

# *Embracing the Future of Digital Libraries within Theological Libraries*

PAUL A. TIPPEY, ASBURY THEOLOGICAL SEMINARY

**I**n 1938, H. G. Wells unveiled his vision of a “world brain,” saying: “The time is close at hand when any student, in any part of the world, will be able to sit with his projector in his own study at his or her own convenience to examine any book, any document, in an exact replica” (1938, 77). In 1990, fifty-two years later, the first digital libraries began to appear. Although digital libraries are still evolving, the technical obstacles that dominated the first phase of digital library development have generally been overcome through advances in computers, networking, and algorithms (Lesk 2012). Universal access to “any book, any document” as envisioned by H. G. Wells is now both technically feasible and economically possible; however, significant social and legal barriers still remain.

In the coming years, digital theological libraries will provide access to a wide variety of resources, integrating content from diverse sources including images, texts, video, etc. These digital libraries will provide a seamless environment where research is transformed by the ability to filter, manipulate, and interact with materials like never before. Users of digital libraries will be both consumers and producers of information, both individually and in collaboration with others. With each of these changes, both past and future, the role played by libraries and information professionals must evolve. This chapter will examine the four principal barriers (technical, economic, legal, and social) to the development of digital theological libraries in order to prepare theological librarians for the challenges we face as we redefine the role of our profession in the days ahead.

## *Technical*

H. G. Wells' 1938 vision set the stage for the development of the "world brain." In order for this to happen, the issue of machine translation, as well as that of information retrieval, had to be solved. In the 1960s, technical difficulties still existed at each stage of the process. The input stage consisted of keystroke documents in order to get the documents into a machine-readable form, a process that was vulnerable to input errors, the computation stage could only handle small collections, and the output stage was limited by retrieval systems. By the end of the 1960s, enough technology existed to build the first retrieval systems. With the establishment of computer typesetting and online access, commercial systems started to appear with Boolean search mechanisms. During the next two decades, essentially all production of published documents migrated to computers, and it became customary for a machine-readable copy of all new text to exist (digitally native content).

In the early 1990s, a breakthrough occurred that changed the future of digital libraries: algorithms for indexing and searching were created. Before this, the first internet-based searching-systems were based on manual indexing and hierarchical structures similar to traditional libraries. The advent of algorithm-based searching allowed large amounts of text to be inputted and every word to be indexed automatically. At the same time, professional scanners were available for publishers, print shops, and larger organizations to digitize traditional materials. This digitization process scanned the materials and saved them as images, and then optical character recognition (OCR) software was used to convert the image into an editable text with reasonable accuracy for searching capability.

In the 1990s, most people still preferred physical media. Screen reading was perceived as difficult and inconvenient, and people wanted the feel of paper and even the smell as they used the material (AntonBergen 2008 cleverly illustrates these attitudes). Even then, however, people were drawn to digital materials, as they could be instantly accessed from their desks and searched at the word level.

The popularity of digital materials grew as these advantages became better known. Perceptions really began to change when journals began to shift to digital format in the early 2000s. At first, most journals chose between paper or electronic versions; then they offered both versions. Now, twenty years later, some journals are shifting completely to electronic versions. Perceptions of digital books have also changed, particularly after the advent of the Amazon Kindle in 2007. By 2011, Amazon was selling more electronic books than physical books (Savitz 2011).

With the technological advances of the past several decades, technology has become less of a barrier to digital libraries. Technology has become less expensive, more reliable, quicker, and in some cases even automatic. The future is unknown,

but with breakthroughs like Google's claims of "quantum supremacy," artificial intelligence could be the next breakthrough (Metz 2019), bringing us closer than ever to realizing Wells's vision of the "world brain."

## *Economic*

It was unclear in the beginning, how this "world brain" was going to be economically supported. Many models have been used by commercial publishers such as monthly or yearly subscription fees, per-minute fees, access fees for signing up new users, transaction fees for downloading and advertising, and the cost charged per page. The per-minute and the access-for-signing-up-new-users models of sharing eventually collapsed. The per-minute model possibly collapsed due to the fact that plenty of people are willing to provide information for free, caring more about recognition than cash, or perhaps simply for the good of society. Regardless of the reason, the collapse of this early model of economic feasibility raises the question: how is this "world brain" going to be paid for (Lesk 2005, 597)?

Predominantly, the same sources that paid for information on paper are paying for it digitally: libraries, readers, and even authors and grants. Generally, publishers have converted paper publishing to electronic publishing to simplify their production process and, at the same time, increase their sales. Libraries (including theological libraries) have been using their acquisition budgets to purchase electronic copies in order to give users better services and avoid shelving costs. Meanwhile, individual readers can often buy current books on the Kindle or other electronic readers thanks to publisher programs. In addition, readers can buy individual articles if the library does not subscribe to the whole journal. Since electronic publishing can be done one copy at a time, self-publishing has been exploding for both books as well as scholarly articles.

Additionally, there has been a significant increase in open access options. When retrospective scanning is not provided by publishers and it falls within copyright laws to do so, many libraries will often do it themselves, sometimes funded by grants or donations, creating digital repositories of materials in their special collections. Some libraries have even established their own open access presses.<sup>1</sup>

The rapid spread of open access publishing is reshaping the very nature of academic publishing. In general, the funds needed for open access come from the authors, grants, donations, or library budgets. We do not yet understand whether a shift to open access will save theological libraries more in subscription fees than it costs them in repository operations. However, it is clear that an argument could be made that open access articles provided by students and faculty have a significant

economic impact. In addition, open access articles are cited on average nine times while toll-access articles are cited on average six times, which can influence tenure and promotion decisions, affecting both individual faculty and institutions (Norris, Oppenheim, and Rowland 2008).

The change from print to digital resources may have been slow for theological libraries, but a great number of articles and books are available online today for free. Even more are available for purchase in a digital format. Clearly, economic problems continue to be a challenge for digital libraries, but, at the same time, the digital library is now economically possible.

## *Social*

The largest issues facing the “world brain” are actually social in nature. Currently, the quantity of available material is outpacing the quality of the material. For example, the number of books being published is exploding, despite the fact that the number of books being sold is falling fast. Lesk (2012) asks, “How do we avoid a world in which junk information is taking over because the new world has less effective refereeing and reviewing?” Having access to more resources is good, but they need to also be usable resources.

A related issue has arisen through the use of the same search algorithms used to solve the technical problems explored above. Many algorithms are based on people’s reviews and the number of downloads, but just because a resource is being used does not mean it is the best resource for a given situation. Additionally, when libraries use MARC records provided by a vendor or within some federated searching service, searches may inappropriately privilege that vendor’s own resources. Furthermore, these third-party algorithms are owned by the company and generally cannot be seen or adjusted by the library. Additional issues arise because effective filtering tends to show people only what they agree with already.

Another social issue is the dependency of the “world brain” on private companies or non-profit organizations that rely on donors. Private companies have less of a responsibility to keep resources available; according to the website *Killed by Google* ([www.killedbygoogle.com](http://www.killedbygoogle.com)), for instance, Google has discontinued 194 different services to date. Although Google and other companies make some amazing resources available online, there is no guarantee those resources would survive the next dot-com crash. Similarly, non-profit organizations that rely on donors for survival could also be impacted in the long run by shifts in the cultural climate and their donors’ shifting priorities.

According to Lesk (2012, 600) and based on sample study, large-scale book scanning projects like Project Gutenberg (est. 1971), the Million Book Project (est.

c. 2001), Google Books (est. 2004), the Open Content Alliance (est. 2005), and others have scanned pre-1920 US-published books more than six times (Lesk 2012, 598-99). While the rate of scanning has slowed, “this scanning project helped establish some important nodes in what’s become an ever-expanding web of networked research” (Howard 2017).

This ever-expanding web of networked research nodes was helped by Google’s scanning project. However, while Google was embroiled in decade-long litigation, the partner libraries wanted to make sure they kept their digital copies for research as well as for preservation. This desire led to the establishment of the HathiTrust Digital Library in 2008. The HathiTrust Digital Library contains more than 18 million monograph volumes, the majority coming from Google’s scanning project (both public domain and copyrighted works), the Internet Archive, and local digitization efforts.

Another social shift for theological institutions is the significant changes to educational delivery methods which have influenced the growth of digital collections as well. Though the Association of Theological Schools initially had given specific guidance within the Educational Standards, prohibiting distance courses from constituting “a significant portion of a degree program,” they seem to be backing away from this policy and have granted an exception to a number of schools, allowing for degrees offered completely online. To meet the needs of a growing population of distance learners, libraries must expand their access to digital materials through either purchase or digitization of printed materials.

Collaborative projects like HathiTrust could help solve most of these social issues. Quality control could be implemented that is similar to how MARC records were handled in the past or how Wikipedia uses crowdsourcing to sort materials and point readers to valuable, obscure materials. Additionally, companies should allow access to the algorithms or, at the very least, allow the library to have additional control of the algorithm. As for private companies and non-profit organizations, they should form a crowdsourced joint shared research node.

## *Legal*

Despite the economic barriers to digital libraries being largely overcome, considerable legal barriers remain, particularly in the form of copyright law. Copyright law can be understood as an attempt to create an appropriate balance between competing interests. Copyright is not a natural right; it is a privilege granted by Congress, giving limited ownership of intellectual material to creators/authors. Initially, copyright was instituted to encourage the creation of creative works, but it has instead turned into a market place for financial

enrichment (Lessig 2004, 6, 78). It is the foundational goal of copyright to enhance democratic culture and to support civil society as a whole. According to the US Constitution, the purpose of copyright is “To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries” (art. 1, sec. 8, cl. 8).

Throughout history, copyright has adjusted to changing commercial practices and evolving technologies, e.g., lithography, radio, sculpture, cinema, television (both broadcast and cable), and reprography (U.S. Congress 1986). In 1976, Congress instituted copyright for the first time for unpublished manuscripts. Before 1976, creators/authors had to register a work with the Library of Congress and post a copyright notice on the work in order for the work to be protected under copyright. This is no longer the case. As soon as the creation is recorded on paper or some type of medium, it is now under copyright protection. In 1976, the United States Congress defined five exclusive rights possessed by copyright holders:

- to reproduce the work and to exclude others from reproducing;
- to derive new works from the work and to exclude others from making derivative works;
- to distribute copies and to exclude others from distribution copies;
- to perform the work—e.g., a play—publicly and to exclude others from so doing;
- to display the work—e.g., a poster—publicly and to exclude others from displaying it (*U.S. Code 17* [2006], § 106).

In the spring of 2003, the duration of this ownership was extended to the life of the creator/author plus seventy years for works not done for hire. On January 15, 2003, the *Sonny Bono Copyright Term Extension Act* was upheld by the United States Supreme Court (*Eldred v. Ashcroft*, 537 US 186).

Copyright law has directly impacted the rise of digital media and digital libraries, and it can be expected to continue to do so in the future. According to the ALA (2019), “Copyright issues are among the most hotly contested issues in the legal and legislative world; billions of dollars are at stake. Legal principles and technological capabilities are constantly challenging each other and every outcome can directly affect the future of libraries.”

The development of the “world brain” does not align well with standard legal views about intellectual property. Traditionally, the author and/or the publisher were involved in the first-use market. The original purchaser had to buy the publication from a legitimate copyright holder. Once purchased, they could sell to a second-user market. This model assured that the author and/or publisher got their share of the profits from the initial sale, and it allowed the buyer rights to

resell the item. Digital materials or items break down protections in both first-user and second-user markets.

Libraries and archives whose collections are open to the public have their own privileges and restrictions under copyright law, including the right to make copies of copyrighted works as long as there is no commercial advantage and the works are accompanied by a copyright notice. Three copies are allowed for preservation, but digital copies are not allowed outside of the library or archives. Under certain conditions specified in the copyright law of the United States (*U.S. Code* 17 [2006], § 108), libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specific conditions is that the photocopy or reproduction is not to be “used for any purpose other than private study, scholarship, or research.” In addition, the Fair Use provision in sections 106 and 106A allows for “...reproduction in copies or phonorecords or by any other means specified in that section, for purpose such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research” (*U.S. Code* 17 [2006], § 106-106A). If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of “fair use,” that user may be liable for copyright infringement. The institution must reserve the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law.<sup>2</sup>

Fair use is vital to the growth of knowledge and can apply to a full range of materials and activities. Educational purpose alone does not automatically make a request fair use because each of the factors must be analyzed in order to conclude whether or not an activity is lawful. Fair use was designed by Congress to be flexible and adaptable to changing needs and circumstances. The law provides no clear and direct answers about the scope of fair use and its meaning in specific situations.

Despite this inherent flexibility, two specific 1980s court decisions concerning unpublished manuscripts have threatened the use of fair use. The first was the 1985 US Supreme Court ruling in *Harper & Row v. Nation Enterprises* (471 U.S. 539), which determined that the scope of fair use for unpublished materials is narrower than the scope for published works.<sup>3</sup> In 1987, even though traditional fair use allowed the right to quote from other materials specifically for purposes of research, scholarship, and education, the Second Circuit Court ruled in *Salinger v. Random House* that a creator/author could prohibit most uses of his unpublished letters even if deposited in archives. The court even excluded not only reprinting and quoting from the unpublished letters but also the detailed paraphrasing of the material. The court ruled that the original author who deposited their material in the archives only lost control of disposition, but retained full copyright of the material. This decision does not prevent donors from depositing papers within

archives or archives from providing access to these materials, but it does affect how the archives' clientele can use the material.

For archivists, the challenge now is determining what is fair use. For unpublished manuscripts, fair use now depends not only on the four factors of fair use but also on the circumstances of the material in question. Privacy must now be taken into consideration along with copyright, even though they have directly conflicting purposes; while privacy laws want to protect confidential material, copyright seeks to promote the growth of knowledge. Archivists must adopt new strategies to control intellectual barriers in the aftermath of these recent changes.

As this brief overview shows, copyright law is a complex and ever-changing issue.<sup>4</sup> Because copyright applies to nearly every document, archivists and librarians need to be copyright leaders in their institutions, working to establish institution/organization-wide policies. These policies must include support and direction from librarians and archivists who are teaching and tutoring.

Institutional copyright policies must spell out two principles: first, all materials are generally under copyright, including unpublished materials, and second, there are some exceptions with fair use limitations. The policy must also identify specific responsible parties tasked with handling copyright for the institution or organization. These responsible parties must interact with legal counsel and must differentiate between professional association guidelines and actual law. Archivists and librarians must rely on law and not guidelines developed by organizations acting in their own organizational interests.

Institutional copyright policies must enumerate fair use privileges, asserting the full right of fair use allowed in each case. In order to accomplish this, policies need to be written to accommodate the grayness of copyright law, especially as regards fair use. In addition, the policy must include guidelines for seeking permission when fair use is not an option. There is no fast and easy answer for copyright; each case is different, and each case must be understood in its own context. For example, just because the ruling concerning unpublished material threw doubt/concern with fair use does not mean that it applies to every piece of unpublished material. Also, because of copyright expiration guidelines, the unpublished writings of authors will enter the public domain seventy years after death. These developments place another burden on archivists and librarians to record the deaths of writers represented in their collections.

In addition, when archivists or librarians obtain materials, they need to understand that they are not receiving the copyright of the material. Therefore, we must include a written form concerning copyright for all gifts and purchases, detailing the copyright provisions that will be in place upon the author's death. Archivists and librarians should then record this copyright ownership with the collection. Although, if copyright is not obtained with the collection, on the death

of the author/creator the copyright may be equally divided and shared among multiple beneficiaries.

In the digital age, it is important to remember that every user is a creator and every creator is a user; therefore, it is imperative for libraries to be more than just gatekeepers of information. Librarians and archivists are called to help their clientele think critically, ask questions, foster creativity, and create (or simply foster re-creation of) information. We have a responsibility to protect and help clientele understand copyright. We have a duty to obey the law and to protect the agreed rights of the donors, but we also have a responsibility to make the collection as useful as possible for the clientele. Only by identifying and using fair use can we better fulfill both of these obligations.

For additional information and developments, please visit the following resources:

- [www.arl.org/copyright-timeline/](http://www.arl.org/copyright-timeline/)
- [www.copyright.gov](http://www.copyright.gov)
- [www.librarycopyright.net](http://www.librarycopyright.net)
- [www.copyrightoncampus.com](http://www.copyrightoncampus.com)
- [www.copyright.com](http://www.copyright.com)
- [www.creativecommons.org/find/](http://www.creativecommons.org/find/)
- [www.sxc.hu](http://www.sxc.hu)

This section has described an unsolved problem for digital libraries—specifically, copyright law in the United States. Additionally, many issues of digital libraries will involve legal issues beyond copyright in the United States, such as international copyright laws, elaborate contracts, and technological protection software. Could you envision if there was a standard legal framework that was reasonable and straightforward to implement throughout the world?

## *Embracing the Future*

Nearly a century after H. G. Wells unveiled his vision of a “world brain” in 1938, we are closer than ever to seeing that vision realized. Technological breakthroughs, combined with the social need for universal information access, have driven society to look for solutions to economic and legal barriers still facing digital libraries, spurring innovation and changes to copyright laws. There is still much work to do.

Digital libraries are transforming scholarship and research practices by increasing accessibility to materials that would otherwise not be accessible, by having no physical boundaries and providing around-the-clock availability, by

allowing researchers and scholars to use any search term to investigate patterns in large amounts of text through friendly interfaces while increasing speed and accuracy of research, and by allowing print-disabled users to use technologies to read scanned books. In other words, digital libraries increase the preservation and conservation of resources while decreasing the physical space needed to store the same number of resources, as well as decreasing the cost of maintaining a digital library over a traditional library, and finally increasing the networking of resources across other digital libraries. Digital libraries have also spurred on innovation within the library, challenging librarians to learn new skills for learning, research, and creation. This has caused an increased focus on learning and development for digital learning, resulting in a shift from teaching and supporting information literacy face-to-face towards digital teaching and support. In digital libraries, innovation will also lead to even more advancements within data management, resulting in more accurate search results significantly improving the way researchers and scholars discover content.

This ever-expanding web of networked research nodes forms the “world brain” that Wells referred to in 1938. Digital libraries are a signature example of how research libraries have evolved beyond thinking of the isolated ivory tower. Expectations are shifting, and people want resources to be collectively held and available for all.

Innovations in digital libraries will, in turn, have an impact on the physical libraries. Clearly, there will be less public shelf space and more collaborative learning spaces, and the design of the library will be to better facilitate face-to-face interactions as well as digital learning interactions. This may require new, innovative technology that facilitates active learning spaces, media productions, virtual meeting spaces, etc. Additionally, partnerships with other areas of the institution (writing centers, instructional design, information technology, etc.) will need to be established to meet the needs of scholars and researchers. We can't be certain how the growth of digital material will manifest in the future; what is clear, however, is that digital libraries are changing the ways libraries are being used forever.

\* \* \*

*Disclaimer: I am not a lawyer and this should not be considered legal advice. You should seek appropriate counsel for your own situation. And please note, the section on legal challenges is directed toward readers in the United States. If you are conducting business outside the United States, I highly encourage you to find and understand your obligations regarding copyright and legal obligations for digital libraries.*

## Works Cited

- AntonBergen. 2008. "Medieval Helpdesk in English." *YouTube*, February 23, 2008. [bit.ly/30lDne7](http://bit.ly/30lDne7).
- American Library Association. 2019. "Copyright for Libraries: General Information." Last updated March 21, 2019. [libguides.ala.org/copyright](http://libguides.ala.org/copyright).
- Howard, Jennifer. 2017. "What Happened to Google's Effort to Scan Millions of University Library Books?" *EdSurge*, August 10, 2017. [bit.ly/2Th2xJv](http://bit.ly/2Th2xJv).
- Lesk, M. E. 2005. *Understanding Digital Libraries*. New York: Elsevier.
- . 2012. "A Personal History of Digital Libraries." *Library Hi Tech* 30: 592–603.
- Lessig, Lawrence. 2004. *Free Culture: How Big Media Uses Technology and the Law to Lock Down Culture and Control Creativity*. New York: Penguin Press.
- Metz, Cade. 2019. "Google Claims a Quantum Breakthrough That Could Change Computing." *New York Times*, October 23, 2019. [www.nytimes.com/2019/10/23/technology/quantum-computing-google.html](http://www.nytimes.com/2019/10/23/technology/quantum-computing-google.html).
- Norris, M., C. Oppenheim, and F. Rowland. 2008. "The Citation Advantage of Open-Access Articles." *Journal of American Society for Information Science and Technology* 59, no. 12: 1963–72.
- Ogden, Cody. 2019. "Killed by Google." [killedbygoogle.com/](http://killedbygoogle.com/).
- Savitz, Eric. 2011. "Amazon Says Now Selling More E-Books Than Print Books." *Forbes*, May 19, 2011. [bit.ly/2FQwcRW](http://bit.ly/2FQwcRW).
- US Congress, Office of Technology Assessment. 1986. *Intellectual Property Rights in an Age of Electronics and Information*. OTA-CIT-302. Washington DC: US Government Printing Office.
- Wells, H. G. 1938. *World Brain*. Garden City, NY: Doubleday, Doran & Co.

## Notes

1. For example, see [place.asburyseminary.edu/firstfruits/](http://place.asburyseminary.edu/firstfruits/).
2. For more information on fair use, please see [www.copyright.gov/fair-use/more-info.html](http://www.copyright.gov/fair-use/more-info.html).
3. For additional information, see [supreme.justia.com/cases/federal/us/471/539/](http://supreme.justia.com/cases/federal/us/471/539/).
4. For additional copyright development and information, see [www.arl.org/copyright-timeline/](http://www.arl.org/copyright-timeline/).