



PORTICO

Portico

An Important Component of a New Preservation Infrastructure

Kevin Guthrie
President, Ithaka

ARL Membership Meeting
October 27, 2005



This Is a Very Challenging Problem

- Technological challenges are plentiful.
- Collaboration is essential.
- Access to archived literature is a key issue for publishers and libraries, but they come at it from different perspectives.
- Financial sustainability must be secured.



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.



Portico's Governance Structure

- Portico is an enterprise being incubated within Ithaka.
- Ithaka is an independent 501(c)(3) not-for-profit organization affiliated with JSTOR. It has its own governing Board of Trustees. This structure enables the community to benefit from connections between JSTOR and Ithaka/Portico while also enabling and facilitating independent innovation in a dynamic environment.
- Portico has established an advisory committee comprised of librarians and publishers with a vested interest in electronic preservation of scholarly journal literature. The first meeting of the advisory committee was in September 2005.



Portico Assumes Archiving Requires

- A trusted archival partner with an appropriate:
 - 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.
 - Collaborative relationships with libraries and publishers: Both parties have preservation needs and will benefit from an archival service.
- Adoption and ongoing watch of preservation and publishing standards, best practices, and format evolution with action taken well in advance of technology or format obsolescence.



What Portico Is

- An organization with a mission and singular focus to provide a permanent archive of born-electronic scholarly journals to ensure that over the long-term a valid, reliable copy of the work exists and is accessible.
- A centralized archive that is open to all peer-reviewed journals. Archived journals may have a print version in addition to an electronic version or they may be available only in electronic form.
- 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.



Portico's Approach to E-Journal Archiving

- Publishers deliver to Portico the “source files” of electronic journals (SGML, XML, PDF, etc).
- 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.
- Where possible, Portico utilizes open standards (NLM DTD) and open source tools (JHOVE).
- Portico supports the development of independent repository certification mechanisms and is participating in the Center for Research Libraries Audit and Certification of Digital Archives project.



Portico's Approach to E-Journal Archiving

- By normalizing content at the time of deposit, Portico:
 - Stores files in non-proprietary formats that can be effectively managed over the long term.
 - Can identify problematic or missing files and conduct follow-up intervention with publishers.
 - Is positioned to render files to supporting libraries, under appropriate conditions.
- Portico retains the source files for the long term; the normalized files will be migrated as needed to new formats.
- Portico offers access to archived content to participating libraries under specific trigger event conditions.



Portico Archive Service & Access Model

- For participating libraries, trigger events initiate campus-wide access. Trigger events include:
 - When a publisher ceases operations and no entity purchases and makes titles accessible
 - When a publisher ceases to publish a title and it is not offered by another publisher
 - When back issues are removed from a publisher's site and not available elsewhere
 - Upon catastrophic failure by publisher delivery platform for a sustained period of time
 - When a publisher chooses to rely upon Portico to meet perpetual access obligations.
- Pre-trigger event, select librarians at participating libraries are granted password-controlled access for verification purposes.



Sustainability

- Diversified revenue streams are important to the longevity of the archive. Support for the archive comes from the primary beneficiaries of the archive - publishers and libraries.
- Government agencies and charitable foundations are also expected to provide support.
- JSTOR, The Andrew W. Mellon Foundation, Ithaka, and the Library of Congress are investing significantly in Portico's infrastructure development.



Sustainability

- Publisher 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)
- Library Contributions
 - A one-time Archive Development Contribution to fund infrastructure development and to prepare for future migrations.
 - An annual Archive Support Contribution to defray the cost of adding new content as it is published.



Urgent Call for Action: 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.



Kevin Guthrie
President, Ithaka

www.ithaka.org
kg@ithaka.org

