Case Study on Sustainable Economic Models: Portico

## Amy Kirchhoff and Kate Wittenberg

Introduction

Portico[1] is a not-for-profit digital preservation service and is among the largest community-supported digital archives in the world. In 2010, Portico became the first preservation service to be independently audited by the Center for Research Libraries (CRL) and certified as a trusted, reliable digital preservation solution that serves the needs of the library community. Working with libraries, publishers, and funders, Portico preserves e-journals, e-books, and other digital scholarly content to ensure researchers and students will have access to these resources in the future. Portico is a service of ITHAKA,[2] a not-for-profit organisation dedicated to helping the academic community use digital technologies to preserve the scholarly record and to advance research and teaching in sustainable ways.

As of August 2011, Portico was preserving more than 17 million journal articles, nearly 13,000 books, and nearly 1.5 million items from digitised historical collections (d-collections, for example digitised newspapers of the 18th century) with more than 12,000 journals the archive.   Portico's approach to preserving this content addresses the key goals of digital preservation: usability—the intellectual content of the item must remain usable via the delivery mechanism of current technology; authenticity—the provenance of the content must be proven and the content an authentic replica of the original;

discoverability—the content must have logical bibliographic metadata so that the content can be found by end users through time; and accessibility—the content must be available for use to the appropriate community. Portico meets the rigour of these goals through a migration-based strategy; Portico will migrate or transform preserved content from one file format to another as technology changes. Portico supplements and supports this migration policy by preserving the original source files along with the migrated versions. In addition, Portico has developed a technology and application independent archive (for example, the Portico archive can be exported into a standard file system with all the information necessary to understand the contents of the archive in organised files).

History of Portico

Since the 1990s, digital publications have become an increasingly important component of scholarly communications, as evidenced by the increasing percentage of the university membership of the Association of Research Libraries (ARL) expenditures on digital resources:
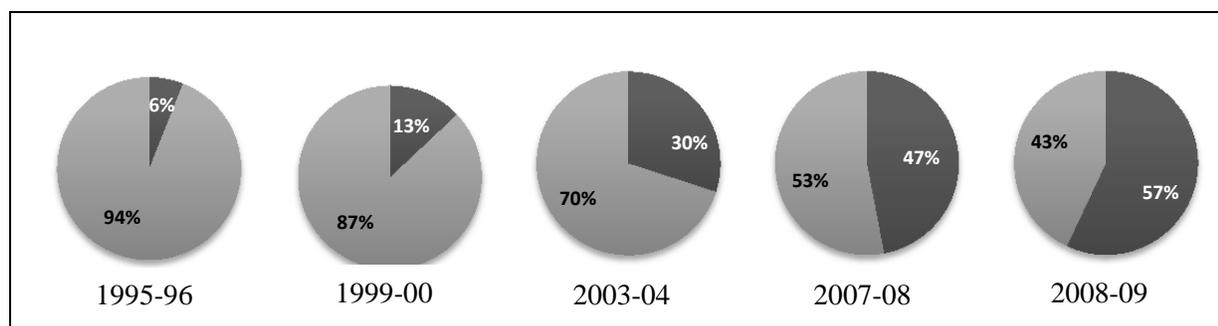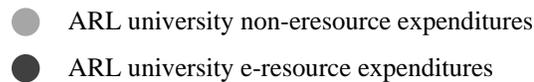
Figure 1: ARL University Library Materials Expenditures (Kyrillidou and Morris, 2011: 20-21)

From the beginning of digital publication, the academic community expressed the need for reliable long-term preservation of these new publications (Garrett and Waters, 1996) and specifically stressed the need for preservation of electronic journals (Waters, 2005). Without a reliable system for preservation, neither publishers nor libraries would be able to transform into modern organisations with full reliance on the digital format—a format often preferred by students and scholars for its ease of access and use. Without effective preservation, both publishers and libraries were locked into a dual-format world with all the expense that duplication imposed.

In 2002, JSTOR[3] established the Electronic Archiving Initiative ('E-Archive') with a grant from The Andrew W. Mellon Foundation. The purpose of the initiative was to further Mellon's seminal E-Journal Archiving Program by developing the organisational and technical infrastructure needed to insure the long-term preservation of and access to electronic scholarly resources. The E-Archive team worked with publishers and libraries to develop an approach that balanced the needs of each group while creating a sustainable digital preservation model. In 2005, Portico was officially launched as a result of these efforts with support from JSTOR, Ithaka, The Andrew W. Mellon Foundation, the Library of Congress, and numerous library and publisher participants. Portico's initial focus was on e-journal preservation.
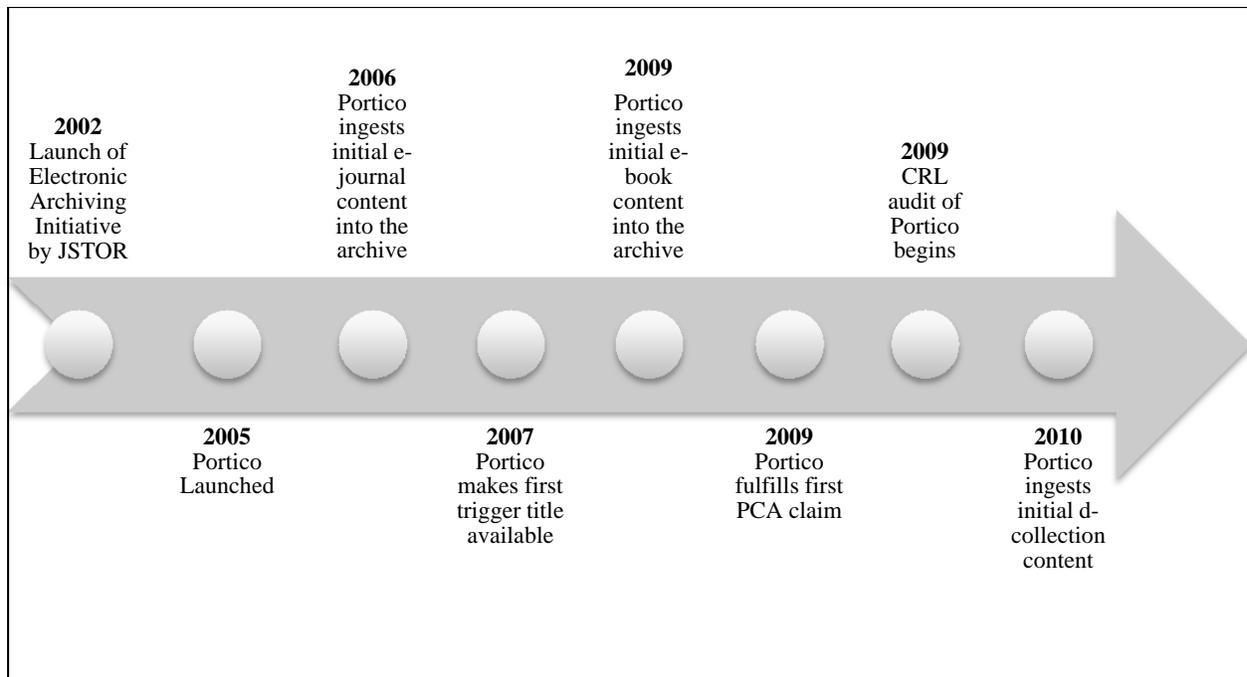
Figure 2: Portico History

As of 2011, Portico offers three separate preservation services to the academic community: e-journal preservation, e-book preservation, and d-collection preservation. The economic models supporting each service differ. The Portico e-journal preservation service and the Portico e-book preservation service are financially supported by publishers and libraries. The Portico d-collection preservation service is financially supported by publishers only.

Economic model development—electronic journal preservation

In early 2002, when JSTOR launched E-Archive, existing market mechanisms had failed to produce a preservation approach for e-journals that garnered the support of a wide range of publishers and libraries, and this suggested a not-for-profit approach was needed. Any reliable preservation mechanism would need wide uptake to secure the

economies of scale and diversity of revenue base required to ensure long-term sustainability of the preservation service. Portico took a system-wide view that included the perspectives of large and small, commercial and university presses, and a variety of large and small, US and international libraries.

In 2002, digital preservation was not well understood, but it was recognised as an important issue to be addressed. With print publications, libraries traditionally held the responsibility for archiving the scholarly record. With the advent of digital scholarly communications, publishers took on preservation responsibilities because the digital files remain in their custody rather than on the library's shelves. In order to develop the e-journal preservation service, Portico staff worked with publishers and libraries to identify and understand the technological and business issues. These issues informed the development of the e-journal preservation business model as Portico strove to make certain it served the needs of the entire academic community.

The infrastructure required for digital preservation is extensive and carries significant costs. E-journals are particularly expensive to preserve because of the wide variety of formats used in the content and variety of packaging schemes implemented—it is simply expensive to transform and preserve e-journals in a manner that ensures their future usability. We realised early on that both publishers and libraries must be invested in a preservation solution to make it sustainable. We needed cooperation from publishers to secure the preservation rights and to ensure the timely deposit of content. Initially, publishers had significant concerns about placing their content in the care of a

third party. Portico overcame this concern with careful stewardship of the content and transparency about archival policies, and third party preservation is now widely seen as a best practice for creators of digital publications.

Libraries wanted assurance that when they needed access, the preserved content would be available. Libraries also needed the ability to audit and verify the integrity of the preserved content. Portico took an approach that balanced the needs of libraries and publishers: access to the content for scholarly use at an institution is only gained when special circumstances or trigger events occur, and participating libraries have access to an audit site for verification purposes. In addition, publishers have the option to allow Portico to fulfill post-cancellation or perpetual access requests to preserved content. Libraries and publishers make an annual contribution to the e-journal digital preservation service, sharing in the cost of the infrastructure required to preserve this content.

One of the ongoing concerns for any digital preservation entity is that of the free-rider problem:

> Preserved digital assets are nonrival in consumption because once one party preserves the assets, they are for all intents and purposes preserved for all. In these circumstances, the incentive for any single party to incur the cost of preservation is weakened, since the other parties can free ride on the benefits. (Blue Ribbon, 2010)

Institutions may not recognise the need to contribute to digital preservation as the long-term benefit may not be immediate and is for the future academic community. Portico addressed this reality in the e-journal preservation business model by including trigger event and post-cancellation access, whereby only library participants in Portico gain access to the content. In addition, early library adopters of the Portico e-journal preservation service were given a founders discount on their annual fees.

Libraries and publishers join Portico for three year terms. The agreements automatically renew at the end of the term, unless Portico is otherwise alerted. A three year term gives Portico the flexibility to make changes in the pricing model (for example increasing or decreasing annual fees), if it should need to do so in the future. Should a publisher choose to leave Portico, all content deposited remains in the archive with the existing trigger and post-cancellation access mechanisms in place for the preserved content. Portico currently has 365 US and 348 international library participants in its e-journal preservation service, with 131 publisher participants.

Economic model development – electronic book preservation

E-books are increasingly important vehicles for scholarly research and personal enjoyment. By December 2010, e-books made up '9 to 10 percent of trade-book sales' (Bosman, 2010) and as seen in the statistics gathered by the Association of American Publishers and International Digital Publishing Forum, trade retail e-books sales growth increased dramatically over 2009 and 2010 and is set to continue with exponential growth (Forum, 2011). Indeed, on May 19, 2011 Amazon 'announced that since April

[2011], it sold more e-books for the Kindle than it has print books – by a ratio of 105 Kindle books to 100 print books – and that's both hardcover and softcover combined' (Knapp, 2011). This dramatic growth in the trade e-book market is mirrored in the academic market where in 2010 'almost all academic libraries … say they currently offer ebooks to users'. (Journal and Journal, 2010).

In the face of this demand for e-book access, the need for preservation of this content is rapidly increasing. Indeed, the urgency around preservation of e-books in today's world is greater than the urgency around preservation of e-journals in yesterday's world because, by-and-large, libraries are not purchasing books in both print and in electronic form (ebrary, 2007). In response to this increasing need for preservation, Portico preserved the first e-books in the archive in 2009 and formally launched an e-book preservation service in 2011.As with the e-journal preservation service, Portico invested a considerable amount of time working with the library and publisher communities to discuss the need for e-book preservation and develop an appropriate business model. Defining the term 'e-book' proved particularly difficult (Portico eventually settled on a definition that includes any book in digital format, whether born digital or digitised.) Much like e-journals, e-books are published by a wide variety of large and small, commercial and university presses, and licensed and purchased by a diverse group of large and small, US and international libraries. Thus, it was logical to extend the dual support model of e-journal preservation to the Portico e-book preservation service, with the same trigger access and post-cancellation or perpetual access models in place.

There are several notable differences between Portico's e-book and e-journal preservation services. One difference is that the cost to libraries for e-books is reduced because Portico is able to leverage the infrastructure already developed for processing e-journals in the preservation of e-books. Another difference is that in addition to an annual fee, publishers must also pay an initial set-up fee that varies based upon the complexity and consistency of their e-book content, packaging, and metadata. Through working with e-journal content, Portico learned that the effort involved with preserving content varies tremendously based upon the consistency of the supplied data, and we concluded that e-book publishers must pay a variable set-up fee to defray the costs of the initial tool development targeted specifically to their content. As with the e-journal preservation service, Portico offers a founders discount for early adopters of the e-book preservation service.

Economic model development – digitised historical collection preservation

The next frontier in digital preservation of scholarly content for Portico was developing an economic model to support the preservation of digitised historical collections. Digitised historical collections, or d-collections, are different from e-books and e-journals in a number of significant ways. D-collections are typically sold, not licensed, to libraries. Libraries purchase these collections with a substantial up-front fee and then pay the publisher an annual access fee to provide their users with access to the publisher's web interface to the content. Many libraries also receive physical copies of the content on tape or DVD. Another differentiator from e-books and e-journals is that, given the expense of individual d-collections, the number of libraries purchasing a d-

collection is much smaller than the number of libraries that might license or purchase any given e-journal or e-book. In addition to the differences in how d-collection content is licensed and sold, the actual content and metadata in a d-collection is typically straightforward and consistent (which is quite different from e-journals and e-books.)

When Portico began developing an economic model to support preservation of d-collections, we followed the same development process we had with previous services, including extensive discussions with a variety of publishers of digitised historical collections and libraries that purchased these collections. The resulting business plan has some similarities with the e-book and e-journal preservation services, but also some significant differences.

As with the other preservation services, d-collection preservation includes trigger event and post-cancellation access. And, as with the e-book preservation service, publishers are charged an initial set-up fee that is based upon the complexity and consistency of their e-book content, packaging, and metadata. As we discussed d-collection preservation with publishers of d-collections, libraries that purchased d-collections, and libraries that had not purchased d-collections, we began to understand that the original models by which libraries purchase d-collections were significantly different from how libraries license e-journals and how they license or purchase e-books and this necessitated a new preservation business model.

With the d-collection preservation service, only the publisher pays an annual fee and trigger event or post-cancellation access will be provided only to those libraries which purchased the collection from the original publisher (and that access will be provided regardless of the library institution's participation in Portico). The d-collection preservation model allows the publisher to have a single preservation plan for their content that meets the needs of all of the owners of the collection—not just those owners that are also Portico participants—and accommodates the reality that the vast majority of libraries will never purchase this content.

Lessons learned

Portico has developed and implemented three preservation services over the course of nine years. As a result of this experience, we have learned a number of lessons that are broadly applicable to others developing business models:

- Lesson one: The first business model developed for a digital library service is unlikely to be the final one, and it is important for the preservation agency and its constituents to respond to changes in the needs and expectations of libraries and publishers—expect to iterate during the business development phase and even beyond.
- Lesson two: The constituents, participants, or customers of the digital library service must be involved in the development of the business model.
- Lesson three: a single organisation or agency may need to implement different business models to meet the requirements of different content types and stakeholders.

The future

The academic community relies on an increasing amount and diversity of digital content, much of which is in need of digital preservation. While robust preservation options are now in place for traditional forms of scholarly materials, the community is collectively working to resolve long-term preservation of many other content types, including datasets, materials digitised locally by a library or scholar, reference databases, collaborative works, informal forms of scholarly communications (such as blogs, working papers, twitter feeds and so on.), and work that includes multimedia elements. Portico is exploring future approaches and services that will respond to these new preservation needs. We are also exploring the connections among preservation entities—independent archives like Portico, national libraries, and specialised digital repositories—that might be useful, and what form those collaborations might take.

---

[1] http://www.portico.org

[2] www.ithaka.org

[3] http://www.jstor.org/

References

Access, Blue Ribbon Task Force on Sustainable Digital Preservation and (2010), 'Sustainable Economics for a Digital Planet: Enusring Long-Term Access to Digital Information', Feb 2010 at

http://www.jisc.ac.uk/media/documents/publications/reports/2010/brtffinalreport.pdf
accessed Oct 4, 2010.

Bosman, Julie (2010), 'Christmas Gifts May Help E-Books Take Root', *New York Times,*
sec. Booksat
http://www.nytimes.com/2010/12/24/books/24publishing.html?sq=ebook%20sales%20gr
owth&st=cse&adxnnl=1&scp=1&adxnnlx=1304450152-aI+T0nqS3a0m21fXCL2dKg
accessed May 3, 2011.

ebrary (2007), 'ebrary's Global eBook Survey', at
http://www.ebrary.com/corp/collateral/en/Survey/ebrary_eBook_survey_2007.pdf
accessed Oct 4, 2011.

Garrett, John and Waters, Don (1996), 'Preserving Digital Information: Report of the
Task Force on Archiving of Digital Information', (Task Force on Archiving of Digital
Information), 64 May 1, 1996 at http://www.clir.org/pubs/reports/pub63watersgarrett.pdf
accessed Sep 16, 2011.

International Digital Publishing Forum (2011), 'Industry Statistics', http://idpf.org/about-
us/industry-statistics accessed Oct 11, 2011.

Knapp, Alex (2011), 'What Do Amazon's E-Book Sales Mean for the Future of Books?',
Forbes, May 19, 2011at http://blogs.forbes.com/alexknapp/2011/05/19/what-do-
amazons-e-book-sales-mean-for-the-future-of-books/, accessed May 25, 2011.

Kyrillidou, Martha and Morris, Shaneka (2011) *ARL Statistics 2008-2009* [online text],
Association of Research Libraries at http://www.arl.org/bm~doc/arlstat09.pdf, accessed
Oct 4, 2011.

Library Journal and School Library Journal (2010), 'Survey of Ebook Penetration and
Use in U.S. Academic Libraries, Executive Summary', (Library Journal and School
Library Journal) Nov 2010 at
http://www.libraryjournal.com/csp/cms/sites/LJ/info/Reportpurchase.csp accessed Oct
4, 2011.

Waters, Donald (2005), 'Urgent Action Needed to Preserve Scholarly Electronic
Journals', October 15, 2005 at http://www.arl.org/bm~doc/ejournalpreservation_final.pdf
accessed Sep 16, 2011.