Bringing digitized special collections into the research workflow through JSTOR

CNI Fall Membership Meeting | December 12, 2023

Presented by:
Kevin Guthrie, President, ITHAKA
Bruce Heterick, SVP, Open Collections & Infrastructure, ITHAKA
When Bill Bowen envisioned JSTOR in 1995 he reached for a very big idea through a relatively small initiative.

While our immediate goal was to convert backfiles of important scholarly journals into digital form, the big idea started the transformation of part the infrastructure that supports research, teaching, and learning – from buildings and shelves to servers and networks.
We know …

- increasingly, faculty want to teach with their and other libraries’ digitized special and primary source collections;

- increasingly, students are being asked to, or want to, do research with these unique collections

- journals and books are increasingly available electronically, but only a very small portion of a library’s special/distinctive/primary source collections have been digitized

- the tapestry of tools, services, systems that have traditionally been deployed to build, manage, share, and preserve these collections are grounded in physical infrastructure
We know …

- consumers, faculty, students, will be looking to do their work digitally
- libraries will need to provide the collections and services they used to provide physically, through the network
- there are economies of scale associated with some of these activities, while some of them remain local (although AI will threaten these…)
- libraries will need to decide whether they will build these services themselves, collaborate with other libraries to create them, or buy these services from others
Academic and research libraries are in the midst of a major strategic reinvention to enable the next generation of instruction, learning, and scholarship.

Libraries’ collections must be increasingly digital, and libraries are transitioning to a services-first model that requires new kinds of talent and infrastructure.

Libraries must demonstrate value to their home institutions, and to do so their locally owned collections must have impact, but most of the current systems for access are outside today’s research, teaching, and learning workflows.

Libraries need to make, build, or buy systems and platforms to manage this digital work, with the twin goals of controlling their costs and increasing impact.
More than ever before, we see how the digital infrastructure we’ve worked together to build has the power to deliver more equitable and impactful access to knowledge and education.
What we are doing

- Together, we have been building infrastructure for the past two decades that help libraries make the transition from print to digital (JSTOR and Portico)

- Our access platform (JSTOR) is in the middle of the “research and learning workflow”; it is a starting place for a substantial number of faculty, students, and researchers in the humanities and social sciences

- Our preservation platform (Portico) is widely recognized as a gold standard for the managed preservation of scholarly content

- We are now making this infrastructure available to institutions - as a trusted, not-for-profit partner - to help institutions build, manage, share, and preserve these collections in a more impactful and cost-effective manner
Hypotheses

1. Make the open collections more discoverable to researchers globally

2. Make the research process more seamless

3. Help institutions fulfill their mission as stewards of their digital collections for the long-term
Open Community Collections initiative

- Charter period ran from January 2020 - June 2023
- Invited institutions to make their digital collections available on JSTOR for no fee
  - 310 institutions shared 1,850 collections on JSTOR by the initiative’s conclusion
- Also explored additional services:
  - Using JSTOR’s cataloging tool - previously developed with Artstor - for building and publishing collections to JSTOR
  - Providing managed preservation of libraries’ collections in Portico - previously only available for publishers’ content
Discovery and reach

📖 **62%** of item requests coming from open web

🔍 **13,000+** unique institutions accessing from 225+ countries/territories

↥ **43%** of usage from sessions starting at JSTOR

🔍 **31%** of usage from discovery on Google
What is working:

- the “JSTOR effect”
- amplified indexing by Google
- connecting secondary literature to the primary source content being contributed by institutions
- making audio and video available with auto-generated transcripts
- making it easy for institutions to bring content built in other repositories into where it can be found and used on JSTOR
- institutions are seeing Portico as a viable long-term preservation alternative
Building a system-wide infrastructure for archives

Charleston Conference, November 9, 2023
Queens College (CUNY): Driving targeted discovery

- In March 2022 published the “Greek News” collection to JSTOR
- 900+ issues of a bilingual newsweekly covering 2003-2021
- Collection stewarded jointly by the Hellenic American Project and Special Collections and Archives.
- Sharing the collection on JSTOR led to increased discovery and use from Greece
Success at Queens

- Donors, grants, renovations, increased staffing, library school fellowships, great relationship with administration and development office (fundraising capacity)
- Increased visibility directly related to the JSTOR platform (featured on NPR)
Putting archives and special collections on a firm footing

• Beyond Queens, some CUNY campuses already publishing successfully on Open Collections.
• “New Deal” with JSTOR: parlaying savings in our collections bill into a sustainable system-wide infrastructure.
• No need for anyone to engage in a difficult procurement (archivists were worried about the end of the pilot)
• New commitment to work together and compromise
• Campus libraries also saw collections savings at a key austerity moment.
SCELC JSTOR Shared Collections Digital Cooperative

Teri Oaks Gallaway
SCELC Executive Director
SCELC: Additive impact

- Analysis done in collaboration with SCELC, Caltech, University of San Francisco, and University of the Pacific in Spring 2022
  - Presented at the International Conference on Open Repositories in June 2022
- Source repositories maintained usage trends after making collections available on JSTOR (and sometimes increased)
- Collections on JSTOR are reaching different audiences
- Usage on JSTOR was additive (and still growing) to source repositories
Why seek to build Consortial Services around JSTOR community/shared collections?

- Similar Missions / Equity Objectives: Find an affordable, sustainable way to serve ALL libraries

- Critical mass of 16 library participants to start

- JSTORs Fit-for-purpose approach will make it easier to grow
  - No Migration from other platforms is required
  - Can be alternate platform or only one
  - Can support options for a wide range of libraries

- JSTOR is developing a growing number of lightweight, unbundled services that SCELC should be able to support with its current expertise
Skidmore: full stack

- JSTOR Forum to build the collections
- Shared Collections to provide access to the collections via JSTOR
- Preserved Collections to manage the long-term preservation of the collections via Portico
Content

JSTOR Archive Collections
Books at JSTOR
Artstor Digital Library
Research Reports
Open Shared Collections
Reveal Digital

Services

Portico and Preserved Collections
Shared Collections
Audio/Video
Syndication
Harvesting

Tools

Constellate (TDM)
JSTOR Forum (cataloging)
genAI chat
ML and AI for metadata enhancement