

Building to Preserve:  
An Overview of Portico

Eileen Fenton  
Executive Director, Portico

Digital Preservation Management Workshop  
Cornell University  
May 2006



## Topics

- Impetus for Portico
- Developing the Portico model
- Portico's archival approach
- Building an archiving service



## Portico's History

- In 2002, JSTOR initiated a project known as the Electronic-Archiving Initiative, the precursor to Portico.
- The goal was to facilitate the community's transition to reliance upon electronic scholarly journals by developing a technological infrastructure and sustainable archive able to preserve scholarly e-journals.
- Portico was launched in 2005 by JSTOR and Ithaka, with support from The Andrew W. Mellon Foundation.



## Portico's Mission

To preserve scholarly literature published in electronic form  
and to ensure that these materials remain available  
to future generations of scholars, researchers, and students.



## What Portico Is

- Portico is a not-for-profit organization with a mission and singular focus to provide a permanent archive of electronic scholarly journals.
- Portico is centralized archive that is open to all peer-reviewed journals.
- A community-based, cooperative approach to the digital preservation challenge.



## Fundamental Assumptions

- A trusted archival arrangement requires:
  - Mission: Preservation should be mission-central.
  - Economic model: Long-term sustainability is best secured through diverse funding sources.
  - Technological infrastructure: Technology must support unique preservation functions and be updated as technology evolves.
  - Relationships with libraries and publishers: Both parties have preservation needs and will benefit from an archival service.
- Material is preserved for eventual use and access. A completely and perpetually “dark” archive is not desirable.
- Access conditions must be articulated by the archive.



## Original Access Model

- Publishers establish a “moving wall” of up to 10 years for each archived title. The moving wall enables the archive to provide enough access to demonstrate to its vested constituents that the archived content is safe and does so without threatening publishers’ revenues from existing subscriptions and products.
- Publishers permit Portico to provide “perpetual access” beyond the moving wall to those supporting libraries who can not receive satisfactory access from the publisher upon canceling a subscription.
- For libraries supporting the archive, campus-wide access is granted in accordance with the moving wall and without regard to library collections or subscriptions.



## Original Access Model

### Barriers:

- “Moving wall” access too threatening to publishers’ current e-journal business model
- Not all publishers fill post-cancellation access claims and perceived Portico as interfering with their business model
- Libraries did not require – or desire - duplicative access for material already accessible via a licensed resource.



## Original Economic Model

### Publishers:

- Set-up Fee (one-time, per publisher) to help fund initial conversion tools development.
- Deposit Fee (annual, per journal) to defray the cost of adding new content as it is published.

### Libraries:

- Archive Development Fee (one-time, per library) to provide for infrastructure development and future migrations.
- Annual Archive Support Fee to support ongoing archive operations.



## Original Economic Model

### Barriers:

- Two-part fee too complex and resulted in a level that publishers were unwilling to expend since archiving is not part of their core activities.
- Deposit fee per journal creates a disincentive for publishers to deposit their full list of journal titles in the archive limiting the archive's effectiveness.
- Combination of one-time and annual payment raised meaningful barrier for some potential library participants putting archival support and protection beyond their reach.



## Portico's Access Model

- Portico offers access to archived content to only those libraries supporting the archive financially.
- Access is offered only when specific trigger event conditions prevail **and** when titles are no longer available from the publisher or other sources.
- Trigger events include:
  - When a publisher ceases operations and titles are no longer available from any other source.
  - When a publisher ceases to publish and offer a title and it is not offered by another publisher or entity.
  - When back issues are removed from a publisher's offering and are not available elsewhere.
  - Upon catastrophic failure by publisher delivery platform for a sustained period of time.



## Portico's Access Model

- For all libraries supporting Portico, trigger events initiate campus-wide access regardless of whether a library previously subscribed to the publisher's offering.
- Until a trigger event occurs select librarians at participating libraries are granted password-controlled access for archive audit and verification purposes.
- Libraries may also rely upon the Portico archive for post-cancellation or "perpetual" access, **if** a publisher chooses to name Portico as one of the mechanisms designated to meet this obligation.



## Sources of Support

- Support for the archive comes from the primary beneficiaries of the archive - publishers and libraries.
- Contributing publishers supply content.
- Publishers committed to archiving more than 3,200 journals in Portico to date include:
  - Elsevier
  - American Mathematical Society
  - John Wiley & Sons, Inc.
  - University of Chicago Press
  - Oxford University Press
  - American Anthropological Association
  - Berkeley Electronic Press
  - UK Serials Group
  - Symposium Journals
  - SAGE Publications
  - SIAM



## Sources of Support

- Publishers make a financial contribution.
  - An annual Supporting Publisher Contribution to fund initial conversion tools development and to defray the cost of adding new content as it is published.
  - Contributions are tiered and vary according to the size of the publisher's annual journals revenue (subscription and advertising).
  - Annual contributions range from \$250 to \$75,000.



## Sources of Support

- Libraries are asked to make an Annual Archive Support payment to defray ongoing operations, maintenance and enhancement of the archive's technological infrastructure, and content migrations as technology evolves.
- The Annual Archive Support payment is based upon a library's total Library Materials Expenditures (LME) reflecting Portico's role in protecting a critical component of library collections.



## Annual Archive Support (AAS) Levels Individual Library

Total LME	AAS	Total LME	AAS
\$25-30 million	\$24,000	\$3-4 million	\$8,900
\$20-25 million	\$19,000	\$2-3 million	\$7,800
\$15-20 million	\$17,500	\$1-2 million	\$6,700
\$13-15 million	\$16,100	\$750k-1 million	\$5,400
\$11-13 million	\$15,200	\$500-750k	\$4,200
\$9-11 million	\$14,300	\$250-500k	\$3,100
\$7-9 million	\$13,000	\$150-250k	\$1,500
\$5-7 million	\$11,500	<\$150k	1% of LME
\$4-5 million	\$10,000		



## Sources of Support

- All libraries that initiate support for Portico in 2006 and 2007 are designated "Portico Archive Founders."
- Libraries that initiate Portico support in 2006 receive a 25% savings each year for five (5) years on their Annual Archive Support payment.
- Libraries that initiate Portico support in 2007 receive a 10% savings each year for five (5) years on their Annual Archive Support payment.
- To recognize cost savings achieved by working with groups of libraries, Portico offers modest savings to consortia and state-chartered systems.



## Portico Advisory Committee

- John Ewing, American Mathematical Society
- Kevin Guthrie, Ithaka
- Daniel Greenstein, California Digital Library
- Anne R. Kenney, Cornell University Library
- Clifford Lynch, CNI
- Carol Mandel, New York University
- David M. Pilachowski, Williams College
- Rebecca Simon, University of California Press
- Michael Spinella, JSTOR
- Suzanne E. Thorin, Syracuse University Library
- Mary Waltham, Publishing Consultant
- Craig Van Dyck, John Wiley & Sons, Inc.



## Portico's Approach to E-Journal Archiving

- Portico preserves the intellectual content of the journal, including the text, images, and limited functionality such as internal linking. “Look and feel” and publishers’ value-add features are not preserved.
- Publishers deliver to Portico the “source files” of electronic journals (SGML, XML, PDF, etc) including all component elements such as tables, illustrations, and supplements.
- Portico relies upon a format-based migration strategy. Portico converts or “normalizes” the files from their original proprietary format to an archival format based on the NLM Archival DTD and deposits the content in the Portico repository.



## Portico's Approach to E-Journal Archiving

- Portico retains the source files for the long term; the normalized files will be migrated as needed to new formats.
- Where possible, Portico utilizes accepted models (OAIS), open standards (NLM DTD, METS, PREMIS) and open source tools (JHOVE) and supports the development of independent repository certification mechanisms.
- Community expressions of best practice or archival expectations also inform Portico's approach.



## Urgent Action Needed to Preserve Scholarly E-Journals: Qualified preservation archives provide a minimal set of well-defined services.

- Receive files in a standard form from a reliable source.
- Store files in non-proprietary formats.
- Use a standard means of verifying the integrity of files.
- Limit the processing of files, in order to keep costs down, but provide sufficient processing in order to adequately render files.
- Restrict access except under specific conditions.
- Offer transparent means of auditing archival practices.



## From Model to Service

- A collaborative, iterative process is key.
  - Helps to build buy in.
  - Requires time and patience.
  - Can create confusion.
- Understanding and responding to the needs and motivations of key constituents is critical.
  - Libraries
  - Publishers
  - Scholars



## From Model to Service: Incentives

- Libraries
  - Mission relevance is critical.
  - Assurance of ongoing access to e-content is a felt need.
  - Cost reduction through cooperative effort is important; opportunity cost reduction is also valued.
- Publishers
  - Responding to library demand is a key driver.
  - Mission relevance varies across publishers.
  - Cost reduction through cooperative effort is important; opportunity cost reduction is also valued.



## From Model to Service: Content Concerns

- Libraries
  - Content from smaller publishers is perceived to be at significant risk.
  - Content licensed at significant cost is perceived to be important to protect.
- Publishers
  - Many publishers seek holistic archival solutions for content.
  - Many publishers are now seeking a multi-layered archival strategy.



## From Model to Service: Access Concerns

- Libraries
  - Access today is perceived as secure, but access over the long-term is regarded as at risk.
- Publishers
  - The primary concern is maintaining access today; long-term access is a secondary concern.
- Translating these concerns to contractual terms required care and compromise.



## From Model to Service: Economic Viability

- Libraries and Publishers must assess archival options based on relevant factors such as:
  - Mission
  - Economic model
  - Governance
  - Technological infrastructure and capacity
- Archive transparency is key.
- Leadership from early adopters also critical.
  - Early support for Portico comes from JSTOR, Ithaka, the Andrew W. Mellon Foundation, and the Library of Congress.
- Broad community support is also needed.



## From Model to Service: Technological Viability

- Libraries and Publishers must assess technological viability based on factors including:
  - Technological infrastructure and capacity
  - Conformance to best preservation practices
  - Commitment (and capacity) to update technology as needed.
- Archive transparency is again key.
- Formal repository certification and audit mechanisms can support this assessment.



## Benefits of Meeting the Preservation Challenge

- Facilitates the transition to reliance upon electronic resources. May enable system-wide savings through reduced printing, processing and storage of print resources.
- Provides a means of assuring access to e-resources over the long term and protects against gaps in library collections.
- Shared infrastructure or “virtual stacks” reduces preservation costs system wide.
- Electronic formats offer very real opportunities to incorporate preservation concerns at the creation stage thereby reducing costs over the long term.
- Success preserving one digital format may help to ensure the future of other preservation efforts.



Ask me questions!

Eileen Fenton  
[eileen.fenton@portico.org](mailto:eileen.fenton@portico.org)  
[www.portico.org](http://www.portico.org)

