

Portico

An Electronic Archiving Service

Eileen Fenton
Executive Director, Portico

iPRES 2006
October 10, 2006



Meeting the Preservation Challenge Demands Cooperation

- Published scholarly materials require robust preservation.
- Libraries require a means to protect against gaps in e-journal collections. Access today is generally secure; future access is uncertain.
- “Individual libraries cannot address the preservation needs of e-journals on their own.”*
- By cooperatively supporting a limited number of highly reliable archives operating under different technological, economic and governance models, libraries can develop a robust preservation network.

* *E-Journal Archiving Metes and Bounds: A Survey of the Landscape*
www.clir.org/pubs/abstract/pub138abst.html



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 secure 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.
- The Portico archive is open to all peer-reviewed scholarly journals.
- A community-based, cooperative approach to addressing the e-journal preservation challenge.



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) shortly after initial publication.
- Portico converts or "normalizes" the source files from their original proprietary format to an archival format and deposits the content in the Portico repository. To date 138,000+ articles have been preserved in the archive.
- Portico's normalization proceeds carefully and deliberately. The process is focused on long-term preservation requirements not immediate access.



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 and make a financial contribution ranging from \$250 to \$75,000.
- Publishers from across the spectrum have thus far committed to archive more than 5,000 journals in Portico. Publishers come from across the spectrum, for example:
 - Elsevier (commercial)
 - Oxford University Press (university press)
 - American Institute of Physics (scholarly society)
 - The Berkeley Electronic Press (e-only publisher)



Sources of Support

- Libraries are asked to make an Annual Archive Support (AAS) payment to defray cost of ongoing operations, maintenance and content migrations.
- The AAS payment is based upon a library's total Library Materials Expenditures and ranges from \$1,500 to \$24,000 annually.
- All libraries that initiate support for Portico in 2006 and 2007 are designated "Portico Archive Founders" and make a significantly reduced AAS payment. Savings are also available to systems and consortia.
- To date 80 participating libraries are Portico Archive Founders; 50 are in the process becoming Founders.
- Participants range from Grove City College to the University of California system.



*E-Journal Archiving Metes and Bounds: A Survey of the Landscape**

- Reviews 12 e-journal archiving programs summarizing key factors relevant to library decision-makers.
- Offers recommendations to publishers, libraries and e-journal archives.
- Recommendations to Publishers:
 - Work with at least one digital archiving partner and communicate overtly about archival arrangements.
 - Work with archival partners to ensure title coverage is readily known.
 - Extend critical rights to archival partners.

* <http://www.clir.org/pubs/abstract/pub138abst.html>



E-Journal Archiving Metes and Bounds: A Survey of the Landscape

- Recommendations to Libraries:
 - Press publishers to make acceptable archival arrangements.
 - Share information about e-journal archiving strategies and decisions.
 - Participate in at least one e-journal archiving initiative.
 - Collectively press for archival programs that meet the needs libraries across the spectrum and for transparent practices from archives.
 - Develop a registry recording which publications are archived by which initiatives.
 - Lobby archival programs to participate in a collaborative network.



E-Journal Archiving Metes and Bounds: A Survey of the Landscape

- Recommendations to E-Journal Archiving Programs
 - Present public evidence of the minimal level of services necessary for a well-managed collection.
 - Be overt about details of content included in program.
 - Ensure that after content is ingested it cannot be removed.
 - Conduct a study to determine what rights and responsibilities are necessary to reflect in contractual archival agreements.
 - Negotiate rights that consider the impact of content eventually entering the public domain.
 - Form a network of support and mutual dependence among archival programs.



Key Factors for Library Decision-Makers

1. "The repository should have both an explicit mission and the necessary mandate to perform long-term e-journal archiving."
 - Portico's mission carries a clear preservation focus and mandate.
2. "Rights and responsibilities associated with preserving e-journals should be clearly enumerated and remain viable over long periods."
 - Portico's library and publisher agreements clearly specify the rights of the library, publisher and archive.
 - Once deposited content remains in the archive. Access terms are clearly spelled out.
 - Agreements were developed through iterative process involving libraries and publishers.



Key Factors for Library Decision-Makers

3. "The repository should be explicit about which scholarly publications it is archiving and for whom."
 - A list of participating publishers and committed titles is available at www.portico.org. Both are updated as new publisher agreements are made and processed.
 - Coverage is more challenging information to share.
 - Date range: Publisher's intentions may differ from file availability.
 - Volume/issue: Use of these conventions changing as e-publishing practice evolves.
 - A list of participating libraries will be posted shortly.



Key Factors for Library Decision-Makers

4. "E-journal archiving programs should be assessed on the basis of their ability to offer a minimal set of well-defined services."

– Receive e-journal files in standard form

- Portico accepts standard formats.
- Portico collaborates with publishers to ensure completeness.
- Some elements such as "look and feel" are not captured in source files.

– Store files in a non-proprietary formats

- Content is converted to archival format based on the NLM DTD.
- Where files can not be converted; byte preservation must be applied.



Key Factors for Library Decision-Makers

4. "E-journal archiving programs should be assessed on the basis of their ability to offer a minimal set of well-defined services."
 - Use standard integrity checks
 - Virus scan
 - Checksum
 - JHOVE
 - Limit processing to contain costs but assure that files can be located and rendered
 - Pre-emptive migration assures rendition.
 - Limiting the number of formats to maintain reduces archival complexity.



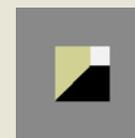
Key Factors for Library Decision-Makers

4. "E-journal archiving programs should be assessed on the basis of their ability to offer a minimal set of well-defined services."
 - Guard against loss through redundancy
 - The Portico archive is replicated offline with copies held in distributed settings under both commercial and not-for-profit control.
 - Mirror sites are not yet implemented; more appropriate for high-access resources.
 - Offer open, transparent means of audit
 - Portico participated in CRL project to develop audit process.
 - Participating libraries are also empowered as auditors.



Key Factors for Library Decision-Makers

5. "A repository should negotiate with publishers to ensure that the digital archiving program has the right, and is expected, to make preserved information available to libraries under certain conditions."
 - As a "dim" archive Portico secures the right to receive, normalize, migrate and display content for long-term preservation purposes.
 - Access conditions are specified in Portico's publisher and library agreements.



Key Factors for Library Decision-Makers

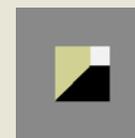
6. "Repositories must be organizationally viable."
 - Financial stability is pursued by securing start up and ongoing support.
 - Initial support is received from The Andrew W. Mellon Foundation, JSTOR, Ithaka, and the Library of Congress
 - Ongoing support is secured through agreements with participating publishers and libraries.
7. "Repositories will work as part of a network."
 - Portico is an active partner in the Library of Congress' National Digital Information Infrastructure Preservation Program (NDIIPP).
 - Informal collaborations continues with other archiving entities.



Conclusions

- “Digital preservation represents one of the grand challenges facing higher education.”*
- If we are to meet this challenge, all parties engaged in scholarly communication – publishers, libraries, and archives – must contribute to the solution.
- Cooperation and collaboration are critical.

* “Urgent Action Needed to Preserve Scholarly Electronic Journals”
www.arl.org/osc/EjournalPreservation_Final.pdf



Questions?

Eileen Fenton
eileen.fenton@portico.org
www.portico.org

