

Portico

An Electronic Archiving Service

Eileen Fenton

Executive Director, Portico

What Works In Archiving?
Society for Scholarly Publishing
November 15, 2006



The Need

- “Digital preservation represents one of the grand challenges facing higher education.” *
- The challenges are many: technical, organizational and economic.
- Diverse solutions increase the likelihood of system-wide success.
- All parties – publishers, libraries, and archives – must contribute to the solution.

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



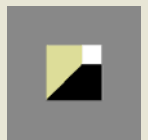
Portico's History

- 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.
- A collaboratively developed, cooperative archiving model which addresses the technological, organizational and economic components of the long-term preservation challenge.
- A managed, operational archive with geographically distributed replication.



Portico Advisory Committee

- John Ewing, American Mathematical Society
- Kevin Guthrie, Ithaka
- Daniel Greenstein, University of California
- 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: Content Scope

In scope:

- Electronic scholarly, peer reviewed journals
- Intellectual content of the journal, including text, tables, images, supplemental files
- Limited functionality such as internal linking

Out of scope:

- Full features and functionality of publisher's delivery platform
- Ephemeral look and layout of the online journal



Portico's Approach: Migration Supplemented with Byte Preservation

- Publishers deliver “source files” of electronic journals (SGML, XML, PDF, etc.) to Portico.
- Using specialized software Portico converts proprietary source files from multiple publishers to an archival format suitable for long-term preservation. Portico's preservation format is based on the NLM Archiving DTD.
- Source and normalized files are deposited in the archive. Once deposited, content must remain in the archive.
- To date more than 200,000 articles are preserved in the Portico archive.



Portico's Approach: Active Archive Management

- For security the archive is replicated, and copies are geographically distributed.
- Portico monitors formats and will develop and test migration tools as necessary.
- When needed, Portico will migrate the archive to new formats or technologies.
- Portico conducts internal audits of archived content and participates in developing external audit processes.



Portico's Approach: Access

- Portico offers access to archived content to only those libraries supporting the archive.
- Portico's delivery infrastructure leverages JSTOR's existing technology and investment.
- Access is offered only when specific trigger event conditions prevail **and** when titles are no longer available from the publisher or other sources.



Portico's Approach: Access

- 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 Approach: Access

- Trigger events initiate campus-wide access for all libraries supporting the archive 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 rely upon the Portico archive for post-cancellation 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 an annual financial contribution ranging from \$250 to \$75,000 depending upon journal revenues.
- To date more than 5,200 journals have been promised to the Portico archive. Participating 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 make an Annual Archive Support (AAS) payment based upon total library materials expenditures. AAS payments range 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.
- To date more than 100 libraries are Portico Archive Founders. Participants range from Grove City College to the University of California system.



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

- 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.
- Libraries:
 - Press publishers to make acceptable archival arrangements.
 - Participate in at least one e-journal archiving initiative.

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



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

- Archives:
 - Present public evidence of the minimal level of services necessary for a well-managed collection.
 - Ensure that after content is ingested it cannot be removed.
 - Be overt about details of content included in program.

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



Conclusion

- All parties with a vested interest in scholarly communication – publishers, libraries, and archives – have a role to play.
- Cooperative support of a limited number of archives operating under different technological, economic and governance models, can yield a robust preservation network.
- The time to address the “grand challenge” of e-journal preservation is now.



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

