool of active developers
Threats

• Declining interest from developers
• Steep learning curve for newbies
• Fedora losing mindshare as strategic choice for new application needs
• Codebase viewed as difficult to change
• New use cases not easily accommodated
Turning Point

- The Fork in the Road
- OR12
- Diverse Stakeholders form Project Steering Committee
- The message
- The “ask”
Objectives

- Preserve the strengths of the architecture and community
- Address the needs for robust and full-featured repository services (that we now understand very well)
- Provide a platform in the repository ecosystem for the next 5-10 years
Use Cases

Steering Group, Tech Group, and Advisors identified some 30 use cases

Large overlap - four major topics:
- manage research data
- improve administrative features
- Increase performance
- interact with linked open data/semantic web
Major Requirements

• Performance
• Scalability
• Flexible storage policies
• Support for Research Data Management
• Thinner, lighter kernel
• Hydra and Islandora adaptors
• Testing goals and metrics
• RDF support
• Clear upgrade path
More Major Requirements

- support for dynamic metadata
- globally unique and reliable identifiers
- improved audit trail/capturing events
- straightforward and fun to use API
The Project Team

- Project Steering Group
- Project Advisory Group
- Technical Team
- Tech Lead – F/T
- Product Manager F/T
- Working groups
  - Fundraising
  - Hiring
  - Technical
  - Governance
Fedora Futures IS Fedora 4

• Unify the current project community with the historical committer base
• Bring Fedora sponsors closer into the decision-making process
• Create on-going model for governance based on sponsors, steerers, project team, advisors, working groups
Case Study

- University of New South Wales
  - Playing critical role on project
  - Developing major pieces of Fedora 4
  - Testing sprint output
  - Contributing to project as a Silver level Sponsor
  - Engaged with entire project team
Fedora 4 wants you!

And you!

And you!

And you!
How you can participate

• Become a sponsor
  www.duraspace.org/sponsor

• Contribute developers to the project
  – www.dspace.org/contributors
  – wiki.duraspace.org/display/FF/Fedora+Futures+Home