Small is beautiful: the value proposition for libraries as publishers using open source systems

Dr Edmund Balnaves, Prosentient Systems, ebalnaves@prosentient.com.au
Hilton Gibson, Stellenbosch University, hilton.gibson@gmail.com
Wouter Klapwijk, Stellenbosch University, Wklap@sun.ac.za
“Don't ask for guarantees. And don't look to be saved in any one thing, person, machine, or library. Do your own bit of saving, and if you drown, at least die knowing you were headed for shore.”

Ray Bradbury. Farenheit 451
The motivation

Ensuring the visibility of the publishing, report and creative output of your institution (why should Google get all the credit?)

The conundrum of media multiple media outlets & the publishing explosion

The digital library as KM resource and information dissemination
Not a mega project anymore

- There are now many established, low cost, digital library systems allowing establishment of out of the box open source digital library solutions

- **Low cost & low risk:** out of the box open source solutions are available, self installable or hosted through providers
Different approaches to content acquisition

- **Simple workflow** - online submission via API, structured word document or web submission with optional accept/reject curation workflow

- **Automated** harvesting of web content from multiple sources with ranking and filtering;

- **System integration** with other open source systems (e.g., DSpace and Koha)
Closing the portcullis

- The tensions between open and closed information sources
- Eg Fee-based gateways to news and online resources replacing previously free systems
- Fee gateways to web services (e.g. Google maps and Google search)
- The example of LogMeIn
The open source benefit

- Source code provided == control
  (you can solve problems directly yourself)
- Flexibility in deployment – unrestricted by licencing control
  - External hosting
  - Internal hosting
  - Multiple instances
  - Unlimited users
- Incremental improvements over time through community involvement
- Code snippets and examples
- Tends to be “open” in other ways:
  - Many services layers
  - Leverages other open tools
The value proposition

Institutions are:
• Losing control of their own publications
• Losing their own publications
• Losing the credit for their own publications

However:
• Commitment to an open access repository is a multi-lifetime commitment with some costs

Hence:
• The value proposition – quantifying the cost/benefit
Ballarat Health Services
Putting your health first

Digital Library
Welcome to the Ballarat Health Services Digital Repository.
This is a service provided to you by the Ballarat Health Services Library.

Search
Enter some text in the box below to search the Repository.
• Ballarat Health Service - Small health service and hospital library, servicing a regional community in Victoria, Australia
• Initial setup cost: (Bare repository) - $US2200 - 08/2012, including training (1 day)
• Website:  http://bhsdlib.intersearch.com.au/bhsjspui/
• Collection size:  710 items
• Ongoing submission management: $US6669 / annum
• Hosting and software support:  $US1300 / annum
• Average views per annum:  1,222,400
• Cost per view:  6.5 cents($US)
About the Repository

ePublications@CSNSW is an open access repository of the research and scholarly output of Corrective Services New South Wales staff and students, and historical and archival material about the department. Staff and students of CSNSW are invited to contribute digital materials for long-term preservation and world-wide electronic accessibility.

Contact: Toni Kennedy
Toni.Kennedy@dcn.nsw.gov.au
Case study 2

- Number of items in the collection: 753 items
- This project involved setup of the repository and scanning of the existing library collection
- Initial setup cost: $US1465 – Jan 2005, incl training (1 day)
- Hosting and software support: $AUD1800 / annum
Case study 2 – collection development

- OCR capture - scanner purchase and installation - $AUD3664 (capital).
- Scanning done 2-3 documents per day by each the front desk staff as part of their daily work profile. No additional staff for the role, plus assistance by low-security inmates.
- Nominal cost (no additional staffing taken on - 3-4 items added to repository per month) $AUD1000 / annum
• Average views per annum: 145968
• Cost per view: .9 cents ($US)
Case study 3

- Stellenbosch University Library: Institutional Research Repository: [http://scholar.sun.ac.za](http://scholar.sun.ac.za)
- Self-hosting value proposition for an institutional research repository, hosted and published by an academic library.
Journal Hosting service

- Stellenbosch University Library: hosting service for Open Access journals using Open Source software
- **Open Journal System (OJS) – Public Knowledge Project (PKP)**
- As on 1 August 2015: hosts 20 Open Access journal titles
- Requirement: member of editorial board must be associated with Stellenbosch University
- Requirement: signed Memorandum of Understanding (MoU)
Journal Hosting service: technical

- Separate domain registered to host journals: http://abc.journals.ac.za
- Each journal is assigned its own server installation
- Server Operating System – Open Source linux (Ubuntu LTS)
- Operational backup of journal contents onto two different server platforms
- Long-term preservation of journal content with LOCKSS or Portico system underway
“Hosting” is defined as the installation of application software (i.e. DSpace) on a server and the systems administration thereof, it does not cover operational maintenance (e.g. metadata management) nor long-term digital preservation.

The Library and Information service is not responsible for the content in the repository.
## Repository Hosting operational cost

### Start-up cost (once-off)

<table>
<thead>
<tr>
<th>Description</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cloud server provisioning (Cloud Service provider or SU IT Data Centre)</td>
<td>R 500.00</td>
</tr>
<tr>
<td>2. Cloud server preparation (OS) – usually requires 2 hours</td>
<td>R 2 000.00 / hour</td>
</tr>
<tr>
<td>3. DSpace installation – usually requires 2 hours</td>
<td>R 2 000.00 / hour</td>
</tr>
<tr>
<td>4. Handle and SSL installation – usually requires 1 hour</td>
<td>R 2 000.00 / hour</td>
</tr>
<tr>
<td><strong>Total cost (typical scenario):</strong></td>
<td><strong>R 10 500.00</strong></td>
</tr>
</tbody>
</table>

### Training cost (initial training only)

<table>
<thead>
<tr>
<th>Description</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSpace configuration, basic UI customization, training, and handover</td>
<td>R 4 000.00</td>
</tr>
</tbody>
</table>

### Administration cost (monthly)

<table>
<thead>
<tr>
<th>Description</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Server hosting <em>(with Hetzner ISP first 50GB backup space is free)</em></td>
<td>R 900.00</td>
</tr>
<tr>
<td>Systems and application administration – set monthly rate capped at R 4 000</td>
<td>R 4 000.00</td>
</tr>
<tr>
<td></td>
<td><strong>R 4 900.00</strong></td>
</tr>
</tbody>
</table>

### Other cost (annual)

<table>
<thead>
<tr>
<th>Description</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain Name renewal (FQDN)</td>
<td>R 100.00</td>
</tr>
<tr>
<td>Handle and SSL renewal</td>
<td>R 4 000.00</td>
</tr>
<tr>
<td><strong>Sub-total:</strong></td>
<td><strong>R 4 100.00</strong></td>
</tr>
</tbody>
</table>

**Contingencies:** follow-up training, support

<table>
<thead>
<tr>
<th>Cost</th>
<th>per hour</th>
</tr>
</thead>
<tbody>
<tr>
<td>R 2 000.00</td>
<td></td>
</tr>
</tbody>
</table>
The cost of hardware per annum, is therefore:

- \((R250,000/4)\) [One production server amortised over 4 years] + \(((R100,000/4) \times 2)\) [Two backup servers amortised over 4 years] = R112,500 pa

The cost of personnel per annum, is therefore:

- R500,000 [1 x OSCD] + R350,000 [1 x OSCM] + (2 x R250,000) [2 x OSCL] + (2 x R250,000) [2 x OSCS] = R1,850,000 pa
The total cost per annum is therefore:

\[ \text{R112,500 [Hardware]} + \text{R1,850,000 [Personnel]} = \text{R1,962,500 pa} \]

Therefore cost per item downloaded on SUNScholar for 2014 is:

\[ \text{R1,962,500 [Total Cost] divided by 102000 [No of items downloaded]} = \text{R19.24 rounded out to the nearest cent.} \]
Installing your own system

How much programming skill do you need?

• 10 years minimum experience

• OR

A SENSE OF ADVENTURE
A system building journey

- Evaluate currently available tools and systems
- Open source evaluation methodologies
  - Code review
  - Database review
  - Community review
- Deciding on a package approach or a toolkit approach
- Mucking about
New paradigms

- Library as a publisher
- Open source, commercial, cloud, bespoke
- Librarian as a system integrator
Ways of harvesting and system integration

- RSS, an interface popular for news syndication,
- OAI/PMH - a protocol for bibliographic and record interchange between digital repositories
- JSON-based information sources.
- Schema.org and other tag embedding
- Endnote/zotero/Pubmed imports
- XPath with XML and PHP coding to isolate portions of a web page of interest
- Search API’s such as Google Custom search
- E-mail APIs (especially IMAP processing tools for PHP).
Libraries as innovators

- Powerful new toolsets

- In sourcing, outsourcing, cloud hosting and system integration

- Rich pickings on slim budgets

- Changing role: library as publisher
Further Information?

ejb@prosentient.com.au

Prosentient Systems Pty Ltd
72/330 Wattle St, Ultimo NSW 2007