

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

Authors:

Edmund Balnaves, Prosentient Systems, ebalnaves@prosentient.com.au

Hilton Gibson, Stellenbosch University, hilton.gibson@gmail.com

Wouter Klapwijk, Stellenbosch University, wklap@sun.ac.za

Abstract

Open source digital library systems have developed in parallel with open access to present a unique opportunity for libraries to directly represent the intellectual output of their organisations by owning their own delivery systems for the digital library content.

Open source digital library systems have reached a state of maturity that allows them to be implemented with minimal cost and technical overhead, while delivering the organisation consistent management and representation of their public documents and intellectual content. Stellenbosch university has published metrics on the value proposition for managing such systems.

Many organisations have published content such as reports, journal publications and other resources through their websites. Website makeovers and domain changes typically result in lost information over time, with the average longevity of such links not more than 5 years. This paper explores the value proposition for the library working with research and communications areas of the organisation to properly manage the publication of organisation output, including the support for open access publishing of academic output. The value proposition is presented qualitatively and quantitatively through examples from the university and special library sector.

Introduction

The last decade has seen the adoption of open access as a principle for publishing in many institutions, and in some cases mandated at national level (citation here). There are various methods by which open access can be achieved, including publication through open access journal or book publishers. This paper focusses on the value proposition relating to institutions that choose their own open access repository using open source systems.

Open Source as an enabler for Open Access

One of the great double acts that underlie modern libraries has been the rise of open source systems (OSS) and open access (OA). Adoption of OSS and OA has been extraordinary over

Transforming libraries with open digital technologies
IFLA Satellite conference.
Stellenbosch University, South Africa
August 13, 2015 – August 14, 2015

the last ten years, has progressed in parallel, and in similar timeframes. The synergies go deeper than this. First, open source can be an enabler for the adoption of open access in an institution. Equally a successful OA project can justify the ongoing improvement of the OSS implementation. Second, OSS can provide a level of certainty for an institution in their operation costs. The larger the community of adopters of open source the stronger the overall support. Third, the OSS can provide a level of security in that there is no proprietary lock-in and the code is visible (and therefore can be corrected). The functional depth of this security will be improved by the work of those adopting the open source model.

Finally, open source systems can provide a cost-viable model for implementation of open access in smaller institutions. A common confusion is that open source means “free”. While it may be lower cost, no information technology system operation is free. The ongoing nurturing of a system, software upgrades over time, support for customisations and enhancements, server administration, network costs are just a few of the base-line elements of managing an information system. Nevertheless, the amortisation of the software support across a wide installed base makes for an effective cost model for smaller institutions.

Libraries as Publishers – Managing Open Access through Digital Libraries

The physical print and serial collections, and the library as a “lending” collection, is being supplanted by the library as a mediator of access to effective electronic journal and print resources, and of the library as a manager of the organisation digital resources. Many institutions have a continuous publishing output, of reports, research and information. This publishing is often “outsourced” through the traditional academic model of publishing, or is published ephemerally through the institutional website and lost with site upgrades and redesigns. The “outsourcing” model of academic publishing relies on the direct relationship between the author of research and the publisher, typically leaving out the institution that has contributed the infrastructure, and services to complete this research. The published work is then “owned” by the publisher with very little residual rights to the author, and usually none to the institution. The Open Access movement represents an effort by authors and institutions to gain a better ownership or at least internal knowledge leverage over the research publishing of the institution and their other publications (eg annual reports, images, etc). Similarly, reports and internal publishing activity is often ephemerally delivered on websites whose longevity is poor. A website makeover will often not preserve older material and with such makeovers typically at least every 5 years, there is gradual loss of knowledge for the institution.

A digital library system presents an opportunity to gain better internal knowledge management of documents, and for better visibility of organisation publications. For example, Great Barrier Reef Marine Park Authority use DSpace as both a publication resource and for staff papers that have to be specially requested. They publish a range of documents, some with full open

access, some with limited access. The Department of Minerals and Resources in NT similarly publishes geological surveys through DSpace with a map-based interface.

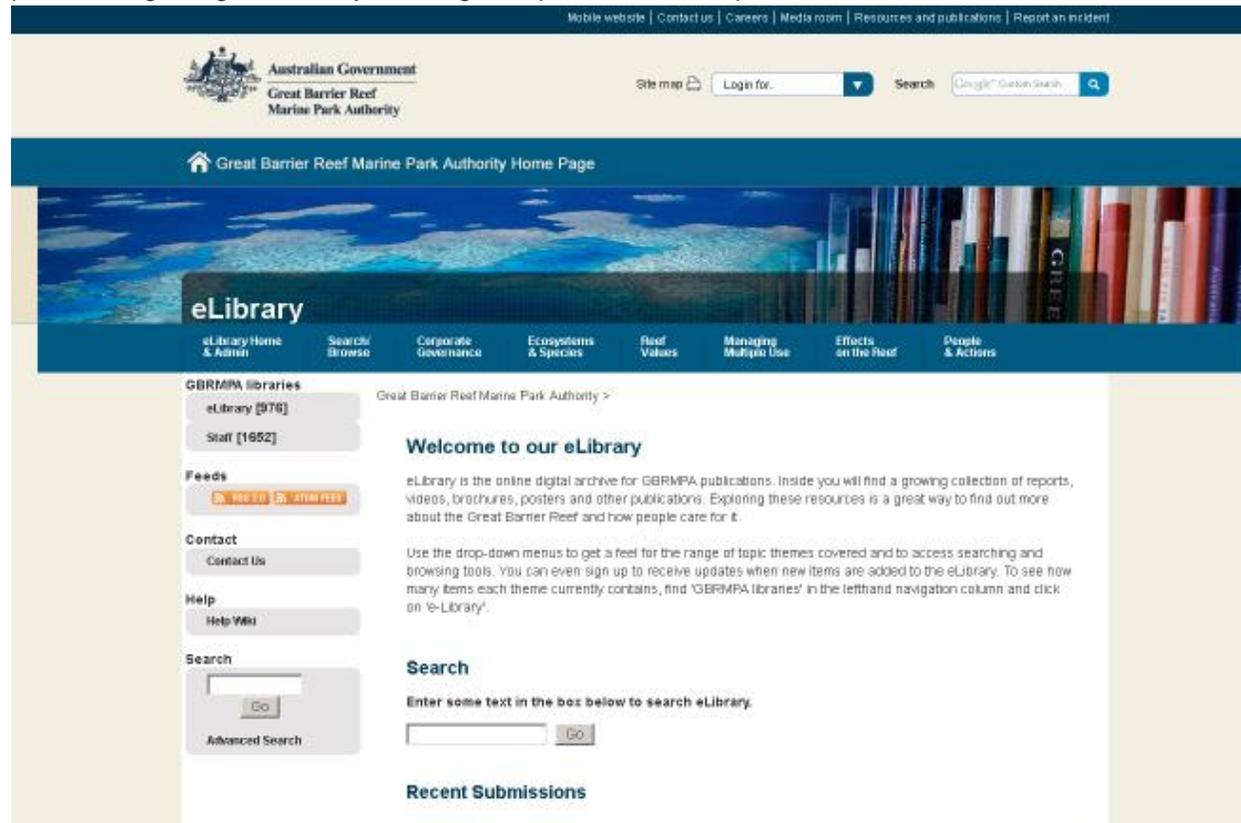


Illustration 1: DSpace as a collection of open and closed electron resources:
<http://elibrary.qbrmpa.gov.au/jspui/>

Research teams are also interested in building repositories that allow them to pool interesting articles over time and share this research. The “Living Systems” digital library, for instance, uses DSpace as a unified repository for their publications and to harvest from the Zotero reference databases for joint research articles. (Illustration 2)

Value Proposition

An institution that hosts its own digital library using open source faces a range of cost factors including:

- The hardware / hosting environment cost
- The software cost for maintaining/upgrading the software (no software is static, it must be upgraded regularly for a range of compliance and functionality issues).
- The time and effort to submit items to the digital library

Transforming libraries with open digital technologies
IFLA Satellite conference.
Stellenbosch University, South Africa
August 13, 2015 – August 14, 2015

However, if the institution is not hosting their own records, then there is a cost to otherwise obtain these documents, as well as a potential opportunity cost for the missed visibility of these documents where other institutions do not make these documents visible over time.

The value proposition in this document illustrates the costs against the access benefits for the institution hosting its own repository.

Non-quantitative benefits can also include:

- greater visibility of the research output of the institution (by distributed harvesting of content through to national and commercial aggregators).
- better longitudinal visibility of the research work of the institution
- making available reports and other content not otherwise available
- keeping the historical record of the institution

Case Study 1:

Ballarat Health Service - Small health service and hospital library, servicing a regional community in Victoria, Australia

Initial setup cost: (Bare repository) - \$AUD3000 - 08/2012, including training (1 day)

Website: <http://bhsdlib.intersearch.com.au/bhsjspui/>

Collection size: 710 items

Management of the system (submission review, repository management, user management):
\$AUD9100 / annum

Hosting and software support: \$AUD1800 / annum

Average views per annum: 1,222,400

Cost per view: 8.9 cents (\$AUD)

Case Study 2:

Corrective Services Academy, NSW Australia - NSW government department.

Transforming libraries with open digital technologies
IFLA Satellite conference.
Stellenbosch University, South Africa
August 13, 2015 – August 14, 2015

Number of items in the collection: 753 items

Website: <http://csa.intersearch.com.au/csajspui/>

This project involved setup of the repository and scanning of the existing library collection

Initial setup cost: (Bare repository) - \$AUD2000 - Jan2005, including training (1 day)

OCR capture - scanner purchase and installation - \$AUD5000 (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

Hosting and software support: \$AUD1800 / annum

Average views per annum: 145968

Cost per view: 1.2 cents (\$AUD)

Case Study 3:

Stellenbosch University Library : Institutional Research Repository : <http://scholar.sun.ac.za>

This case study attempts to introduce the self-hosting value proposition for an institutional research repository, hosted and published by an academic library.

The values being proposed:

- It is eminently more economical and sustainable in the long term for academic libraries to build capacity and become online open access scholarly communication publishers themselves without resorting to outside commercial services ever again and risking another very expensive "serials bundle crisis" that was and is precipitated by very greedy commercial academic journal publishers.
- In addition, libraries are uniquely positioned to be in the "forever business" of maintaining the digital scholarly record because most libraries are already staffed by information professionals and therefore the library only needs a few simple hires to transform into a digital research publisher and long term digital curator.

Transforming libraries with open digital technologies
IFLA Satellite conference.
Stellenbosch University, South Africa
August 13, 2015 – August 14, 2015

Assumptions

The following is assumed:

- The academic library is the responsible publisher and therefore employs and trains the required personnel.
- Normal monthly overheads, such as building rental and maintenance, utilities and insurance etc... are deemed to be absorbed by the academic institution as part of normal operations, with or without this value proposition and therefore not a factor. They are a BIG factor for commercial publishing entities and this is where an academic library benefits the most, it and the institution are already absorbing institutional overhead costs!
- The proposition involves normal running costs for the library and does not cover costs for ad-hoc or continuing digitisation projects. Digitisation is assumed to be costed separately.
- The open system focus is on the ability to provide "green" open access publishing infrastructure and support, the immediate deposit, immediate access, model (ID/IA) as per Steven Harnard's definition.
- The open system is based strictly on open standards and open source software, so that there are no intellectual property or service contract costs, payable to third parties, at all.
- The central IT department provides internet infrastructure and data centre support, and that a service agreement exists for such services, between the library and the central IT department.

To be able to make comparisons with other library publishing systems and to provide a single simple metric, cost per item downloaded (CPD) will be calculated.

- Number of items downloaded in 2013 as per our Piwik stats = 102000
- Production Server Hardware Cost = R250,000 to be amortised over 4 years, the length of a typical hardware warranty.
- Backup Server Hardware Cost = R100,000 to be amortised over 4 years, the length of a typical hardware warranty.
- Salary - OSCD - Open Scholarly Communications Director (New Hire) = R500,000
- Salary - OSCM - Open Scholarly Communications Manager (New Hire) = R350,000
- Salary - OSCL - Open Scholarly Communications Librarian (New Hire) = R250,000
- Salary - OSCS - Open Scholarly Communications Systems (New Hire) = R250,000

The cost of hardware per annum, is therefore:

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

Transforming libraries with open digital technologies
IFLA Satellite conference.
Stellenbosch University, South Africa
August 13, 2015 – August 14, 2015

The cost of personnel per annum, is therefore:

$$R500,000 [1 \times OSCD] + R350,000 [1 \times OSCM] + (2 \times R250,000) [2 \times OSCL] + (2 \times R250,000) [2 \times OSCS] = R1,850,000 \text{ pa}$$

The total cost per annum is therefore:

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

Therefore cost per item downloaded on SUNScholar for 2014 is:

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

Observations

- At an average 2014 R/US\$ exchange rate of R10.00 to the dollar, the CPD = \$1.92 approx.
- The more that items are downloaded, the cheaper it becomes per item!
- The more full text items available for download, the better! This makes better economical use of the resources.
- If existing library staff, who are digitally competent and are passionate about open access, could be trained and reassigned, then the CPD drops radically - since no new hires are required.
- Assuming that each academic library in South Africa pays R40 million per year to commercial publishers on average, then a fully open scholarly communication publishing system only costs 5% of the annual e-resources budget!

Conclusion

Self hosting of digital libraries by non-government, government and research institutions is highly cost economical, whether hosted internally for opting for externally hosted solutions. Out of the box software is available, and is open source. The software is well established, functional and stable. The open source licensing gives flexibility in deployment over time, whether using internally hosted servers or servers provided by hosting providers.

The cost per item downloaded, in relation to the large commercial publishers, is extremely economical. We do not have any direct comparison regarding the cost per download for commercially published items.

The benefits of the institution hosting its own repository include:

- Better ownership and self-branding of the content and therefore better visibility of the institution and its longitudinal history of publishing and reports
-
- Physical custody of all academic research and teaching digital assets.
- Significantly more financial budget control because the institution is not at the mercy of commercial publishers
- Complete control of the publishing workflow which allows the institution to customise the system to fit the purpose for which it was created, rather than trying to customise a one-size-fits-all system.

However, the institution must itself commit to the long-term management of the repository, including:

- Training of librarians in new scholarly publishing competencies.
- Training of librarians to advocate for open access research publishing.
- Technical support and hosting over time (financial commitment)
- Hiring persons in academic libraries that have expert internet publishing technology skills.
- Adopting internet publishing technology as standard and supported practice for academic libraries.
- Participating in the open source community in various ways

References

Balnaves, E. (2008). Open source library management systems: a multidimensional evaluation. *Australian Academic & Research Libraries, AARL*, 39(1), 1-13.

Balnaves, E. (2013). From OPAC to Archive: integrated discovery and digital libraries with open source. *IFLA Congress, Singapore*, [\[Online version\]](#)

Balnaves, E. (2012). *Information and Communication Technologies in Parliamentary Libraries*. [\[Online version\]](#)

Transforming libraries with open digital technologies
IFLA Satellite conference.
Stellenbosch University, South Africa
August 13, 2015 – August 14, 2015

http://wiki.lib.sun.ac.za/index.php/SUNScholar/Self-Hosting_Value_Proposition
http://wiki.lib.sun.ac.za/index.php/SUNScholar/Capacity_Building
http://wiki.lib.sun.ac.za/index.php/SUNScholar/Install_Ubuntu/S01#Hardware
http://wiki.lib.sun.ac.za/index.php/SUNScholar/Disaster_Recovery

Transforming libraries with open digital technologies
IFLA Satellite conference.
Stellenbosch University, South Africa
August 13, 2015 – August 14, 2015