Experimenting with JSTOR as Open Infrastructure for Primary Source Collections

Bruce Heterick
Senior Vice President, Open Collections & Infrastructure
ITHAKA

Clifford Anderson
Associate University Librarian for Research and Digital Strategy
Vanderbilt University

Jean & Alexander Heard Libraries
ITHAKA and Primary Sources

- 25 years bringing content together in humanities and social sciences
Content on JSTOR Platform

- Open Community Collections
- Reveal Digital
- Research Reports
- Artstor Images on JSTOR
- Books at JSTOR
- JSTOR Archive Collections

Timeline:
- 1997
- 2010
- 2015
- 2018
- 2019
- 2020

Y-axis:
- Open
- Closed
ITHAKA and Primary Sources

- **25 years** bringing content together in humanities and social sciences

- Journals (1997)
  - -> Books (2010)
  - -> Research Reports (2015)
  - -> Reveal Digital (2018)
  - -> Images via Artstor (2019)
  - > Contributed Content from Institutions (2020)

  **All underpinned with long-term preservation from Portico**

- What we heard from our participating libraries:
  - Faculty increasingly want to teach with primary sources;
  - Students - particularly early career researchers - want to do research with primary sources

- JSTOR has traditionally been a “research” platform; can it also become a “teaching” platform as well?
The Problems

- **Difficult to discover:** seldom in places where research starts
- **Out of context:** unconnected to the secondary literature (journal articles and book chapters)
- **Expensive to maintain:** ongoing investment required to keep relevant with changing research techniques
- **Limited Reach:** don’t typically reach a global audience
- **Long-term preservation:** not a priority
Two Working Hypotheses

1. Making this content available in the JSTOR channel will increase its usage footprint and make the open collections more discoverable to researchers globally.

2. Co-locating these primary source collections alongside the important secondary literature (journal articles, book chapters, research reports) already on the JSTOR platform will allow researchers to find this content in context and make the research process more seamless.
Vanderbilt on Community Collections

Vanderbilt University

Search 1,539 items in 5 collections

Search
Usage Statistics

4,505
Total Item Requests

30%
Items Accessed

105
Countries

35
Institution Landing Page Views

Item Requests by Collection
Challenges
Progress in Growing Use and Collections

 GTA 850+  
collections published to JSTOR

 G 200+  
unique contributing institutions

 F 62%  
of item requests coming from non-JSTOR IP ranges (i.e. open web)

 E 7,760  
unique institutions accessing from 180+ countries (2021)
Steady Growth in Participants and Collections

The graph shows the steady increase in both Institutions and Collections from 1/1/2020 to 1/1/2022. The line for Institutions remains below the line for Collections throughout the period.
Community Collections Cumulative Usage: 2020-2022

+175%
Important Learnings Thus Far

- Image content still dominates what is in repositories: Reinforced the importance of having robust metadata in making relevant image content discoverable in a vast sea of text content.

- Google referrals have grown from 25% (2020) to 40% (2021) of traffic; and are increasing; the JSTOR platform can be beneficial for Google ranking of this content.

- “Open” is working.

- Preservation is an immediate priority.

- Metadata enhancement/enrichment is essential on the dimensions of discovery, preservation, usability for TDM, accessibility, and inclusiveness.
Next Steps

- Continue adding collections from new and existing institutions
  - Charter Period ends December 31, 2022
  - **Institutions can still participate for no fees**

- Finalize business models for Hosting Service and Preservation Service;
  - **perhaps make Preservation Service available later in 2022**

- Continue platform work required to more substantially connect the secondary literature with these primary source/special/distinctive collections

- Continue work to understand the value of these collections for text analytics
  - **ITHAKA launching Constellate TDM service**

- Define and build Metadata Enhancement Service with library partners