

Considering “Sameness” of Monographic Holdings in Shared Print Retention Decisions

Jennifer Hain Teper

In addition to the pressure of operating in a steady state of insufficient funding, academic libraries face incessant pressure to use space differently. As a result, libraries are aggressively withdrawing materials to relieve cramped shelves and reduce overall collection footprints. Selection for withdrawal may be based on various factors, but of concern is the withdrawal of materials for which copies are currently held in shared print repositories. Recent publications point to the need for thoughtful and strategic evaluation of shared print for quality and completeness, plus the evaluation of copies considered for withdrawal to ensure the perseverance of our print heritage. This study focuses on the comparison of forty-seven monographic titles cataloged as identical items that show broadly varying differences in editions, printings, condition, and preservation and repair. Survey data collected includes information about bibliographic accuracy, printing and binding variances, completeness, physical damage, chemical deterioration, provenance, and presence in the HathiTrust. The results show wide variability in the accuracy of cataloging records, historical use, physical condition of the materials, and the ability for those materials to be successfully digitized in the future. These results are illustrative of the strong potential for variation in “identical” bibliographic holdings among the broader academic library community.

For decades, libraries have made preservation and withdrawal decisions based largely on local information, considering shared or national-level holdings only in reference to identifying scarcely held materials. However, as libraries increasingly accept digitization as a trusted form of access for many titles, and as the demand for library space for user services and other functions increases, approaches to evaluating and prioritizing materials for preservation, print retention, or discard must take a wider perspective.

Currently, many academic and research libraries participate in shared print repositories—where one item serves as a physical copy for many institutions. While models and partnership agreements for print repositories vary, they share the commonality that a given title is selected for retention in agreement with a larger group. That title is retained either at the home institution or in a centralized location so that other institutions may choose to withdraw their copies to gain shelf space. Identification of titles for shared print agreements often focuses on low-use content, materials for which electronic access is available, or

Jennifer Hain Teper (jhain@illinois.edu) is Velde Professor and Head, Preservation services, University of Illinois.

Manuscript submitted August 16, 2017; returned to author for revision October 9, 2017; revised manuscript submitted January 22, 2018; returned to author for addition revision May 25, 2018; revised manuscript submitted July 6, 2018; accepted for publication July 31, 2018.

both. Although the condition of the selected physical item identified as the archived copy may be evaluated to meet minimum guidelines, rarely is it compared to other copies held locally to select the best or most historically accurate copy that, ideally, is undamaged and in an original binding. Additionally, libraries are making local retention and preservation decisions based on low OCLC holdings, assuming that this means scarcely held content. Libraries rely heavily on the accuracy of our shared cataloging records in OCLC's WorldCat. Yet many, including the University of Illinois at Urbana-Champaign, are painfully aware of both the inaccuracy of many of our local records and institutional holdings information and those found in the larger OCLC network. Still, many projects that focus on "last print copy" retention decisions, deaccessioning widely held titles, or making preservation/conservation/reformatting judgments that rely heavily on the accuracy of this data as it is the best data available for such choices.

Currently, decisions made by libraries regarding print retention affect the preservation of physical collections. Many academic libraries are moving towards a future with non-special collections print holdings occupying significantly reduced real estate in patron-focused areas, with access to many of the physical volumes provided through shared print holdings or through retention in remote storage.¹ However, without forethought and collaboration, that future could involve discarding books with potential value to our shared print heritage in lieu of lesser (damaged or incomplete) copies, simply because comparative data on like titles was not available or reviewed. Such "value" could be in the form of variances in imprint or edition, with important signatures or marginalia, original and historically/intellectually valuable bindings, or those that had received costly preservation and conservation treatments to extend their long-term usability (like deacidification). There are many who argue fervently for the value of the book as an object, such as Stauffer through his Book Traces project, a CLIR-funded program that has set out to find and record historical readers' interventions in the University of Virginia Library's circulating collections and around the United States.² It stands to reason that if we withdraw or "deduplicate" a large portion of our print heritage, information will be lost. That information may lie in fine bindings, historic provenance, or important but subtle variance between editions, if not properly cataloged as different editions. Will that loss disservice the scholarly community or the population at large?³

A more tangible argument for why we should concern ourselves with the quality of the materials we are maintaining or withdrawing is as a safeguard against faulty digitization or a researcher's need to reference the original, physical work as it was published. Texts belonging to our cultural canon ought to have reliable copies that serve

the role of "leaf master," to quote Frost, to back up their digitized expressions.³ While the quality of digitized texts is constantly improving, vast numbers of books scanned through large-scale digitization efforts such as the Google Books Project have errors ranging from small to significant. Many of these are unintentional flaws either inherent in the source content used or resulting from the scanning process, while others are intentional decisions, such as cases in which large foldouts are not scanned because the complexity of capturing or compositing large images slows down the scanning process.⁴

Regardless of the motivation, there is a clear reason to consider the quality and completeness of archived copies of printed books and the quality and completeness of the digitized content upon which we are increasingly reliant. However, the definition of what might possibly be viewed as "acceptable" in quality and quantity of copies may differ significantly depending on the decision to retain. For digitization backup, we need to identify and retain copies that are complete, in usable condition, and provide an ample gutter margin should reimaging be required. To guard against the more variable loss of cultural heritage, copies must be assessed for persistent value individually, and an ideal number of archived copies may not be definable. While this study does not argue for either retention strategy, it attempts to bring attention to the potential risks of any pursuit of shared print management and local withdrawal of print holdings.

To better understand and evaluate the perceived risks and variability in US shared print holdings, the author designed a survey to review a sample of circulating monographic titles dating between 1851 and 1922 held in common across the Big Ten Academic Alliance (BTAA). The BTAA is an academic consortium consisting of the University of Illinois, University of Chicago, Indiana University-Bloomington, University of Iowa, University of Maryland, University of Michigan, Michigan State University, University of Minnesota, University of Nebraska, Northwestern University, Ohio State University, Pennsylvania State University, Purdue University, Rutgers University, and University of Wisconsin. The purpose of the survey was to gather data on both physical and bibliographic quality of each university's holdings. Serials were excluded since monographs were believed to display more potential for bibliographic-level cataloging errors, variant editions, and preservation actions. Circulating materials were selected for their greater likelihood to be considered for withdrawal, but also for the stronger likelihood of heavy use and damage due to a longer circulation history. The date range 1851 to 1922 was selected since it is the most common range of holdings available digitally (being in the public domain) and still held in circulating collections (e.g., not yet transferred to special collections). It was anticipated that the sample

results would illustrate the degree of variability in quality of our physical holdings and the dependability of our professional reliance on the accuracy of OCLC records for the given titles.

Literature Review

Since large-scale digitization initiatives such as the Google Books Project and the Internet Archive began scanning large numbers of US libraries’ holdings, there has been concern about the future of print in libraries. Some, such as Grafton, have painted dire futures, while others within the preservation community focused on how widespread access to digital content is changing preservation and conservation selection and priorities, such as Pickwood’s “Library or Museum? The Future of Rare Book Collections and Its Consequences for Conservation and Access” and Conway’s “Preservation in the Age of Google.”⁵

Another area of influence is the idea of “minimum” holdings, or better defining scarcity in holdings for prioritization related to retention and preservation. The keystone of several seminal papers in this area is Yano’s “Optimizing the Number of Copies and Storage Protocols for Print Preservation of Research Journals” concerning the results of a study completed several years earlier in support of research for Ithaka S+R.⁶ Yano was commissioned by Ithaka S+R to produce a statistically valid evaluation and recommendation of the minimum number of copies needed, using different storage and use scenarios, to guarantee the perseverance of a print copy of a journal title held in JSTOR. From this analytical study came Schonfeld and Housewright’s 2009 study “What to Withdraw? Print Collections Management in the Wake of Digitization” and Nadal and Peterson’s “Scarce and Endangered Works: Using Network-Level Holdings Data in Preservation Decision-Making and Stewardship of the Printed Record.”⁷ Both of these frequently referenced studies use Yano’s research to project longevity for titles and use those projections to suggest better withdrawal practices or selection for preservation activities.

The idea of comparing “identical” books was also considered by an Andrew W. Mellon Foundation–funded study at the British Library called “The Identical Book Project” in which four hundred identical book titles in six libraries across the UK were assessed physically and chemically to evaluate paper condition and degradation over time in different locations.⁸ This work, however, primarily focused on paper strength relative to location in the UK, not overall condition of the materials. Stauffer, the faculty lead behind the Book Traces project, recently published another study that discusses the comparison of “identical books.” In his 2016 paper, “*My Old Sweethearts: On Digitization and the Future of the Print Record,*”

Stauffer reviews ten bibliographically identical copies of the 1902 publication *My Old Sweetheart* as a case study of the potential for loss as libraries withdraw individual print holdings.⁹ Stauffer asserts that materials printed between 1830 and 1923 are the most at-risk as they are predominantly out of copyright, in poor condition, and little used. He points the small sample’s variance in bindings, publisher information, text, preliminary text and endleaves, illustrations, and usage marks.

In another area of study, many refer to the need for print retention to serve as backups for poor quality, incomplete, or faulty digital copies. Conway’s “Preserving Imperfection: Assessing the Incidence of Digital Imaging Error in HathiTrust” addresses this concern.¹⁰ Conway reports the results of a study of the image quality of a thousand-item sample of 1.25 million volumes in the HathiTrust consisting of English-language books and serials published before 1923 that were scanned and processed by Google between 2004 and 2010. The results of his study find that there was an average of 2.42 errors per page, though many of these were minor, and 1.5 percent were what Conway classifies as “severe errors” leading to contextual loss of information. However, a much more substantial proportion of “whole volume errors,” such as missing pages, fully obscured pages, or pages out of order, was found. Of the books reviewed, 46.8 percent contained at least one of these types of errors, though not all errors meant loss of content. More importantly, the study examined the relationship between the physical condition of the original source volumes and its impact on the quality of the resulting digital scans. In this part of his study, Conway records the basic statistics on his sample of 860 physical, source volumes reviewed for overall binding integrity, narrow gutters, embrittlement, paper damage, printing errors, and annotations.

There is much research published in the past decade assessing the value of shared print retention and its possible approaches. A few publications stand out as particularly relevant. Kieft has been a key player in many conversations regarding shared print. In his 2010 paper “A Nation-Wide Planning Framework for Large-Scale Collaboration on Legacy Print Monograph Collections,” he and co-author Payne present a summary of what a potential framework for collaborative management and preservation of print monographs might entail and the strengths and weaknesses of such a framework.¹¹ Similarly, Malpas’s 2011 *Cloud-Sourcing Research Collections: Managing Print in the Mass-Digitized Library Environment* laid significant groundwork for the establishment of a more organized and collaborative network of large-scale print and digital repositories for the long-term preservation and access of low-use print books through a focused data analysis of OCLC holdings and the HathiTrust.¹² Although many of the data comparisons between the HathiTrust and academic library

holdings are now outdated, Malpas importantly calls the proposed repository system for print retention a “print preservation repository,” valuing not only a commitment to *retain* but also a commitment to *preserve* shared print holdings.

The CRL has positioned itself as a leader in discussions of shared print management for serial holdings. Their 2015 report *Print Archiving and Shared Print in North America: A Preliminary Analysis and Status Report* is the outgrowth of the findings of a 2015 meeting: “Preserving America’s Print Resources II: A North American Summit.”¹³ Though this study focused on serial holdings, many of the challenges the report addresses hold true for any physical print resource. Information available from current shared print initiatives falls short of the necessary level of detail to support sound risk assessments and decision-making for preservation, retention, and disposition of materials; information regarding the varying commitments of partners in shared print projects is unavailable or vague; and little data is available about the environmental conditions in which libraries store archived materials, thus calling into question if these commitments are simply to “retain” or to “preserve” content. Most recently, a similar call for action toward a more organized, national approach was issued by the Modern Language Association with its 2016 white paper “Concerted Thought, Collaborative Action, and the Future of the Print Record.”¹⁴ The authors argue for the creation of a cohesive system, including both governance and brick-and-mortar structures, using existing high-density book storage facilities and new purpose-built facilities to oversee the management of print collections.

Many of these writings on shared print cite the importance of copy-specific preservation information in the MARC record, most often citing the MARC 21 field 583 Action Note as a possible home for such copy-specific condition or treatment-related information.¹⁵ While discussions about the sharing of preservation information are numerous, little has been published about the use of MARC 583. McCann’s 2013 paper “Conservation Documentation in Research Libraries: Making the Link with MARC Data” presents the results of a survey about how institutions are currently recording preservation actions in MARC 583, most specifically focusing on conservation documentation of special collections materials and how it might be more comprehensively documented.¹⁶

Survey Design and Methodology

The first step in designing the survey was to identify how many monographic titles were held in common across the fifteen consortia members. After running reports against OCLC, the author compiled a list of 251 records identified

as physical monographs in OCLC published between the dates of 1851 and 1922 and held by all consortial institutions. Of the 251 records found, the author selected a random sample of 52 titles from the list using a random number generator, giving a statistical confidence of 90 percent with a margin of error of 10 percent for title-level data. For item-level data interpretation, the total population of 3,765 commonly held individual items (15 copies for each title), and the constituent 780 items requested for review, the author predicted an item-level confidence and tolerance to be 94±3 percent. However, due to several instances where microformats and electronic formats displayed as books in the OCLC report—an actual sample of 47 titles resulted in a slightly broader margin of error of 90±11 percent, and title-level data confidence at 92±3 percent. A full list of the titles and publication information for all titles reviewed is provided in appendix A and an image of the University of Illinois’s copies of the titles can be seen in figure 1, which shows the general age, size, and condition of the titles considered.

The study used interlibrary borrowing services to obtain as many of the titles as possible from the partner institutions. Due to reasons such as non-circulating status, items being checked out, or library renovation projects, not all items could be borrowed during the period in which the research was conducted. Of the possible 705 items, 625 (89 percent) were reviewed. Data collected in the assessment considered cataloging record accuracy, nearness to an “as-published state,” printing variances, completeness, provenance, condition, preservation actions taken, and openly available digital surrogacy. While some records were found to be RDA compliant and some were not, for reference in relation to the cataloging accuracy and completeness review, the purpose of the cataloging record evaluation was not RDA compliance but rather to discern significant enough differences in publisher, date, and/or edition information that a patron or library employee looking at the record alone might reasonably confuse one title for another, or potentially withdraw an item based on an incorrect match. The author photographed all items, both individually and with all copies of a given title for a side-by-side comparison. The full survey tool is available in appendix B.

Survey Data

Various manipulations of the collected data yielded revealing trends. The most useful view of the data is an item-level examination of each data point collected (instances of each in an individual book). Some considerations, such as available digital content, were at a title level. The author attempted to aggregate the data by broad subject areas (as defined by the LC call number classifications on the



Figure 1. Examples of each title examined as part of the survey, as held by the University of Illinois.

items), but in nearly all cases, the titles in a given subject area were small enough in number to make this view of the data unusable. Even in the broadest of classifications, of the total forty-seven titles reviewed, only one was in the subject area of agriculture; seven were in biological sciences; three were in business and economics; one was in geography and Earth sciences; five were in history and auxiliary sciences; twenty were in language, linguistics and literature; one was in library science, generalities and reference; one was in performing arts; two were in philosophy and religion; three were in the physical sciences; and three were in sociology. Data was filtered by institution to determine whether trends could be observed for particular institutional practices. The data presented below draws predominantly from the aggregated total data. Some views of potential trends both by subject area and by institution are presented at the end of this paper.

Cataloging Record Accuracy

There were several instances of miscataloged items that were linked to the incorrect OCLC number. In all cases these were due to later or variant editions, and did not include instances of potential printing variances over later reprints of the same edition as this information was collected separately. Overall, 3.4 percent of items had some variance in publisher name, place of publication, or copyright dates. Later publication dates without changes in publisher, place of publication, or copyright were considered later printings of the same edition and therefore not a miscataloged different edition. Eight percent of the total books reviewed were later reprints of the original publication, which, while correctly sharing the same OCLC number and record, are still potential points of printing variance. While properly cataloged, these often displayed minor printing variances over the subsequent printings, including the presence of

publisher advertisements, prologues, or other differences largely in the books' front and end matter. An additional 1.0 percent of the items were preservation photocopies of the original text with varying degrees of reproduction quality, which should have been cataloged as new editions, resulting in a total finding that 4.4 percent of the items surveyed should have been cataloged using different OCLC records than those on which they were found. Another 2.7 percent displayed variances that would often not be considered different editions, bibliographically, such as "library editions" or "handmade editions" where copies were on higher-quality paper and often signed and numbered. Such physical variances, though disparate from variation in the bibliographic qualities of a given item, are nonetheless of interest to those who value books as objects and find meaning in variation between items' material components.

Nearness to "As Published" State

Just over half (56.8 percent) of the total volumes retained their original covers (this includes repaired covers with replaced spines), while 43.2 percent were rebound in their entirety. Of those rebound, 4.3 percent were issued as paperbacks with their original covers bound in with the text or mounted to the cover of the new hardback binding. Of the 43.2 percent lacking original bindings, 40.0 percent had buckram bindings and 2.2 percent were in older-style library bindings, half-bound in leather and marbled paper. The remaining 1.0 percent were rebound in a conservation lab, which is discussed in the section titled "Preservation Actions."

Printing and Binding Variance

Four percent of the total (or 7.0 percent of those with original covers) had variant covers. While some of these variances correlated with the library or handmade editions previously noted, others had no other distinguishable variance from the other pieces for that title except book cloth color or material. See figure 2 for an example of such variance.

Provenance

Eighteen percent of the items reviewed showed some evidence of provenance, either through a bookplate stating that an item was part of a particular collection or a gift of a certain person or was signed or otherwise inscribed by an identifiable previous owner. In most cases the provenancial information was relatively brief, and a few items included tipped-in letters or long inscriptions by the author



Figure 2. An example of binding variance. While the front-most book is a different edition, the rear two are identical except for their covers.

(0.8 percent of the total, or 3.6 percent of those showing provenance).

Completeness

The majority of the materials reviewed (95.7 percent) were complete, while the remaining 4.3 percent were missing some form of content. The most common missing content concerned 3.2 percent of materials that were missing plates or text within the body of the work, followed by 2.9 percent missing half title pages, and 1.1 percent missing title pages. In total, 1.1 percent of materials were missing more than one defined category of content. Not counted as missing content, but noted nonetheless, 8.6 percent of items were originally published with advertisements at the rear of the publication, which were lost or not included when an item was rebound.

Condition

A great deal of information was collected on the condition of materials. While not as important as completeness when selecting for print retention, it is common sense that materials in better condition are preferable for long-term print retention, especially if that damage hinders the readability or future digitization potential of the item at hand.

The openness and width of gutter margin of an item was reviewed and measured to ensure that future digitization efforts of a specific item would not lose text on the gutter margin nor require damaging disbinding of the bound artifact. Reduced gutter margin (e.g., text running far down

Figure 3. Distribution of Severity of Damage Noted

Cover Damage	Cover to Text Attachment	Visible Embrittlement	Pages Torn	Underlining/ Highlighting/ Marginalia	Losses	Page Detachment	Water Damaged/ Stained/ Foxed	Broken Text Block
None/wear Only	Sound	None	None	None	None	None	None	None
67.7%	72.8%	77.6%	78.4%	59.7%	93.9%	89.8%	93.8%	96.7%
Slight	Weak	Slight	Slight	Slight	Slight	Slight	Slight	Slight
23.8%	18.2%	16.5%	16.6%	18.9%	5.9%	9.0%	3.2%	1.6%
Moderate	Part Detached	Moderate	Moderate	Moderate	Moderate	Moderate	Moderate	Moderate
7.5%	5.1%	5.1%	3.7%	7.40%	0.0%	1.1%	1.8%	0.3%
Severe	Detached	Severe	Severe	Severe	Severe	Severe	Severe	Severe
1.0%	5.3%	1.6%	1.8%	12.0%	1.6%	1.8%	1.2%	1.4%

towards the spine) is often due to rebinding especially with the practice of oversewing, which was a popular library binding practice through the 1980s in which pages were sewn together through the sides of the gutter instead of through the fold. Of the entire collection reviewed, 63.8 percent of the materials maintained their original sew-through-the-fold page attachment method, while 26.9 percent were oversewn. An additional 0.6 percent were found to be double-fan adhesive bound (a later library binding practice of page attachment), and 0.8 percent were found side sewn. However, despite a substantial number being oversewn or side sewn (a total 27.7 percent), only 7.4 percent of the total (or 26.7 percent of those with restrictive binding structures) had margins that were too narrow to digitize without likely text image loss (measured at a visible gutter of less than ¼ inch if found in any part of the text).

Physical damage to the volumes was also evaluated, including the condition of the covers, cover-to-text attachment, and damage to the text block. While many items showed evidence of their age through wear (scuffs, scratches, and minor corner or headcap strain), 32.3 percent showed damage (defined as breaks or tears) to their covers, with most being only slight damage (see figure 3 and appendix B for a full description of all assessment questions and definitions of what was considered “slight,” “moderate,” and “severe” damage).

The majority of materials exhibited sound cover-to-text attachment, yet 10.4 percent were either partially or completely detached. This is significant because, for large-scale scanning workflows, detached covers can seriously impede the ability to scan an object as it makes the book more challenging to secure to the cradle for imaging. Damage to the text blocks was evaluated on various considerations, including paper embrittlement, tears and losses on pages, page detachment, and text blocks split into two or more

pieces. Perhaps the most significant of these in considering future usability is embrittlement. Utility of materials is dramatically decreased as the flexibility and strength of the pages decreases. The resulting fractures and potential losses of text result in difficult and/or possibly incomplete digital capture. There are many ways to test paper for degrees of embrittlement—the most common is a “double fold test,” in which a corner of a page is folded back and then forward, testing the durability of paper over repeated folds. Since visibly destructive testing on actively circulating books held by other institutions was not deemed acceptable as part of this study, and other options for analytically testing paper strength were not available, embrittlement data was collected only based on visual observations of damage (breaking edges or fracturing paper off existing sewing structures). If destructive testing such as a double fold test had been completed, a much higher percentage than the 22.4 percent found to show signs of embrittlement would likely have been noted. However, with nearly one quarter of those items reviewed noted as being exceptionally brittle, this percentage is substantial in its own right as these items are exceptionally brittle and already actively fracturing. Page damage, as evidenced by tears and breaks in the paper, is often closely related to the strength of the paper but can also result from heavy use or abuse. Therefore, it is not surprising that 22.2 percent of books reviewed had some tears (tears were not counted unless they ran into the text or measured at least an inch long), while 77.8 percent of items had no torn pages across the publication. Few losses (absence of a portion of a page) were noted, with 7.5 percent having losses of any type, with the vast majority being minor amounts of paper loss, resulting in little to no text loss. Page detachment and broken text blocks (where the sewing has broken midway through a text block, rendering it in two pieces) were also reasonably rare, with 10.1 percent

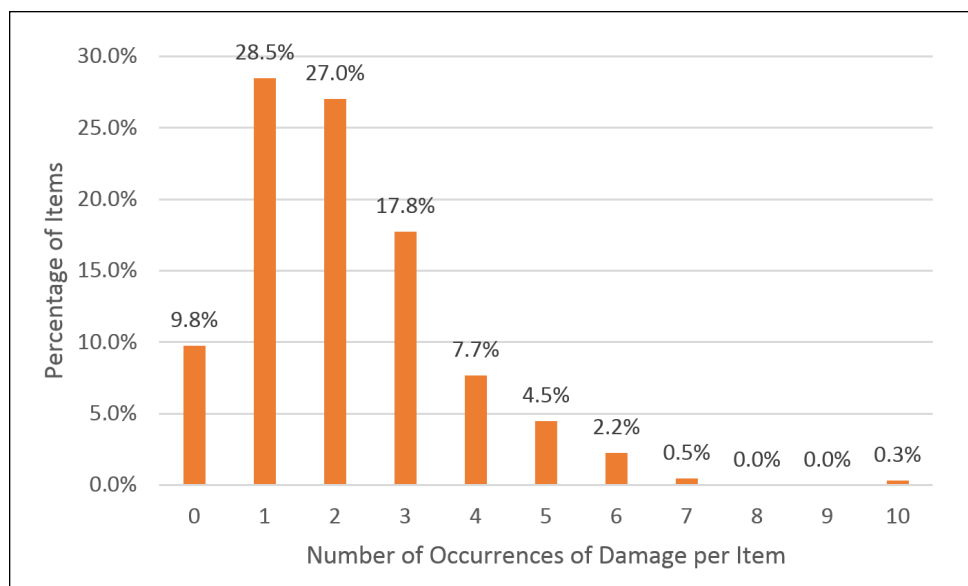


Figure 4. Instances of Multiple Occurrences of Damage on Individual Items.

of materials having some number of detached pages and 3.0 percent having broken text blocks.

Visual distractions, such as writing or staining on the pages, were considered as they can interfere with, or even obscure, the text. Such distractions were the most common type of damage observed. More than 40 percent (40.2 percent) of the materials had some level of writing on them (not inclusive of provenancial markings). Of those with markings on them, 18.9 percent were slight, 7.4 percent were moderate, and 12.0 percent were severe (with marks covering or obscuring text on at least ten or more pages). On a positive note, the majority of the marks were in pencil, which could be fully or partly removed at some future date. Water damage and staining occurred far less frequently than markings, with only 6.2 percent of materials exhibiting notable water damage or staining.

While all individual damage categories provide valuable information about the potential usability of the sampled titles, the mode of data presentation necessarily isolates each form of damage from the other. A reasonable assertion is that one occurrence of damage is often not independent of other types of damage. For instance, high use is likely to cause not only a greater likelihood of underlining, but also more tears, stains, and cover damage. Poor quality paper that has become brittle is likely to directly correlate with a much higher likelihood of tears, losses, and detached pages. Therefore, to better understand whether each instance of damage was isolated or, more likely, occurred in aggregate, each item was individually evaluated to record the total number of damage types observed per piece. Through this analysis, a relatively small percentage of

items were completely undamaged (9.8 percent). The majority (55.5 percent) of items showed only one (28.5 percent) or two (27.0 percent) types of observed damage per item. Occurrences of three damage types were noted in only 17.8 percent of items and significantly less for four (7.7 percent), five (4.5 percent), and six (2.2 percent) types of damage occurring within one item. Less than 1 percent of items observed displayed multiple damages of seven types or more (see figure 4). This means that, while 60 percent of the items surveyed showed instances of more than one damage type, only 15.2 percent of items were recorded in four or more damage categories, indicating that, while multiple instances of damage per book are common, severely damaged books with many types of damage were significantly less common and few books were in what is professionally called “terrible shape.” This observation indicates that, while the majority of items (nearly 85 percent) are either unbroken or show only a few categories of damage (with many of these related to paper quality), a significant proportion are severely damaged and would be poor choices as copies of record in a shared print repository environment.

Preservation Actions

Defining what was considered a “preservation action” was challenging since what was accepted as common preservation treatment forty years ago may not currently be considered acceptable preservation practice. The author decided to consider any effort to repair an item, whether with pressure-sensitive tape or through a well-performed modern conservation treatment, as a preservation action. In total, 18.9 percent of materials had received some sort of preservation action, the most common (8.6 percent) being internal hinge reinforcement or repair either through the replacement of endsheets or the addition of a reinforcing paper or tape layer. Paper repairs were a close second in frequency, with 8.0 percent of materials showing some sort of paper repair, most often with some sort of pressure sensitive tape. Another 6.2 percent of materials had received spine repair (rebacking) either independent of, or in concert with, internal hinge reinforcement repair. Enclosures were

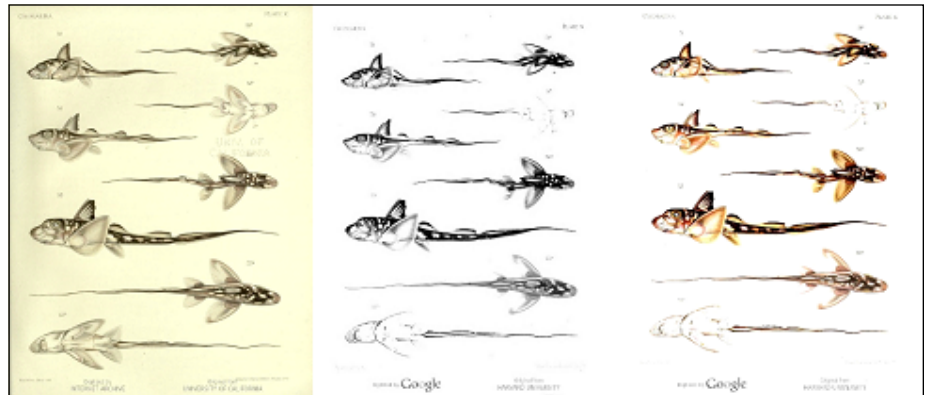
Figure 5. Distribution of Preservation Actions

Paper Repaired with Tape (various types)	Internal Hinge Reinforcement/ Reattachment	Rebacked	New Case	Book Tape	Box	Envelope	Shrink Wrapped	Other Enclosure	Deacidified
8.0%	8.6%	6.2%	1.0%	3.2%	3.4%	1.1%	0.6%	2.2%	0.3%

relatively common, with a total of 7.3 percent having some sort of protective enclosure, though the items held in these enclosures were frequently in poor condition and unrepairable due to severely embrittled paper (see figure 5 for a summary of all preservation actions observed).

Digital Surrogacy

Lastly, the availability of digital surrogates in the HathiTrust was investigated for each title. Whereas the availability of digital content does not likely have a direct influence on the condition of the items surveyed given their age and relatively recent digitization, choices about whether to maintain a print item, especially if it is damaged, may be driven by availability of a reliable (i.e., in a trustworthy digital repository) and complete digital surrogate. Factors considered when evaluating the digital surrogates included whether the surrogate was captured in color or in black-and-white and how this related to the accurate representation of the original publication; whether the digital image was missing content when evaluated against the physical object; and whether variant editions were digitized and tagged incorrectly as the edition being evaluated. A total of 231 digital files were found in the HathiTrust when the forty-seven titles were searched for by OCLC number, producing an average of just under five (4.91) available files per title.¹⁷ From this, 24.7 percent of the digital files were in color or grayscale and 75.3 percent were bitonal (black-and-white). The high proportion of bitonal files is a direct result of the relatively high proportion of Google Books’ project output in the HathiTrust, which has largely produced bitonal images.¹⁸ If considered on a title-by-title basis, ten titles (23.3 percent) were only available as bitonal images. Fifteen (31.9 percent) of the forty-seven titles contained significant fine detail or colored image content that is compromised in a bitonal scan. However, just two of these titles were only available as bitonal files. For examples of image quality loss due to bitonal imaging, see figure 6. The presence

**Figure 6.** Comparison of various color and bitonal images on the digital copies of *Chimeroid Fishes*.

of foldouts was also noted in three titles (6.4 percent). In observing available digital content for those three, two titles had four distinct digital copies available in HathiTrust and one had six (fourteen copies total). One title had no available digital copies of the foldouts, while the others had either two of four or two of six with the foldouts included, for a total of only 28.6 percent of digital copies including foldouts.

Data Interpretation

The author found that cataloging errors were less common than anticipated. The 3.4 percent of errors found were all due to variant editions being cataloged using the wrong OCLC record. An additional 2.7 percent had either “Library” or “Deluxe” editions published by the same publisher in the same year, which would not always be noted as a different edition, yet were physically different from other copies with the same OCLC number. This means that, of the sample observed, approximately 6.1 percent of the books reviewed were variant from the standardly held title sharing that OCLC number, though the intellectual content of these variances may not be significantly different. Although this is accurate cataloging, one cannot assume that all books sharing the same OCLC number are physically identical copies.

Figure 7. Comparison of Paul Conway's Physical Condition Findings in His 2013 *Preserving Imperfection: Assessing Incidence of Digital Imaging Error in HathiTrust* to Condition Findings within this Study

Data Point	Conway Study	Current Study	Notes on Difference
Binding Condition			
<i>Sound</i>	80.5%	72.8%	
<i>Loose</i>	13.8%	18.2%	
<i>Not intact</i>	5.0%	10.4%	
<i>Missing</i>	0.7%	0.0%	
Gutter Margin			Measured for legibility from margin in current study inclusive of curvature of page. Measured at 1 cm. from gutter in Conway
<i>Fine</i>	74.9%	92.6%	
<i>Narrow</i>	25.1%	7.4%	
Text Block			
<i>Intact</i>	80.2%	83.3%	
<i>Pages missing</i>	1.0%	3.2%	
<i>Pages loose</i>	10.8%	10.2%	
<i>Broken</i>	5.8%	3.3%	
Embrittlement			
Not brittle	45.3%	87.6%	Measured by visual observation only in current study, and by destructive double fold tests in Conway
<i>Brittle</i>	54.6%	22.4%	
Page Damage			
<i>Undamaged</i>	89.4%	78.4%	
<i>Damaged</i>	10.6%	21.6%	
Annotations			
<i>None</i>	96.4%	59.7%	
<i>Some</i>	3.6%	40.3%	

Only 56.8 percent of the books reviewed maintained their original bindings. While many of the bindings were not overly decorative, some were highly embellished or illustrated, and a small percent (4.0 percent) had variant cover designs whose existence was not evident except when compared side-by-side as illustrated in figure 2. There is a loss of originality in the objects themselves by having the items rebound. This loss may not be relevant to future users mainly interested in the book's intellectual content, but to those studying the history of publishing and readership, the use of cover illustrations and variant covers for marketing is significant. Rebinding imposes another layer of risk by altering the original page attachment. All instances of oversewing (27.0 percent of pieces reviewed) occurred in rebound books. This new sewing structure dramatically decreased the visible inner margin and functional openability of books, leading to a stronger likelihood of problematic image capture if those copies are used for future digitization, and higher risk of text loss if the paper is or will become brittle.

The historical value of observable physical evidence of ownership or provenance is often debatable, but in a rare few cases, these markings hold significant and undeniable historical value. Whereas 17.8 percent of those items reviewed had some sort of marking indicating previous ownership, only 0.8 percent of items claimed evidence of any historical significance as subjectively deemed relevant by the author. These cases were comprised entirely of letters or inscriptions from the author themselves.

The completeness and condition results collected in this study were relatively consistent with the similar condition data collected in Conway's *Preserving Imperfection*, which sampled a combination of serial and monographic titles of approximately the same publication date range digitized through Google. Comparisons of Conway's data to the data collected in this study are provided in figure 7.

Overall, the data collected in this study showed a slightly greater likelihood for damage than the items Conway reviewed. There are two significant differences between the populations in the two studies. The first is

that Conway's study included both monographs and serials. The second difference is that the titles reviewed in this study were *all* held by BTAA Libraries and were therefore presumably a widely held title. While extensive holdings do not necessarily correlate directly to use, the fact that a title was widely purchased and retained by a large number of libraries indicates a broader interest in the title compared to the more scattered and sometimes esoteric titles included in Conway's study as selected by the Google digitization program, and therefore potentially higher use. If this correlation is accurate, the higher observed rates of binding damage, paper damage, and annotations are symptomatic of higher levels of use over time. Until more research is done on the relationship between widely held items and the frequency of their individual use, the supposition that such use correlates directly to potential damage is merely a hypothesis. The most significant difference in the overall populations of various types of damage observed was in the embrittlement rate. As noted earlier, this study measured embrittlement through visual observations only, such as repeated edge tears, losses, and fractures along the gutter margin. Comparatively, Conway's study performed the more destructive double fold tests, observing how many folds the paper would withstand before fracturing. Had similar tests been performed on the sample observed for this study, it is probable that the embrittlement rate would have been much closer to Conway's observed 54.6 percent than this study's 22.4 percent. In either case, embrittlement of the paper of pre-1923 publications on wood pulp paper is a considerable concern. Even if the lower 22.4 percent is considered more accurate, the likelihood of current or future loss of textual content and significant difficulty in future image capture is of considerable concern for nearly one quarter of the texts reviewed.

The observation of instances of items showing multiple occurrences of damage as opposed to isolated single instances of damage revealed that 15.2 percent of the items reviewed had four or more types of observable damage occurring in one item. This is a relatively high rate of significant damage and is likely corollary to the proposed higher-than-average use of these items. While the use data collected from this survey was inconclusive, other data observed supports this assumption, such as the rate of preservation action. At most institutions, preservation treatment is driven by use and the 18.9 percent of materials observed that sustained some sort of preservation actions is, at least anecdotally, higher than anticipated in a more randomized sample.¹⁹ However, no recent studies of preservation or repair in general collections could be found to support this assertion.

Unfortunately, the sample was too small to extract any meaningful data regarding trends by subject area. See figure 8 for the dispersal of sample titles across broad subject areas. Some possible trends appeared through this

Figure 8. Dispersal of Titles by Broad Subject Classifications

Agriculture	1
Biological Sciences	7
Business & Economics	3
Geography & Earth Sciences	1
History & Auxiliary Sciences	5
Language, Linguistics, and Literature	20
Library Science, Generalities & Reference	1
Performing Arts	1
Philosophy and Religion	2
Physical Sciences	3
Sociology	3

attempted analysis that may be worth further investigation. Since the number of items observed in individual subject areas was too small for analysis, it is possible to group together the humanities and arts-related topics (language, linguistics and literature, performing arts, philosophy, and religion) against all other subject areas for a very base-level comparison. This rather blunt tool reveals some interesting data. Of the 4.4 percent of miscataloged items, including preservation photocopies, nearly all of those (98 percent) were arts and humanities titles. Occurrences of damage or incomplete texts, though slightly higher in the arts and humanities, was not significantly higher than those observed in the sciences. Items in the sciences were 5 percent more likely to retain their original cover, while items in the arts and humanities (directly related to the stronger likelihood of having been rebound) were 5 percent more likely to have a tight gutter margin. Additionally, items in the arts and humanities were 7 percent more likely to have torn pages and 9 percent more likely to have some level of annotations or markings on the pages.

While perhaps of more interest to the individual participating institutions, aggregation of the data by institution showed a potential for certain trends by institution. To conclusively state this, a larger sample is needed from each institutional collection, as the sample size for this study is too small to conclusively show trends for the larger collections. The data summarized in figure 9 shows a wide distribution of occurrences of damage, preservation actions, and "as published" state. This type of profiling, using a broader sample, would be useful when considering cooperative shared print planning, to better strategize for selection of collections most likely to be intact and in good condition if the time-consuming item-level review of materials is not to be undertaken.

Lastly, the data collected may shed light on a very current question in print retention planning: How many archived copies are enough? Again, the sample is too small

Figure 9. Summary of Data Collected by Institution

Institution	Instances of "As Published" State	Instances of Noted Damage	Preservation Actions Noted	Items Reviewed
1	82 (med high)	156 (high)	45 (med high)	46
2	71 (med low)	125 (med high)	37 (medium)	40
3	81 (med high)	125 (med high)	31 (med low)	42
4	47 (low)	68 (low)	35 (medium)	29
5	75 (medium)	109 (medium)	33 (med low)	42
6	71 (med low)	103 (med low)	27 (low)	39
7	66 (med low)	65 (low)	35 (medium)	43
8	68 (med low)	114 (medium)	34 (med low)	44
9	76 (medium)	87 (med low)	26 (low)	40
10	90 (high)	73 (low)	40 (med high)	43
11	83 (med high)	120 (med high)	42 (med high)	44
12	84 (med high)	127 (med high)	51 (high)	46
13	69 (med low)	70 (low)	45 (med high)	39
14	84 (med high)	109 (medium)	26 (low)	44
15	65 (med low)	104 (med low)	48 (med high)	45

Figure 10. Summary of the Probability of Archiving a "Good" Copy if only One Copy of the Title is Randomly Selected to be Archived

Probability Range (chance of archiving a "good" copy)	# Titles in That Probability Range	% of Titles in That Probability Range
0% chance	4	9%
1–10% chance	2	4%
11–20% chance	6	13%
21–30% chance	4	9%
31–40% chance	7	15%
41–50% chance	3	6%
51–60% chance	10	21%
61–70% chance	2	5%
71–80% chance	5	3%
81–90% chance	3	1%
91–100% chance	1	2%

to draw statistically valid conclusions, but it is apparent that there are some trends that point to a need for further study. To do this, the author calculated the probability of randomly archiving a "good" condition copy based on the condition rankings collected through the survey sample. The probability of randomly selecting a "good" copy from the total number of copies for each title was determined using the following calculations: If one copy is selected, the probability of randomly placing a good copy into an archive is the total number of good copies divided by the total

number of copies, or $P = G/T$, where G equals the number of good copies found for each title surveyed, and T equals the total number of books available for that title. This same probability can also be expressed as 1 (being 100 percent probability) minus the probability that all titles selected are "not good" by changing the equation to $P = 1 - ((T-G)/T)$. For the title *A Bibliography of Samuel Taylor Coleridge* from 1903, for instance, eight copies of the fifteen available were in good condition, and the probability of randomly selecting a good copy is $P = 1 - ((15-8)/15)$ or 53 percent probability of randomly selecting a copy in "good" condition for this title. Again, further study is required before this tool could be reliably applied in real-world selection scenarios.

To extend this to anticipate the probability if two or more copies are archived, the calculation changes to $P = 1 - (((T-G) * ((T-G)-1)) / (T * (T-1)))$ if two copies are archived, and $P = 1 - (((T-G) * ((T-G)-1) * ((T-G)-2)) / (T * (T-1) * (T-2)))$ for three copies archived, etc. Again, for *A Bibliography of Samuel Taylor Coleridge*, the probability of archiving a good copy if two copies are randomly selected increases to 80 percent, and if three copies are archived, rises to 92 percent. These calculations were done for each title, calculating the probability of archiving a good copy if one through ten copies were archived. This data, alone, however, shows only title-level probability. But, if considered in aggregate at the number of times all titles showed a certain probability of archiving a "good" copy, we can infer a few trends. For instance, by examining the model of "one copy archived" across all titles, the following is apparent (see figure 10): Showing that the bulk of the titles have a 51 to 60 percent chance of archiving a "good" copy when only one copy is archived, while only 6 percent of the titles have a probability of 71 percent or higher in randomly archiving that copy.

Assuming a desired confidence of at least 71 percent, looking at all models simultaneously (see figure 11), one can extrapolate that the probability of getting a "good" copy increases steadily until five copies are archived and plateaus between 81 to 87 percent of the titles being in that confidence range regardless of how many more copies are archived (with eighty-seven as the maximum in this case because four titles lacked good copies and, mathematically, could not generate a good copy no matter how many copies were archived). However, a significant jump

in probability occurs when three copies are archived. That same jump occurs if the confidence level is raised to 81 percent or higher, but moves to four copies archived if 91 percent confidence or higher is desired.

What this cursory analysis shows is that, at least in a limited sample, the probability of archiving a copy in good condition through random selection increases dramatically with the number of copies archived, possibly as low as three copies. Further research in this area might mitigate some of the question of the value of time-consuming item-level condition review by considering an ideal number of duplicate copies in shared print repositories, statistically reducing the risk of poor-quality copies.

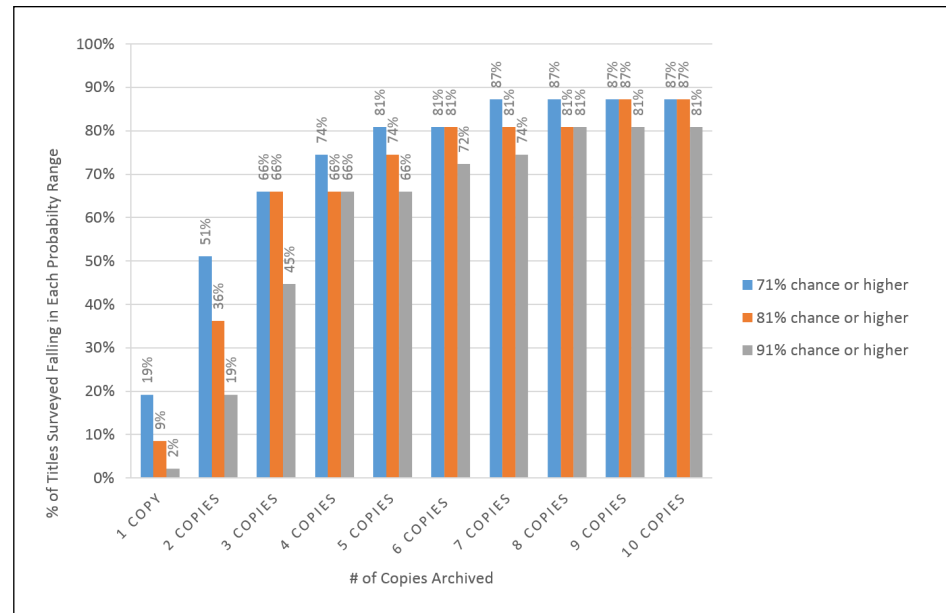


Figure 11. Probability of archiving a “good” copy for all titles with varying numbers of copies archived.

Conclusion

The data collected and analyzed shows that for the types of items reviewed—widely held, pre-1923 monographs—there were several trends that should cause concern for those planning the withdrawal of widely held monographic titles, or selecting individual copies of such items for shared print programs. The most important identified trends include:

- A relatively small but significant likelihood (3.4 percent) of miscataloged editions (especially in the arts and humanities)
- A relatively small but significant likelihood (4.0 percent) of binding variances within a single edition
- A very high occurrence (91 percent) of damage of some type and significant risk (14.4 percent) of more than three instances of damage being found in one title, which represents reduced usability
- A significant likelihood (43.2 percent) of items lacking original bindings, meaning loss of authenticity of the original, as published work
- A relatively small, but significant likelihood (4.3 percent) of items missing content, typically within the text or plates

As institutions undertake shared print projects, resulting in potential for large-scale withdrawal of titles now held by those projects, the data above stresses the risks that libraries are currently taking. By making withdrawal

decisions without item-level review of titles (or incorporating item-level information from shared MARC fields), we are collectively establishing an insecure foundation on which our shared print heritage is being built. The author recognizes that item-level review is logistically impossible in many of these projects; however, this research strongly indicates that further inquiry into the number of copies that must be retained in order to statistically avoid the risk of such losses must be conducted.

Additionally, this research illuminates other areas of potential future research. A comparative study of “unique” items—unique copies as identified through OCLC records—would further expose the potential risks of reliance on OCLC records to denote scarcity or duplicity across institutional holdings. Further research into trends in condition and completeness by subject area could help to focus on subject areas that are prone to miscataloging, damage, or incompleteness, thus targeting limited resources on those collections most likely to be at risk. Lastly, this study shows the potential for strong institutional (or perhaps consortial) trends in condition and preservation action. If a larger-scale research project to review trends in condition and completeness across many institutions were undertaken, data may show certain types of institutions or regions to be more likely than others to possess copies suitable for shared print retention selection—and it is possible that those institutions are not currently contributing copies into such repositories or retention agreements.

References and Notes

1. James Grossman and Geneva Henry, “. . . on the Future of the Print Record,” *The Future of the Print Record: A Multi-Organizational Working Group* (blog), Modern Language Association, January 6–9, 2015, and November 14, 2014, <https://printrecord.mla.hcommons.org/>.
2. For more on the Book Traces project, see www.booktraces.org/.
3. Oya Rieger, *Preservation in the Age of Large-Scale Digitization: A White Paper*, CLIR Publication 141 (Washington, DC: Council on Library and Information Resources, 2008), accessed January 3, 2018, www.clir.org/pubs/reports/pub141.
4. Paul Conway, “Preserving Imperfection: Assessing the Incidence of Digital Imaging Error in HathiTrust,” *Digital Technology & Culture* 42, no. 1 (2013): 17–30, <https://doi.org/10.1515/pdte-2013-0003>.
5. Anthony Grafton, “Apocalypse in the Stacks? The Research Library in the Age of Google,” *Dædalus*, 138, no. 1 (2009): 87–98; Nicholas Pickwoad, “Library or Museum? The Future of Rare Book Collections and Its Consequences for Conservation and Access” in *New Approaches to Book and Paper Conservation-Restoration*, ed. Patricia Engel et al. (Austria: Verlag Berger, Horn/Wien, 2011), 113–30; Paul Conway, “Preservation in the Age of Google,” *Library Quarterly* 80, no. 1 (2010): 61–79.
6. Candace Yano et al., “Optimising the Number of Copies and Storage Protocols for Print Preservation of Research Journals,” *International Journal of Production Research* 51, nos. 23–24 (2013): 7456–69.
7. Roger C. Schonfeld and Ross Housewright, “What to Withdraw? Print Collections Management in the Wake of Digitization. (Ithaca S+R, 2009), <https://doi.org/10.18665/sr.22357>; Jacob Nadal, Annie Peterson, and Dawn Aveline, “Scarce and Endangered Works: Using Network-Level Holdings Data in Preservation Decision-Making and Stewardship of the Printed Record,” accessed January 3, 2018, www.jacobnadal.com/wpcontent/uploads/2011/05/ScarceAndEndangeredWorksv7.pdf.
8. Barry Knight and Velson Horie, “The Identical Book Project,” *International Preservation News* 42 (2007): 18–21.
9. Andrew Stauffer, “*My Old Sweethearts*: On Digitization and the Future of the Print Record,” in *Debates in the Digital Humanities* (2016): 218–29, accessed January 3, 2018, <http://dhdebates.gc.cuny.edu/debates/text/70>.
10. Conway, “Preserving Imperfection.”
11. Robert H. Kieft and Lizanne Payne, “A Nation-Wide Planning Framework for Large-Scale Collaboration on Legacy Print Monograph Collections,” *Collaborative Librarianship* 2, no. 4 (2010): 229–33.
12. Constance Malpas, *Cloud-Sourcing Research Collections: Managing Print in the Mass-Digitized Library Environment* (Dublin, OH: OCLC Research, 2011), accessed January 3, 2018, www.oclc.org/research/publications/library/2011/2011-01.pdf.
13. *Print Archiving and Shared Print in North America: A Preliminary Analysis and Status Report* (Chicago: Center for Research Libraries, 2015), accessed January 3, 2018, www.crl.edu/sites/default/files/attachments/events/PAPR_summit_preliminary_analysis2_revised.pdf.
14. Future of the Print Record Working Group, *Concerted Thought, Collaborative Action, and the Future of the Print Record: A White Paper* (New York: Modern Language Association, 2016), accessed January 3, 2018, <https://printrecord.mla.hcommons.org/>.
15. Jim Hinz and Babette Gehrich, “Documenting Library Conservation Treatments: Using the 583 Action Note Field in the MARC Record,” *The Book and Paper Group Annual* 25 (2006): 59–64; Library of Congress Network Development and MARC Standards Office, *Preservation & Digitization Actions: Terminology for MARC 21 Field 583* (Washington DC: Library of Congress, 2004), accessed January 3, 2018, www.loc.gov/marc/bibliographic/pda.pdf.
16. Laura McCann, “Conservation Documentation in Research Libraries: Making the Link with MARC Data,” *Library Resources & Technical Services* 57, no. 1 (2013): 30–50.
17. OCLC number is one of the metadata requirements for ingestion into the HathiTrust, though practical experience shows that many files in the HathiTrust do not meet all the metadata requirements. It is therefore likely that more digital files than those found exist for the titles searched. However, to ensure that only the appropriate edition of each title was assessed, title/author searches were not performed as there was a greater likelihood of accidentally including variant editions that were difficult to distinguish from the item being evaluated and were not clearly identified as the appropriate edition.
18. Approximately one-third of the corpus of public domain content in the HathiTrust is derived from Google Books content, see www.hathitrust.org/datasets.
19. Jennifer Hain Teper, “Selection for Preservation: A Survey of Current Practices in the Field of Preservation,” *Library Resources & Technical Services* 58, no. 4 (2015): 220–32.

Appendix A. Monographic Titles Selected for Assessment, Listed by Date of Publication

- Twenty years of Congress: from Lincoln to Garfield; with a review of the events which led to the political revolution of 1860.* James Gillespie Blaine. Norwich, CT: Henry Bill. 1884. OCLC # 20498700.
- Walter of Henley's Husbandry, together with an anonymous husbandry, Seneschaucie, and Robert Grosseteste's Rules.* Walter de Henley; Elizabeth Lamond, W Cunningham, Robert Grosseteste. London; New York: Longmans, Green, and Co. 1890. OCLC # 02146299.
- A popular treatise on the physiology of plants for the use of gardeners or for students of horticulture and of agriculture.* Paul Sorauer. London, New York: Longmans, Green & Co. 1895. OCLC # 0151333.
- The fire of love, and the mending of life; or, The rule of living. The first English in 1435, from the De incendio amoris, the second in 1434, from the De emendacione vitæ of Richard Rolle, hermit of Hampole.* Richard Rolle, Richard Misyn, Rev. Ralph Harvey. London: Published for the Early English Text Society by K. Paul, Trench, Trubner & Co. 1896. OCLC # 00374731.
- Histoire de la langue et de la littérature française des origines à 1900.* L. Petit de Julleville, Paris: A. Colin & cie, 1896–99. OCLC # 00930890.
- The Works of John Ruskin.* John Ruskin (Edward Tyas Cook and Alexander D. O. Wedderburn, eds). London, New York: Longmans, Green and Co. 1903–1912. OCLC# 32081530.
- A bibliography of Samuel Taylor Coleridge,* John Louis Haney, Philadelphia: Printed for private circulation, 1903. OCLC # 01244508.
- Compromises.* Agnes Repplier. Boston: Houghton, Mifflin & Co. 1904. OCLC # 01844986.
- Sexual reproduction and the organization of the nucleus in certain mildews.* R. A. Haper. Washington, DC: Carnegie Institution of Washington. 1905. OCLC # 00535542.
- Chimæroid fishes and their development.* Bashford Dean. Washington, DC: Published by the Carnegie Institution of Washington. 1906. OCLC # 02323291.
- Biographia literaria,* John Shawcross. Oxford: The Clarendon Press. 1907. OCLC # 02774821.
- Variation and differentiation in Ceratophyllum.* Raymond Pearl. Washington D. C. Carnegie Institution of Washington. 1907. OCLC # 02360085.
- Roman Holidays: and Others.* William Dean Howells. New York, London: Harper & Bros. 1908. OCLC # 02663185.
- Fennel and Rue: a novel.* William Dean Howells. New York; London: Harper & Brothers Publishers. 1908. OCLC # 01021078.
- Actions and Reactions.* Rudyard Kipling. New York: Doubleday, Page & Co. 1909. OCLC # 00236439.
- A study of the absorption spectra of solutions of certain salts of potassium, cobalt, nickel, copper, chromium, erbium, praseodymium, neodymium, and uranium as affected by chemical agents and by temperature.* Harry C. Jones; W. W. Strong. Washington, DC: Carnegie Institution of Washington. 1910. OCLC # 02336051.
- The Old Order Changeth; A View of American Democracy.* William Allen White. New York: Macmillan. 1910. OCLC # 00854253.
- Clayhanger.* Arnold Bennett. New York: E. P. Dutton. 1910. OCLC # 00918462.
- Shakespeare bibliography: a dictionary of every known issue of the writings of our national poet and of recorded opinion thereon in the English language.* William Jaggard. Stratford-on-Avon: Shakespeare Press. 1911. OCLC # 01978611.
- Railway Economics: A Collective Catalogue of Books in Fourteen American Libraries.* Richard Holland Johnston, Bureau of Railway Economics (Washington D.C). Chicago: Bureau of Railway Economics by the University of Chicago Press. 1912. OCLC # 01437582.
- Regesta regum anglo-normannorum, 1066-1154.* H. W. Carless Davies, R. J. Whitwell, Charles Johnson eds. Oxford: Clarendon Press. 1913–1969. OCLC # 00661506.
- The germ-cell cycle in animals.* Robert William Hegner. New York: Macmillan Co. 1914. OCLC # 2361630.
- Genetic studies on a cavy species cross.* John Adolph Detlefsen. Washington, DC: Carnegie Institution of Washington. 1914. OCLC # 02678826.
- Chief contemporary dramatists: twenty plays from the recent drama of England, Ireland, America, Germany, France, Belgium, Norway, Sweden, and Russia.* Boston: Houghton Mifflin Company. 1915. OCLC # 02666849.
- The song of the lark.* Willa Cather. Boston, New York: Houghton Mifflin Company. 1915. OCLC # 00702452.
- The Cambridge History of American Literature.* William P. Trent; John Erskine; Stuart Pratt Sherman; Carl Van Dorer. New York: G. P. Putnam's Sons. 1917. OCLC # 01090047.
- God the Invisible King.* H. G. Wells. New York: The Macmillan Company. 1917. OCLC# 00383754.
- Outdoor Theaters; the Design, Construction and Use of Open-Air Auditoriums.* F. A. Waugh. Boston: R. G. Badger. 1917. OCLC # 01187029.
- The History of Henry Fielding.* Wilbur L. Cross. New Haven: Yale University Press; London: Humphrey Milford; Oxford University Press. 1918. OCLC # 01593752.
- Credit of the nations; a study of the European War.* J.

- Laurence Laughlin. New York: C. Scribner's Sons. 1918. OCLC # 00597768.
- On contemporary literature.* Stuart Pratt Sherman. New York: Holt. 1917. OCLC # 00674623
- The principles of American diplomacy.* John Bassett Moore. New York, London: Harper & Bros. 1918. OCLC # 00993154.
- Forced movements, tropisms, and animal conduct.* Jacques Loeb. Philadelphia: Lippincott. 1918. OCLC # 01891338.
- Reminiscences of Lafcadio Hearn.* Setsu Koizumi. Boston, New York: Houghton Mifflin. 1918. OCLC # 00478394.
- Dramatic technique.* George Pierce Baker. Boston, New York: Houghton Mifflin Company. 1919. OCLC # 00330380.
- Linda Condon.* Joseph Hergesheimer. New York: Alfred A. Knopf. 1919. OCLC # 00242478.
- Pawns, four poetic plays.* John Drinkwater. Boston: Houghton Mifflin Company. 1920. OCLC # 02476717.
- The unsolved riddle of social justice,* Stephen Leacock. New York: John Lane Company; London, John Lane. 1920. OCLC # 00497082.
- England in transition, 1789-1832, a study of movements.* William Law Mathieson. London, New York: Longmans, Green, and Co. 1920. OCLC # 00907796.
- Life and letters of Henry Lee Higginson.* Henry Lee Higginson, Bliss Perry. Boston: Atlantic Monthly Press. 1921. OCLC # 00234045.
- The Jew and American ideals.* John Spargo. New York, London: Harper & Bros. 1921. OCLC # 00555558.
- The mind in the making: the relation of intelligence to social reform.* James Harvey Robinson. New York: Harper & Brothers. 1921. OCLC # 00255133.
- Fossil Echini of the West Indies.* Robert Tracy Jackson. Washington, DC: Carnegie Institution of Washington. 1922. OCLC # 03133717.
- The revolt against civilization; the menace of the under man.* Lothrop Stoddard, New York: C. Scribner's Sons, 1922. OCLC # 01027004.
- Claudian.* Claudius Claudianus; Maurice Platnauer. London: W. Heinemann; New York: G. P. Putnam's Sons. 1922. OCLC # 00313897.
- The fiscal and diplomatic freedom of the British oversea dominions.* Edward Porritt; David Kinley. Oxford: The Clarendon Press; London, New York: H. Milford. 1922. OCLC # 21007534.

Appendix B. Assessment Data Points Collected and Definitions of Rankings

Storage Location	From ILL slip or book
Circulation history	If known from book
Barcode	
Title	
Author	
Publisher	
Publisher location	
Publisher date	
Other variance	
Facsimile	Y = 1, N = 0
Reviewed	Y = 1, N = 0
If no, reason	Y = 1, N = 0
Complete	
If no, describe	
Original cover	Y = 1, N = 0
Book plate showing provenance	Y = 1, N = 0
Original cover (from paperback release) mounted or bound in	Y = 1, N = 0
Evidence of original binding variance	Y = 1, N = 0
Library binding (older style in 1/4 or 1/2 binding)	Y = 1, N = 0
Library binding (buckram)	Y = 1, N = 0
Cover damage	none/wear only = 0, slight = 1, moderate = 2, severe = 3

Appendix B. Assessment Data Points Collected and Definitions of Rankings (continued)

Storage Location	From ILL slip or book	
Cover to text attachment	sound = 0, weak = 1, part detached = 2, detached = 3	
Tight inner margin or over trimmed (implies text loss if digitized)	Y = 1, N = 0	
Repaired	Y = 1, N = 0	
Rebacked	Y = 1, N = 0	
New case, inhouse	Y = 1, N = 0	
Book tape	Y = 1, N = 0	
Paper repaired with tape (various types)	Y = 1, N = 0	
Internal hinge reinforcement/reattachment?	Y = 1, N = 0	
Deacidified	Y = 1, N = 0	
Box	Y = 1, N = 0	
Envelope	Y = 1, N = 0	
Shrink wrapped	Y = 1, N = 0	
String tied	Y = 1, N = 0	
Other enclosure	Y = 1, N = 0	
Discolored	0-9 with 9 being most discolored	Standardized photography against grayscale calibration card
Brittle (visibly)	none = 0, slight = 1, moderate = 2, severe = 3	slight = minor edge or gutter breakages, moderate = regular edge or gutter breakages, severe = at least 1/3 of book showing edge or gutter breakages
Surface pH gutter	value taken on page 20 with Astro pH tester pen	
Surface pH edge	value taken on page 20 with Astro pH tester pen	
Tears greater than ½ inch	none = 0, slight = 1, moderate = 2, severe = 3	slight = 1 occurrence, moderate = 2-3 occurrences, severe = >3 occurrences
Underlining/highlighting/marginalia	none = 0, slight = 1, moderate = 2, severe = 3	slight = 1 occurrence, moderate = 2-3 occurrences, severe = >3 occurrences
Losses greater than ½ inch	none = 0, slight = 1, moderate = 2, severe = 3	slight = 1 occurrence, moderate = 2-3 occurrences, severe = >3 occurrences
Method of page attachment	sew through fold, oversewn, side sewn, adhesive, other	
Page detachment	none = 0, slight = 1, moderate = 2, severe = 3	slight = 1 occurrence, moderate = 2-3 occurrences, severe = >3 occurrences
Water damaged/stained/foxed	Y = 1, N = 0	
Broken text block	none = 0, slight = 1, moderate = 2, severe = 3	slight = 1 occurrence, moderate = 2-3 occurrences, severe = >3 occurrences
Notes		