Research and Discussion on Large-Scale Digitization of Print Collections: Postprint

Authors: Li Yongmei, Liu Liu, Liu Hong

Date: 2023-10-08T00:00:00+00:00

Abstract

[Purpose / Significance] Large-scale digitization of print collections is an inevitable trend in the development of digital libraries. It not only infinitely expands the depth and breadth of library services, but also greatly promotes the development and advancement of scholarly communication and academic research.

[Method / Process] This study analyzes and examines the global large-scale digitization movement of print collections from the perspectives of development origins, environmental factors, current status, and existing issues, using the HathiTrust project in the United States as an empirical case study.

[Result / Conclusion] Establishing principles for large-scale digitization, resolving copyright issues based on the fair use doctrine, and building a multi-level and multi-dimensional cooperation framework are important initiatives to promote the sustainable development of large-scale digitization of print collections.

Full Text

Research on the Mass Digitization of Print Collections

Li Yongmei[1], Liu Liu[1], Liu Hong[2]

Abstract

[Purpose/Significance] The mass digitization of print collections represents an inevitable development for digital libraries, infinitely expanding both the depth and breadth of library services while profoundly advancing academic exchange and research. **[Method/Process]** This paper analyzes the global mass digitization movement of print collections from the perspectives of its origins, environmental factors, current development status, and existing challenges, using the HathiTrust project in the United States as an empirical case study.

[**Result/Conclusion**] Establishing principles for mass digitization, addressing copyright issues through the fair use doctrine, and constructing a multi-level, multi-dimensional cooperation framework constitute crucial measures for promoting the sustainable development of print collection mass digitization.

Keywords: print collection; mass digitization; cloud library; HathiTrust; fair use

---

Initiated in 2004, the Google Books project sparked a wave of mass digitization of print collections. The establishment of HathiTrust in 2008 pushed this movement to its climax. In February 2014, the HathiTrust Digital Library's collection surpassed 11 million volumes [1], marking a milestone in its development history. By February 2017, HathiTrust's holdings had exceeded 15 million volumes [2], achieving a new leap forward. This spark has now become a prairie fire, as digitization initiatives undertaken by libraries worldwide have greatly promoted the development and progress of academic exchange and research.

1. Origins

Traditional print resources typically undergo three stages—publication, dissemination, and acquisition—before reaching end users, with their value realized through use. For anyone seeking to read or use resources, the desire has always been for faster, more convenient, unrestricted, and completely barrier-free access regardless of time or space, representing the ultimate goal that libraries have continuously pursued. When the number of end users reaches its maximum—that is, when print resources are utilized by as many people as possible—their value is most fully demonstrated. Mass digitization of print resources serves as the essential means to achieve this objective, maximizing the use value of print resources, greatly facilitating the dissemination of human civilization and knowledge, and promoting the depth and breadth of academic exchange and research.

In December 2004, Google announced the launch of the Google Print Library Project, partnering with several major libraries including Harvard University, the University of Michigan, the New York Public Library, Oxford University, and Stanford University. These institutions' collections were estimated to exceed 15 million volumes [3]. The agreement stipulated that Google would provide funding, technical equipment, and scanning and cataloging services for the collections, enabling readers worldwide to search for and access snippets, abstracts, and information about library holdings and market availability from the Google website. Within a few years, Google had scanned nearly ten million books for online search [4].

In April 2008, Dutch libraries pioneered a new approach by signing a declaration of intent with copyright holders regarding the digitization and accessibility of heritage collections. This represented the world's first agreement of its kind be-

tween libraries and copyright owners, with its core provision being that libraries could obtain authorization from copyright holders to digitize their works and make them publicly accessible within library premises for teaching or research purposes [5]. This development strongly propelled the mass digitization movement of traditional collections.

In October 2008, HathiTrust was established as a shared digital repository project jointly created by the Committee on Institutional Cooperation (CIC) library consortium and the University of California library system. The project digitizes and centrally preserves print documents collected by its member libraries [6], marking the vigorous development of the mass digitization movement for print collections.

2.1. Major Developments in Digital Libraries

With the advancement of digitization and networking, an increasing number of resources are now network-based. The convenient, cross-temporal accessibility of online digital resources has made digital reading a new and increasingly mainstream mode of reading. Users' evolving reading habits and behavioral patterns compel libraries to maintain a leading edge in new technologies and user experience—the driving force behind digital library development. The growing demand for digital collections and new requirements for mobile services have made traditional print collections increasingly inadequate for meeting users' online reading needs. Mass digitization has become a crucial component of digital library development, enabling libraries to maximize online access to their collections and services. Simultaneously, mass digitization objectively reduces space and personnel requirements, thereby lowering collection management costs.

Furthermore, as mobile technology advances rapidly, mobile devices are transforming how people transmit and access information. Mobile reading has emerged as a mainstream mode due to its advantages: portability and mobility for reading work, searchability and timely access to content, and persistence and continuity in reading behavior [7]. This trend places higher demands on the diversity and richness of digital library service content. To adapt to the evolving mobile environment, libraries have begun utilizing mobile devices to deliver services and content, allowing users to conveniently access digitized print resources via smartphones. The rise of the big data era has also placed more urgent demands on library digitization development.

2.2. Major Developments in the Open Access Movement

The open access movement is founded on the principle that research results should be accessible to anyone who wishes to read and use them without barriers, thereby maximizing the social, economic, and cultural benefits of scientific achievements. Openness, transparency, and easy access to resources have gradually become a valued principle.

In December 2007, the U.S. government passed legislation mandating open access for NIH-funded research. In 2012, the Federal Research Public Access Act (FRPAA) was reintroduced to the House and Senate, expanding mandatory open access to all federal agencies. The League of European Research Universities also released statements on open research data and open access to research publications [8]. The Wellcome Trust, one of the world's largest research funding agencies and a long-time advocate and leader in open access, required as early as 2006 that original research papers it funded be deposited in Europe PubMed Central for open access within six months of publication. In 2012, it introduced a stronger deposit policy requiring that monographs and book chapters from new projects funded after October 2013, and all funded projects after October 2014, follow the same open access policy as journal articles [9]. These mandatory open access policies have greatly propelled the development of open access.

The impact of the open access movement on academic research has been unprecedented, promoting the digitization of academic achievements while further advancing the rapid progress of the open access movement itself. This has provided users worldwide with broader and deeper public access to literature resources, making the digitization of print collections an inevitable choice for libraries seeking to strengthen their core competitiveness in an open access environment.

2.3. Development Trends in Higher Education

Higher education is facing unprecedented opportunities and challenges, undergoing significant transformation as traditional educational systems and models are increasingly challenged by technological development. Competition among higher education institutions has intensified and is rapidly becoming more international. Universities in Europe and America have established branch campuses in Asia and the Middle East—for example, New York University in China and the UAE, the University of Nottingham in China and Malaysia, and Yale University in Singapore—with numerous transcontinental cooperative education and degree-granting institutions emerging continuously.

Simultaneously, due to widespread use of the internet and tablet computers, online teaching and learning have become popular. Massive Open Online Courses (MOOCs) have risen as a powerful force, gaining increasing popularity through their extensive free content and high-quality instruction, gradually being adopted as alternatives and supplements to traditional university courses [10]. These transformations in higher education and academic research undoubtedly have major impacts on libraries in several areas: collection development and growth, resource acquisition and preservation, and continuously expanding user services [11]. Digitizing print collections, including various special and rare collections, can provide richer network resources for teaching and research, meeting the needs of institutions for online instruction and academic innovation while enhancing competitiveness.

3.1. Current Development Status

As of March 2012, Google had digitized over 20 million books [12], while HathiTrust, after more than five years of development, had reached a collection of 11,262,697 volumes [13]. [Figure 1: see original paper] shows HathiTrust's annual statistics for its digital collections from June 2009 to June 2014 [14]. The figure indicates that during this five-year period, its collection grew by 240%, with an average annual growth rate of nearly 50%, representing rapid expansion.

The figure also reveals that HathiTrust's collection growth slowed after 2012 due to Google reducing its book scanning efforts, reflecting the natural maturation process of the project. The possibility of duplicate scanning increased, thereby reducing acquisition volume, while some institutions broke from their conventional agreements with Google to scan only special collections—for instance, the University of Texas at Austin scanned only its Latin American collection [12].

Meanwhile, HathiTrust has placed considerable emphasis on building its public domain collection, with its volume growing annually and comprising an increasingly larger proportion of the total collection. As shown in Table 1 "HathiTrust Public Domain Collection Statistics (June 2009–June 2014)" [14], by June 2014, 34% of the collection was freely accessible in full text and open to global users. By February 2017, 5,811,123 volumes had entered the public domain, accounting for 38.5% of the total collection of 15,101,537 volumes [15].

HathiTrust's June 2014 statistical report also revealed that the University of Michigan Library and the University of California Libraries ranked first and second in scanning volume, with 4,689,072 and 3,520,634 volumes respectively [13], representing 41.6% and 31.3% of HathiTrust's total collection. Combined, these two institutions account for 73% of the total, demonstrating that they are HathiTrust's primary contributors.

Many HathiTrust member libraries provide digital collection printing services. For example, the University of California Libraries offers reprint services for books in its digital collection that have entered the public domain, with delivery through Amazon. Google has gone further by introducing on-demand printing services, designing a rapid book-printing machine that allows readers to obtain a bound paperback for approximately eight dollars within five to ten minutes [4], making the process both economical and efficient.

HathiTrust continues to expand its membership, enabling member libraries to deposit their locally digitized collections into HathiTrust—another reason for its rapid collection growth. Additionally, HathiTrust continuously develops new tools (such as SIP) to simplify and accelerate the deposit process for member libraries.

Europe has also undertaken numerous digital library construction projects. The British Library, for instance, has implemented extensive digitization projects to digitize its cultural heritage, creating digital surrogates for numerous precious

manuscripts while making them widely accessible to the public. The library has recently participated in several global large-scale digitization projects, including the Google Books project and the brightsolid newspaper digitization project [16].

China's digitization efforts have primarily focused on special collections, with significant progress made particularly in the digitization of ancient Chinese texts. According to statistics, by 2012, public institutions in China possessed over 2 billion characters of digitized ancient texts [17], while international cooperation projects have increased, with the National Library of China undertaking the most initiatives. The most representative international cooperation project among universities is CADAL (China Academic Digital Associative Library). As of February 2014, CADAL's collection reached 2,748,688 volumes/items, including ancient texts, books, journals, and dissertations. Among these, Chinese-language books totaled 1,130,595 volumes (including 236,581 ancient texts), and foreign-language books numbered 767,059 volumes [18], making it the representative project for large-scale digitization of print books in Chinese libraries, allowing any registered user to read and print for free.

3.2. Existing Problems

The rapid development of print collection mass digitization has encountered numerous difficulties, with the following issues becoming obstacles to its progress.

3.2.1. Copyright Issues

All controversies surrounding mass digitization are related to copyright; the history of mass digitization is essentially a litigation history of Google. Beginning in 2005, when the American Authors Guild and the publishing industry sued Google, the company has faced continuous infringement lawsuits from an increasing number of industry organizations and publishers. In March 2011, a U.S. court rejected Google's Book Settlement Agreement [19], casting a shadow over the development of mass digitization. The Library Copyright Alliance issued a statement: "This ruling clearly demonstrates that current copyright law continues to pose significant obstacles for libraries and other partners interested in undertaking mass digitization projects" [20]. HathiTrust has also been troubled by copyright issues, with the American Authors Guild suing its digital library for violating U.S. copyright law. The Digital Public Library of America (DPLA), launched in 2013 to provide free online public access, has also identified copyright as one of its biggest challenges. Evidently, copyright issues remain a major obstacle to mass digitization of print resources.

3.2.2. Long-term Management and Preservation

Data management and preservation represent the most important and difficult tasks for all digitization projects. Digital storage typically includes both bibliographic data and full-text data. Unlike physical resources, digital resources have a very short lifecycle without special protection. As digital libraries mature and

digital collections grow, long-term management and preservation issues have become increasingly prominent. Individual libraries, in particular, show significant deficiencies in awareness and planning for long-term preservation, which is extremely detrimental to the sustainable development of local digital collections. Libraries invest far more in creating new resources than in maintaining or optimizing existing ones—once collections are digitized, they are often not actively managed [21].

Reports from Portico and Cornell University Library indicate that most cultural heritage institutions are only beginning to recognize that their substantial investments in building digital collections must be supported by infrastructure capable of preserving content throughout its lifecycle [22]. An OCLC survey of 169 research libraries in the United States and Canada regarding special collection digitization revealed that 97% had completed one or more digitization projects, while lack of funding, expertise, and administrative support were identified as the greatest obstacles to digital collection management and preservation [23].

3.2.3. Metadata Interoperability Issues

Metadata is an essential element for effectively organizing and revealing digital resources. However, the diversity of institutions undertaking mass digitization projects and the rich variety of content and formats have resulted in multiple coexisting metadata standards, making metadata collection and organization extremely difficult. This creates challenges for integrating digital resources to meet users' unified retrieval needs, inevitably constraining the sustainable development of digital libraries. The key to solving this problem lies in achieving metadata interoperability. Research investigating metadata interoperability among major domestic and international digitization projects—including CADAL, the National Digital Library of China, Google Books, Europeana, Open Library, HathiTrust, the California Digital Library (CDL), and American Memory—shows that while some progress has been made, a fundamental solution remains a long way off [24].

4. Reflections and Recommendations on the Development of Print Collection Mass Digitization

After more than a decade of development, mass digitization has achieved remarkable accomplishments, not only advancing digital library development but also promoting broader and deeper academic exchange and research, particularly in knowledge integration and innovation. However, difficulties and problems encountered during its development have become obstacles. The following reflections and recommendations address how to overcome these challenges and better promote development.

4.1. Establishing Principles for Mass Digitization

In June 2010, the Association of Research Libraries (ARL) released nine principles for mass digitization [25] to guide its member libraries in cooperating with vendors or publishers while digitizing special collections and to encourage their adoption. Drawing on ARL's guiding principles and considering the actual development of mass digitization, we propose principles that Chinese libraries should follow when undertaking print collection mass digitization.

4.1.1. Fair Use Principle Fair use is a mechanism in copyright law that balances the interests between copyright holders and users, representing a compromise between intellectual property protection and public interest. In the digital age, where the intellectual property environment has become increasingly complex, the application of fair use principles is particularly important. However, due to subjective influences, accurately grasping the boundaries of fair use is not easy. When following this principle, libraries must clearly define the objects and purposes of fair use to maximize protection of copyright holders' interests while providing services to the public.

4.1.2. One-time (Scanning) Principle To protect print and special collections, particularly rare resources, the one-time scanning principle should be followed whenever possible. Digitization scanning always causes varying degrees of damage to print resources, requiring libraries to conduct careful organization and planning to develop forward-looking scanning programs.

4.1.3. Preservation of Copies and Update Principle Libraries should obtain and completely preserve all digital files generated from their collections, regularly updating these digital contents. Within the scope permitted by copyright law, they should provide unrestricted local access to these copies for relevant users and exercise fair use rights.

4.1.4. Sharing Principle Libraries' natural mission is to inherit and disseminate knowledge and culture. Against the backdrop of open access, sharing has evolved into a "universal" value in the library community. Sharing digital collections among libraries can prevent redundant investment by other institutions, thereby avoiding waste of resources, funding, and manpower. HathiTrust's shared digital repository project represents the best practice of this principle. Sharing requires cooperation to achieve, which will be discussed in detail below.

4.2. Addressing Copyright Issues Through Fair Use

Copyright issues represent the greatest obstacle in the mass digitization of print collections. Libraries must scientifically and reasonably apply the "fair use" doctrine within the framework of copyright law to resolve copyright issues properly.

The American Authors Guild sued HathiTrust in the Southern District Court of New York for violating U.S. copyright law. In October 2012, the court ruled that HathiTrust Digital Library's scanning and digitization of print resources indeed constituted fair use under U.S. copyright law, representing a significant victory for fair use and American libraries [26]. Subsequently, the American Authors Guild appealed to the U.S. Court of Appeals for the Second Circuit, which in June 2014 upheld the lower court's decision [27]. Most encouragingly, in October 2015, the appellate court ruled that Google's digitization of books was entirely legal, representing fair and normal use of books and benefiting society [28]. This decision concluded the decade-long litigation, declaring Google's ultimate victory and, to some extent, clearing obstacles for mass digitization.

These cases demonstrate that the fair use principle has tremendous guiding significance and practical value in resolving copyright issues. However, because fair use objectively causes some degree of damage to copyright holders' interests, international copyright law imposes strict limitations and criteria to prevent overuse and abuse of fair use.

4.3. Establishing a Multi-level, Multi-dimensional Cooperation Framework

Cooperation is the key to success for library endeavors and an inevitable means to achieve in-depth resource sharing. Both the Google Library Project and HathiTrust are products of cooperation, and their ultimate victories in copyright litigation are inseparable from the active role played by the Library Copyright Alliance.

The Cloud Library Project, jointly implemented by OCLC, HathiTrust, New York University's Elmer Holmes Bobst Library, and the Research Collections and Preservation Consortium (ReCAP), represents an excellent practice. Funded by the Andrew W. Mellon Foundation, this project demonstrates strong cooperation among libraries and between libraries and funders—a forward-looking and sustainable initiative worthy of emulation and promotion. The project's research showed that as of June 2010, 30% (median) of ARL member libraries' collections overlapped with HathiTrust Digital Library. At New York University's Bobst Library, the number of volumes duplicated in both ReCAP and HathiTrust increased monthly from September 2009 to June 2010, reaching approximately 200,000 volumes by June 2010 [29]. If HathiTrust can provide comprehensive services, duplicate print collections can be withdrawn from libraries, immediately saving considerable space and management costs. HathiTrust now holds 15 million digital volumes, 2.42 times the amount in 2010. Conservatively estimated, the duplication rate between ARL member libraries' collections and HathiTrust has reached 60% (median).

The Cloud Library Project aims to achieve highly efficient management and optimal allocation and integration of print collections in a mass digitization environment, thereby promoting the repositioning of collection development and

library service innovation. Although HathiTrust's digital repository already contains 15 million books, comprehensive implementation of a cloud library plan for print collections requires cooperation with other large service institutions to share collection resources. For example, the Library of Congress is a print book repository. Although it became a HathiTrust member in 2011, its digitized collection in HathiTrust comprises only slightly over 100,000 volumes. More comprehensive and in-depth cooperation between HathiTrust and the Library of Congress could cover at least 70% of digitized book resources. Both HathiTrust and the Library of Congress are important partners of the Digital Public Library of America [30], and deeper cooperation among the three would greatly expand the scope and impact of the Cloud Library Project.

HathiTrust officially launched its Shared Print Program in 2016, representing deeper cooperation and sharing based on its shared digital repository project. The program aims to ensure the preservation of both print and digital collections through linking them, thereby reducing collection management costs for member libraries while promoting joint management and sharing of national and even intercontinental collections. The program is being implemented in two phases, with substantial progress achieved in Phase I: over 50 member libraries joining as retention partners were required to confirm their planned retention of print books by April 2017 [31]. The Shared Print Program not only ensures the secure retention of print resources corresponding to HathiTrust's digitized collections but also guarantees long-term availability of print collections for interlibrary loan among member libraries, providing more reliable support for the ultimate realization of the Cloud Library Project.

Intercontinental cooperation is also gaining momentum. Europeana and the Digital Public Library of America announced a partnership enabling users to access the combined rich collections of both systems, with cooperation encompassing interoperable digital models, shared resource standards, and collaborative collection development [32].

The ultimate success of print collection mass digitization requires concerted cooperation among research institutions, universities, researchers, libraries, publishers, copyright holders, and funding agencies to establish a balanced system acceptable to all stakeholders with sustainable development.

References

[1] Update on February activities 2014[EB/OL]. [2017-01-25]. http://www.hathitrust.org/documents/hathitrust-update-201402.pdf.
[2] February 2017 update[EB/OL]. [2017-04-06]. https://www.hathitrust.org/february-2017-update.
[3] Google Books history[EB/OL]. [2017-01-02]. http://books.google.com/googlebooks/about/history.html.
[4] Zou Jiping, Dong Xiaofen. Google Library, or the Googlization of Libraries—A Review and Prospect of Resource Sharing in American Libraries[J]. Library and Information Service, 2010(5): 65-69.

[5] Guo Mingzhu. Traditional Collection Digitization Is Just Around the Corner—Libraries Reach Agreement with Copyright Holders[J]. Library and Information Service Newsletter, 2009(2): 38-39.
[6] Li Yongmei, Yuan Xueliang, Tang Lixing. The HathiTrust Project in the United States and Its Implications for China[J]. Journal of Sichuan Library, 2011(3): 35-37.
[7] Zhao Naixuan, Zhou Jingzhen. Transformation, Innovation, and Cooperation of University Libraries in the Mobile Reading Environment[J]. Library and Information Service, 2012, 56(9): 48-52.
[8] LERU launches statements on open access and open data[EB/OL]. [2017-04-02]. http://www.leru.org/index.php/public/news/leru-launches-satements-on-open-access.
[9] Wellcome Trust extends open access policy to include scholarly monographs and book chapters[EB/OL]. [2017-01-02]. http://www.wellcome.ac.uk/News/Media-office/Press-releases/2013/WTP052746.html.
[10] The New Media Consortium. Horizon report 2013 higher education edition[EB/OL]. [2017-06-01]. http://www.nmc.org/pdf/2013-horizon-report-HE.pdf.
[11] The ACRL Research Planning and Review Committee. Environmental scan 2013[EB/OL]. [2017-02-06]. http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers
[12] HOWARD J. Google begins to scale back its scanning of books from university libraries[EB/OL]. [2017-05-30]. http://chronicle.com/article/Google-Begins-to-Scale-Back/131109/.
[13] Update on June 2014 activities[EB/OL]. [2017-01-26]. http://www.hathitrust.org/updates_{june2014}.
[14] All newsletters and reports[EB/OL]. [2017-01-26]. http://www.hathitrust.org/updates.
[15] Currently digitized[EB/OL]. [2017-04-06]. https://www.hathitrust.org/visualizations_{{{deposited}}{{*vo*
[16] British Library digital preservation strategy[EB/OL]. [2017-06-25]. http://www.bl.uk/aboutus/stratpolprog/collectioncare/digitalpreservation/strategy/DigitalPreservationStrategy_{2013}-16-external.pdf.
[17] Gao Juan, Liu Jiazhen. Issues and Countermeasures for Ancient Book Digitization in Mainland China[J]. Journal of Library Science in China, 2013(4): 110-119.
[18] Total Resources Deposited in CADAL Project[EB/OL]. [2017-03-10]. http://www.cadal.cn/zydt/index1402.htm.
[19] BAND J. A guide for the perplexed part IV: the rejection of the Google Books settlement[EB/OL]. [2017-01-09]. http://www.districtdispatch.org/wp-content/uploads/2011/04/GuideIV-FINALV3.pdf.
[20] LCA releases statement on Google Books settlement rejection[EB/OL]. [2017-05-21]. http://old.arl.org/news/pr/gbs_{statement24mar11}~print.shtml.
[21] MARON N, PICKLE S. Evaluating Digital Investments: The Sustainability of Digitized Special Collections in ARL Libraries[J]. Translated by Zhu Manman. Library and Information Service Newsletter, 2013(3): 15-16.
[22] Portico, Cornell University Library. Preservation of digitized books and other digital content held by cultural heritage organizations[EB/OL]. [2017-06-25]. http://www.portico.org/digital-preservation/wp-content/uploads/2010/01/NEH-IMLS-D-book-model.pdf.

[23] DOOLEY J, LUCE K. Taking our pulse: the OCLC research survey of special collections and archives[EB/OL]. [2017-06-25]. http://www.oclc.org/content/dam/research/publications/li 11.pdf.
[24] Song Linlin, Li Haitao. Investigation and Implications of Metadata Interoperability in Large-Scale Document Digitization Projects[J]. Journal of Library Science in China, 2012(5): 27-38.
[25] Principles to guide vendor/publisher relations in large-scale digitization projects of special collections materials[EB/OL]. [2017-02-26]. http://www.arl.org/bm~doc/principles_{{{large}}*{{scale}}}*{digitization}.pdf.
[26] Molly Raphael. The Digital Shift: The Impact of E-books and Digital Content on Readers and Libraries[J]. Translated by Zhang Yuling. Journal of Library Science in China, 2013(1): 31-35.
[27] Liu Wenqi. Determination of Fair Use for Large-Scale Digital Reproduction under Copyright Law Framework[J]. Electronics Intellectual Property, 2014(12): 69-73.
[28] U.S. Court Decides Decade-Long Lawsuit: Google's Book Digitization Ruled Legal[EB/OL]. [2017-05-05]. http://tech.qq.com/a/20151017/008368.htm.
[29] MALPAS C. Cloud-sourcing research collections: managing print in the mass-digitized library environment[EB/OL]. [2017-01-02]. http://www.oclc.org/content/dam/research/publica 01.pdf.
[30] History[EB/OL]. [2017-09-05]. https://dp.la/info/about/history/.
[31] Shared Print Program[EB/OL]. [2017-06-15]. https://www.hathitrust.org/shared_{{{print}}}_{{program}
[32] Europeana and Digital Public Library of America Announce Cooperation[J]. New Technology of Library and Information Service, 2012(3): 96.

Author Contributions

Li Yongmei: Conceived the research framework and wrote the paper;
Liu Liu: Responsible for data collection;
Liu Hong: Revised the paper.

Note: Figure translations are in progress. See original paper for figures.