

I T H A K A

JSTOR | PORTICO | ITHAKA S+R

**What to Withdraw:
Print Collections Management
in the Wake of Digitization**

Roger Schonfeld
Ross Housewright

January 2010

Our Mission

I T H A K A
JSTOR | PORTICO | ITHAKA S+R

ITHAKA is a not-for-profit organization dedicated to helping the academic community take full advantage of rapidly advancing information and networking technologies. We serve scholars, researchers, and students by providing the content, tools, and services needed to preserve the scholarly record and to advance research and teaching in sustainable ways. We are committed to working in collaboration with other organizations to maximize benefits to our stakeholders.



Our Services



- **Ithaka S+R** works with initiatives and organizations to develop sustainable business models and conducts research and analysis on the impact of digital media on the academic community as a whole.



- **JSTOR** helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive of over 1,000 academic journals and other content. JSTOR uses information technology and tools to increase productivity and facilitate new forms of scholarship.



PORTICO

- **Portico** preserves scholarly literature published in electronic form—more than 10,000 e-journals and 28,000 e-books—and ensures that these materials remain accessible to future scholars, researchers, and students.



Organizational Commitment to Preservation

- JSTOR – actively preserving over one thousand academic journals in both digital and print formats
- Portico – digital preservation service providing a permanent archive of electronic journals, books, and other scholarly content
- Ithaka S+R
 - Extensive work focusing on print collections management during a format transition
 - In addition to our work focused on scholarly journals, we have recently completed a project on government documents
 - Emphasis on developing policy framework to help libraries negotiate a format transition without sacrificing preservation

The Dilemma of Print Preservation

- As an increasingly greater share of collections is digitized, libraries face growing pressure to reduce the size of their corresponding print collections
- But it may be important to keep at least some print copies on behalf of the library community, even if not locally
- The need for remaining print preservation is poorly understood, the responsibility for it is often unclear, and libraries lack nuanced information about print preservation activities elsewhere
- As a result, any given library is unable to identify those materials that are sufficiently well-preserved elsewhere that the local copy is not needed for community preservation needs

What to Withdraw: Vision

- Our objective is to provide information about community-wide preservation needs and activities that can assist library decision-making processes
- The What to Withdraw framework applies to general collections of published scholarly journals – it does not apply to monographs, rare books, newspapers, journals that have been categorized as special collections, or any other collections or format types
- We are not advising any individual library that it should, or should not, retain or withdraw, any of its holdings

What to Withdraw: Activities & Plans

- The overall intellectual framework is contained in the report “What to Withdraw: Print Collections Management in the Wake of Digitization”
- We are now bringing together data from a variety of sources and making them actionable via tools to help libraries use this framework to support decision-making

I T H A K A

JSTOR | PORTICO | ITHAKA S+R

What to Withdraw Framework

What to Withdraw Framework: Overview

- Define rationales for print preservation in the presence of digitized surrogate
- Based on these rationales, categorize journals according to their relative preservation needs
- Develop system-wide print preservation goals for each category
- Determine the levels of print preservation required for each goal

What to Withdraw: Rationales for Print Preservation

Rationales that are relevant for the whole community

- Fix scanning errors
- Inadequate scanning standards & practices
- Inadequate digital preservation
- Unreliable access

- Scholarly needs
- Campus politics

Additional rationales that apply at the local level

What to Withdraw: Criteria

- Our framework considers several criteria in categorizing journals:
 - The quality at which the journal has been digitized
 - The ongoing quality assurance processes to fix errors
 - Digital preservation
 - The relative importance of images in a journal
 - The reliability of digital access, in terms of business model, terms and conditions of access

What to Withdraw: Examples of Categories

	Ideal Scenario	Inadequate Digital Preservation	Image Intensive	Inadequate Digitization
High digitization quality standards?	Yes	Yes	Yes	No
Active error-correction?	Yes	Yes	Yes	Yes
Reliable digital preservation?	Yes	No	Yes	Yes
Image intensive?	No	No	Yes	No
Reliable terms?	Yes	Yes	Yes	Yes
Minimum time horizon for retention of some copies system-wide	20 years; and a candidate for local withdrawal	n/a	n/a	100 years; and may not be a candidate for withdrawal at research libraries



What to Withdraw: Model

- Ithaka S+R commissioned Candace Yano, operations researcher at UC Berkeley, to develop a model for how many copies are needed to meet these preservation goals
- Assumption that dark archives have an annual “loss rate” of 0.1%

What to Withdraw: Model

- Ithaka S+R commissioned Candace Yano, operations researcher at UC Berkeley, to develop a model for how many copies are needed to meet these preservation goals
- Assumption that dark archives have an annual “loss rate” of 0.1%

Scenario	Time Horizon	Probability of Success	Number of “Perfect,” Uncirculating Copies Required
Ideal Scenario	20	>99%	2

What to Withdraw: Some Recommendations

- Local print collections management decisions should be made with awareness of community-wide preservation needs and activities
- Greater information-sharing is needed about community-wide print preservation activities
- Greater information-sharing is needed about quality and reliability of digitized collections

What to Withdraw: Ithaka S+R Next Steps

- Encourage and facilitate greater information sharing about print preservation activities
- Encourage and facilitate greater information sharing about ability of digital materials to substitute for print originals
- Build tools to help libraries integrate this information into decision-making about local print collections management
- Ultimately, provide a single source for comprehensive information about all digitized journals and their preservation status

I T H A K A

JSTOR | PORTICO | ITHAKA S+R

**What to Withdraw:
Decision-Support Tool –
Proof of Concept**

What to Withdraw: Immediate Opportunities

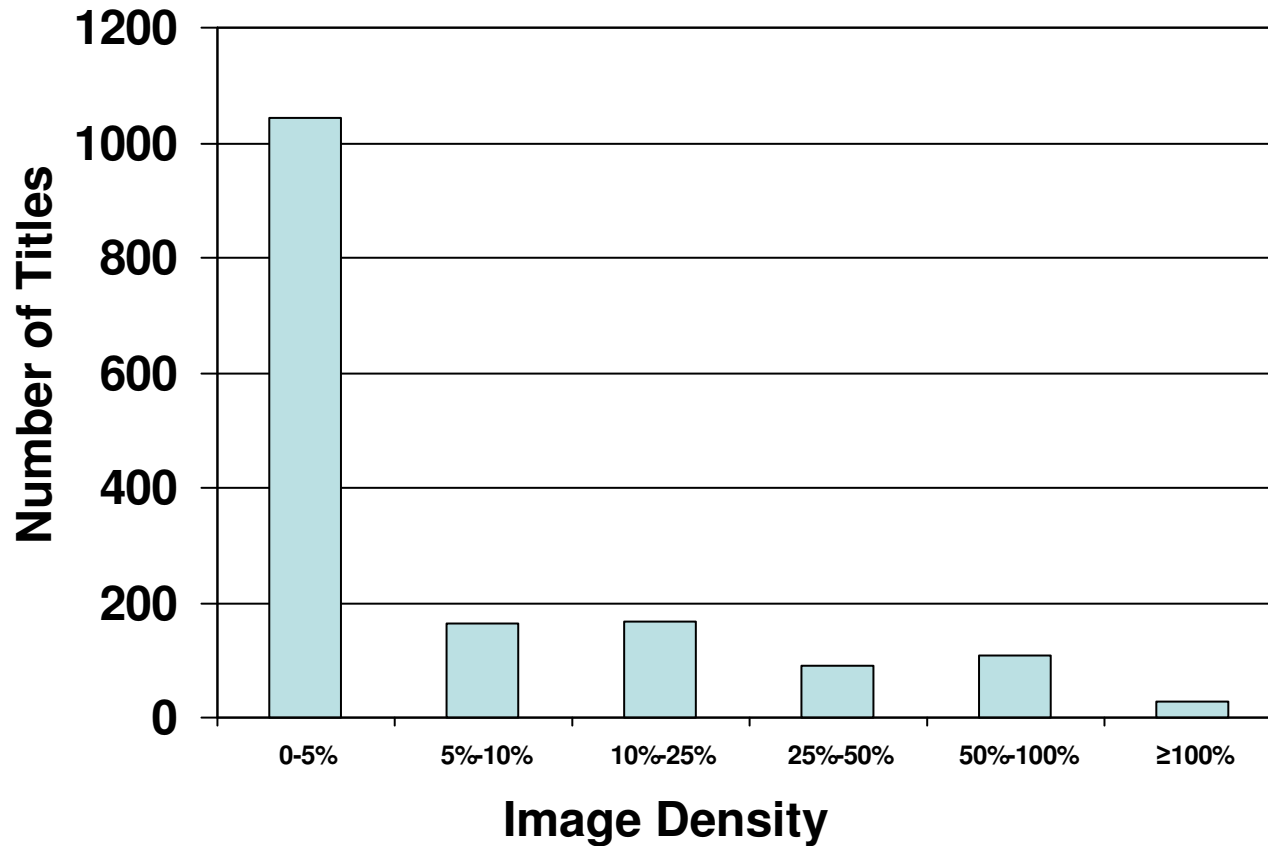
- JSTOR-digitized titles offer an easy opportunity to apply this model in the short term:
 - Widely agreed to be of high quality and reliability, digitally preserved
 - Two page-validated dim to dark archives (at Harvard and UC)
 - Widely held at libraries
 - Easy access to relevant data
- We have developed a tool to provide libraries with additional information about preservation status of JSTOR-digitized titles to support print collections management decisions by identifying titles which:
 - Have relatively few images
 - Are relatively completely held in both Harvard and UC archives

Preservation in JSTOR-Affiliated Archives

- JSTOR has supported the creation of two page-verified dim to dark archives, one at the University of California, another at Harvard
- Two considerations:
 - How much of a given title is held in each archive?
 - How many copies are held between these two archives?

Image Density

Image-to-page ratio offers a proxy for the importance of images in a journal



I T H A K A

JSTOR | PORTICO | ITHAKA S+R

**What to Withdraw:
Decision-Support Tool –
Demo**

Variables

Percent of a title's complete run (measured at the issue level) that must be held in a page-verified dark archive to be considered sufficiently close to "complete" to meet withdrawal thresholds:	90%
Number of page-verified copies of complete run of a title (at the above level of "completeness") required to meet withdrawal thresholds:	2
Threshold for image-to-page ratio (across a title's complete run, as a percentage) below which a title will meet withdrawl thresholds:	5%
Please identify JSTOR collections in which your institution participates:	
Arts & Sciences I	Yes
Arts & Sciences II	Yes
Arts & Sciences III	Yes
Arts & Sciences IV	Yes
Arts & Sciences V	Yes
Arts & Sciences VI	No
Arts & Sciences VII	No
Arts & Sciences VIII	No
Biological Sciences	No
Business	Yes
Business II	No
Ecology & Botany	Yes
Health & General Sciences	No
Ireland	No
Language & Literature	No
Life Sciences	No
Mathematics & Statistics	No
Music	No
How many of the total set of JSTOR-digitized titles meet these criteria?	402
How many of these titles do I currently have access to through JSTOR?	345
(See Titles tab for complete details)	



Titles Listing

ISSN	Title	Passes image criteria?	Passes holdings criteria?	Passes all criteria?	Subscribed?	Actionable?
0001-4273	The Academy of Management Journal	Yes	No	No	Yes	No
0001-4826	The Accounting Review	Yes	Yes	Yes	Yes	Yes
0001-6241	Acta Musicologica	Yes	Yes	Yes	Yes	Yes
0001-6993	Acta Sociologica	Yes	No	No	No	No
0001-8392	Administrative Science Quarterly	Yes	Yes	Yes	Yes	Yes
0001-8678	Advances in Applied Probability	Yes	Yes	Yes	No	No
0001-9720	Africa: Journal of the International African Institute	Yes	Yes	Yes	No	No
0001-9887	Africa Today	Yes	No	No	No	No
0001-9909	African Affairs	Yes	Yes	Yes	Yes	Yes
0001-9933	African Arts	No	No	No	No	No
0001-9992	African Historical Studies	Yes	No	No	Yes	No
0002-0206	African Studies Review	Yes	Yes	Yes	Yes	Yes



Actionable Titles List

ISSN	Title	Passes image criteria?	Passes holdings criteria?	Passes all criteria?	Subscribed?	Actionable?
0001-4826	The Accounting Review	Yes	Yes	Yes	Yes	Yes
0001-6241	Acta Musicologica	Yes	Yes	Yes	Yes	Yes
0001-8392	Administrative Science Quarterly	Yes	Yes	Yes	Yes	Yes
0001-9909	African Affairs	Yes	Yes	Yes	Yes	Yes
0002-0206	African Studies Review	Yes	Yes	Yes	Yes	Yes
0002-7294	American Anthropologist	Yes	Yes	Yes	Yes	Yes
0002-8282	The American Economic Review	Yes	Yes	Yes	Yes	Yes
0002-8762	The American Historical Review	Yes	Yes	Yes	Yes	Yes
0002-919X	The American Journal of Comparative Law	Yes	Yes	Yes	Yes	Yes
0002-9300	The American Journal of International Law	Yes	Yes	Yes	Yes	Yes
0002-9319	The American Journal of Legal History	Yes	Yes	Yes	Yes	Yes
0002-9327	American Journal of Mathematics	Yes	Yes	Yes	Yes	Yes



What This List Is (And Is Not)

- This list can help a library to identify titles that match criteria set by the library, identifying potential easy opportunities for withdrawals
- This list is **not** a picklist for a withdrawal project; any library may appropriately choose to locally maintain any or all of the items this tool identified because of other needs or priorities
- This list can provide one new source of information into a decision-making process; it cannot substitute for that process

Future Opportunities

- To provide libraries with more information to support print collections management choices, we hope to work with partner to expand this tool to cover:
 - a significantly broader range of journal titles
 - a significantly broader set of print repositories,

What to Withdraw: Further Information

- What to Withdraw report and decision-support tool can be found at <http://www.ithaka.org/ithaka-s-r/research/what-to-withdraw>
- Our recent report on the transition facing federal government information can be found at <http://www.ithaka.org/ithaka-s-r/research/documents-for-a-digital-democracy>
- For more information or questions, please contact Roger Schonfeld, rsc@ithaka.org.

I T H A K A

JSTOR | PORTICO | ITHAKA S+R

What to Withdraw: Print Collections Management in the Wake of Digitization

Roger Schonfeld
Ross Housewright