JSTOR and Preserving E-Journals

Eileen Fenton, Executive Director
Electronic-Archiving Initiative

Digital Preservation Coalition Forum   June 23, 2004
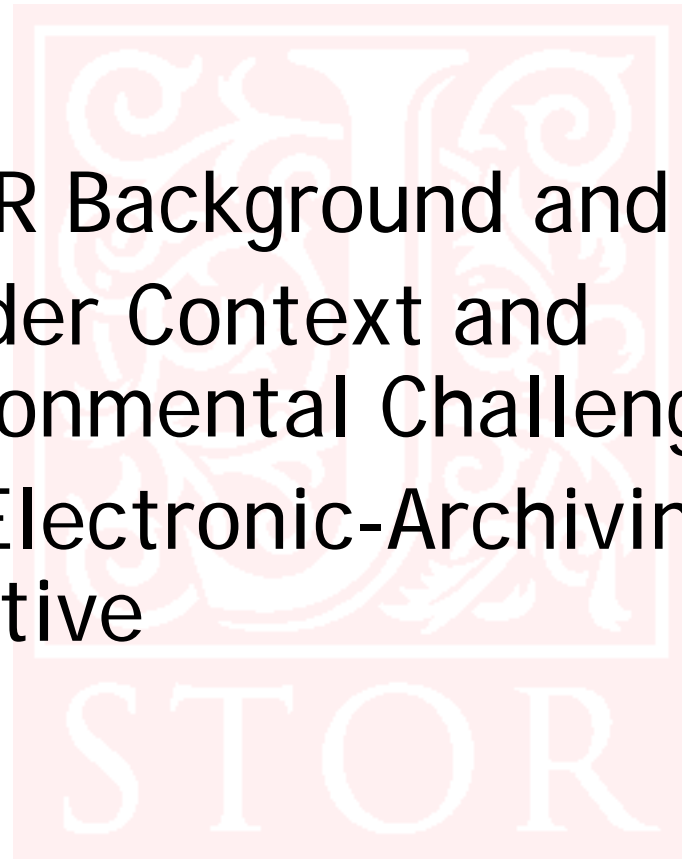
Overview

- JSTOR Background and Context
- Broader Context and Environmental Challenges
- The Electronic-Archiving Initiative

About JSTOR

- Mission-driven goal-oriented not-for-profit organization in the academic community
- JSTOR's mission is to help the scholarly community take advantage of the advances in information technology
- JSTOR takes a system-wide perspective, seeking benefits for libraries, publishers and scholars

About JSTOR

- Reliable and comprehensive archive of important scholarly journal literature
- Primary functions: preservation and access over time
- At present, nearly 400 journals available in ten collections
- Over 15 million pages
- Archival content only
- Focus on the social sciences and humanities

About JSTOR

- 230 participating publishers
- 1,233 U.S. participants
- 810 participants outside the US
  - 86 countries represented

The JSTOR Context

- From the inception of JSTOR the inclusion of electronic journals in the archive was anticipated.

- The Electronic-Archiving Initiative was born out of JSTOR's commitment to serve as a long-term archive of journal literature, a commitment that is format neutral.

- JSTOR approaches this challenge with a system-wide perspective, seeking to reduce costs and improve convenience for all participants in the scholarly communication cycle.

The JSTOR Context

- Archiving electronic resources requires a significant investment in the development of organizational and technological infrastructure.

- Maximum system-wide benefit will be achieved by archiving a broad array of content extending well beyond JSTOR's current scope and mission.

- A new entity is needed. Launching new organizations is beyond the scope of JSTOR's mission.

ITHAKA

The Broader Context: Environmental Challenges

- Rapidly changing information technologies are having a potentially transformative impact on the scholarly enterprise and on higher education

- This period of change presents opportunities and challenges

- Some of these opportunities will be addressed (and should be addressed) by commercial ventures

- Some will not be addressed by, to use Adam Smith's phrase, "the invisible hand" of the market

I T H A K A

The Broader Context: Environmental Challenges

- The not-for-profit sector is responsible for filling gaps not addressed by the free market

- There does not exist a robust mechanism to stimulate innovative not-for-profit organizations addressing mission-critical needs in higher education

- Foundations make grants, but they are not set up to provide the full array of human, logistical and financial resources required to create organizations.

Mission

- Ithaka has been founded to accelerate the creation, development and success of not-for-profit organizations focused on deploying new technologies for the benefit of higher education

- It brings together:
  - Financial resources from (initially) three foundations (Mellon, Hewlett, Niarchos)
  - The experience derived from the creation of JSTOR, including a conviction that organizations such as JSTOR can contribute enormous value to the scholarly community
  - Relationships in all sectors and at all levels of the higher education community (developed at the sponsoring foundations and through JSTOR)

I T H A K A

Four areas of activity

1. Incubating promising and important projects and new ideas;
2. Supporting affiliated organizations through sharing of resources;
3. Conducting research on the impact of advancing technologies on the scholarly community;
4. Providing strategic assistance and advice to other organizations engaged in related activities.

The Electronic-Archiving Initiative

The mission of the Electronic-Archiving Initiative is to preserve scholarly literature published in electronic form and to ensure that these materials remain available to future generations of scholars, researchers, and students.

E-Archive expects to take responsibility for archiving a broad range of scholarly e-journals and journal-like resources.

JSTOR, Ithaka, and The Andrew W. Mellon Foundation are together supporting the development of E-Archive.

Components of a Trusted Archive

1. Mission

   – Mission is critical because it drives resource allocation and routine organizational priorities and activities.

2. Business Model

   – Sustainability is key.
   – The archive must generate funds adequate to cover the work of the archive from sufficiently diversified sources.

Components of a Trusted Archive

3. Technical Infrastructure
   – An infrastructure must be developed which supports the key functions of the archive:
     • Ingest
     • Verification
     • Storage
     • Delivery
     • Migration
   – It must allow for sufficient replication with appropriate geographic spread.

Components of a Trusted Archive

4. Relations with Libraries

   – The archive must meet the needs of the library community and the scholars they serve.

4. Relations with Content Producers

   – The archive must secure the rights necessary to the archival task and must arrange for timely, secure deposit of content.

E-Archive Approach

- "Source File" archive: E-Archive will seek to preserve the source files which comprise publishers' e-journals.

- This approach captures some content which is not presented online (i.e., higher resolution graphics).

- This approach makes it very difficult to capture certain elements such as dynamic advertisements and editorial information.

E-Archive Areas of Activity

1. Define an archival service.

2. Develop a business model which ensures the short-, mid-, and long-term sustainability of the archive.

3. Design and build technological infrastructure.

4. Develop content processing protocols and tools.

5. Research the economic impact of electronic resources on operations costs for libraries and content producers.

Activities to Date

- Engaged libraries in discussions of e-archiving needs and challenges.
  - Emerging themes:
    - There is a widespread desire for a trusted solution to the e-archiving need. This is true for academic libraries of all sizes.
    - Regardless of institution size, librarians believe it is important for their own institution to contribute to the solution of this problem.
    - Librarians recognize that e-archiving raises complex technical and business issues.
    - Librarians are concerned about "perpetual access" to materials that have been "bought and paid for."

Activities to Date

- Seeking publishers' perspectives on the e-archiving challenge.
- Signed ten publishers to participate in pilot, developmental phase.
  - Association of Computing Machinery
  - American Economic Association
  - American Mathematical Society
  - American Political Science Association
  - Blackwell Publishing
  - Ecological Society of America
  - National Academy of Sciences
  - The Royal Society
  - University of Chicago
  - John Wiley & Sons, Inc.
- Analyzing sample e-journal source file data.

Activities to Date

- Created a working prototype archive; production-level archive now in development.

- Developed tools to process and deposit sample content into the prototype archive. Creating verification and normalization procedures.

- With support from the Ithaka Research group, designed and completed a study of libraries' non-subscription costs for print vs. electronic periodicals.

Current Focus

- Finalize business model
  - Can an archival service successfully balance the needs of publishers and libraries?
  - Should the archive provide access to the materials that it holds? If so, how, to whom, and under what conditions?
  - Can E-Archive generate sufficient revenue from its most likely sources (libraries, publishers/content owners, government agencies and foundations)?

- Complete work on production-level archival repository
  - E-Archive must develop workflows and quality control tools which match the complexities of the multiple formats it will archive.
  - The design of the repository must be sufficiently flexible to accommodate e-journal developments not yet seen.

ITHAKA

The Electronic-Archiving Initiative

Eileen Fenton
egfenton@ithaka.org