Google Books: Mass Digitization and the Implications for Public Libraries

Toni Beaton[1]

[1]*Department of Information Science, Dalhousie University*, tn228036@dal.ca

Abstract

Google Books has become a forerunner in the process of mass digitization of turning physical books into online books. With more than 40 million books digitized, Google Books is an online platform that enables worldwide access to a vast variety of literature (Lee, 2019). Although this platform has created an opportunity for more free and open access to books for the public, Google Books has been met with copyright lawsuits from authors around the globe. This research paper covers a brief history of this mass digitization project and considers how Google Books has impacted the traditional services of public libraries, including reference, interlibrary loans, and collection development decision-making. In an analysis of the content, quality, and accessibility of online resources on Google Books, it is recommended that the platform is used only as a complementary resource to a public library's collection, not as a full replacement for online library collections.

Keywords: mass digitization, Google Books, copyright, collection development, public libraries, open access, metadata, interlibrary loans

G oogle has become so ubiquitous within society that this company's namesake has become a verb: "Google it." Integrated platforms such as Google Search, Google Maps, and Google Scholar provide a seamless search experience for the end user. Twenty years ago, Google released its newest project to the public: what would become known as 'Google Books,' an online, keyword-searchable index. Google's mass digitization of academic and public library holdings would enable worldwide access to millions of digitized books, including those within the public domain and copyrighted works. Although Google Books brings many advantages to the end user such as 24/7 online access, the concept was fraught with copyright and legal

issues, resulting in years of ongoing legal battles between Google, authors, and publishers.

Technological innovations have reached far and wide, and public libraries are not immune to this shift to an online society. Public libraries have also innovated to remain relevant, changing their collections to meet patron needs and values. Google Books cannot be ignored as a megalith of online eBook content; however, public librarians, especially those in charge of information resource and collections management decisions, must consider the potential impact of Google Books on a public library's holdings. This research paper aims to highlight that while Google Books may be used as a supplemental resource for public librarians to consult in their collections management decisions, it should not act as a replacement to library collections due to its lack of quality control, privacy and accessibility issues, restrictions on copyrighted books, and inherent ephemeral quality as a digital platform.

Google Books: A Brief History

Google Books (initially marketed as Google Print, and subsequently, the Google Books Library Project), founded by Stanford graduates Sergey Brin and Larry Page, publicly announced its mass digitization project in 2004 (Johnson, 2018; Lackie, 2018; Willett, 2009). Google's original mission statement for this project was to "…digitize every book ever published, index them, and make them fully searchable online for free" (Lackie, 2008, p. 37). Although a seemingly tall order, Google Books is well on its way to achieving this goal with more than 40 million books digitized (Lee, 2019), although the exact current number remains undisclosed. Mass digitization is the "…conversion of whole libraries without making a selection of individual materials" (Coyle, 2006, p. 641). As such, there is no careful selection criteria by Google of which materials are digitized. The goal is for quantity over, in some cases, quality (Johnson, 2018). Google Books is not a producer of content, nor is it posited as a preservation effort; it relies on content from established libraries and aims to be a search index of freely accessible copies of online books (Coyle, 2006; Dougherty, 2010).

The Google Books Project although in many ways dynamic and exciting, has also been highly controversial (Lackie, 2008). When Google first announced their intentions of mass digitization, many authors raised concerns and publicly criticized the

company. Google has been subjected to much scrutiny, including from authors who have sued the company for blatant copyright violations (Byrne, 2012). One legal case stated that Google Books was in violation of fair use dealings, claiming "Google's unauthorized uses are for a commercial purpose; involve verbatim copying, distribution and display of protected expression; are not transformative, and… [may] adversely affect actual and potential markets for copyrighted books" (Byrne, 2012, p. 1). In contrast to online content providers such as JSTOR or Muse, Google Books is "…the only mass digitization project to undertake digital conversion of copyrighted material" (Doyle, 2006; Koehler, 2008; Willet, 2009, p. 143). In the beginning of its mass digitization efforts, Google Books did not discriminate between copyrighted material and material that existed within the public domain. The Authors Guild and the Association of American University Presses sued Google in 2005 for violating copyright law (Lackie, 2008). However, in 2013, a federal district court "ruled in favor of the internet giant" after years of ongoing legal battles (The Economist Newspaper, Inc., 2015, p. 1). In light of these court cases, Google Books has been forced to make policy changes regarding copyrighted materials.

Copyright issues and fair use dealings are complex and fluid. Fair use dealings are defined by copyright laws, which are based on court cases and traditional court rulings (Koehler, 2008). As Google Books includes public domain and copyrighted publications, complexities arise involving author rights (Willett, 2009). One way Google Books gets around digitizing copyrighted material is to only showcase "snippets" to the end-user (Willett, 2009). These snippets are usually only a few pages, or several lines of material on a page. Under the agreement between Google, the American Association of Publishers, and the Authors Guild, "end users [are] able to view up to 20 percent of copyrighted materials" (Willett, 2009, p. 145). Although there are some who question Google's intentions and claim the company is "taking fair use for a ride" (Lackie, 2008, p. 37), Google asserts that as only "snippets" of copyrighted works are shown, including these works in the index increases the work's visibility and discoverability.

Google Books has revolutionized the accessibility of books online. Google claims the "ability to introduce millions of users to millions of titles can only *expand* the market for authors' books, which is precisely what copyright law is intended to foster" (Dames,

2006, p. 19, emphasis added). Interestingly, a study by Nagaraj and Reimers (2023) developed a theoretical framework that demonstrates mass digitization projects, such as Google Books, have the potential to significantly boost demand for physical versions of resources, providing evidence to support Google's claim of nearly 20 years ago. Results of this study found that "digitization increases sales by 4.8 percent and increases the likelihood of at least [one] sale by 7.7 percent…" (Nagaraj & Reimers, 2023, p. 431). However, it's important to note the digitized books used in this study sample were older, in the public domain, less well-known, or published by small or independent publishers. Therefore, studied effects of "cannibalization" on physical sales for digitized copyrighted materials are still unclear (Nagaraj & Reimers, 2023, p. 456).

Digitization Limitations

Despite the grand concept of a massive, open access index that is freely searchable worldwide, digitization does have its limitations, particularly on accessibility and quality control. During the digitization process, books are scanned page by page from cover to cover with the resulting scans displayed in Google Books as images on a webpage. From a user experience perspective, since the pages of each digitized book are simply images, "…there is no capability to highlight and copy text from the page, and [copyrighted] books cannot be downloaded to be read offline" (Coyle, 2006, p. 643). This presents accessibility issues, particularly for users who do not have access to consistent or reliable internet use. Without the ability to connect to Google Books, users will not be able to search, preview, or read any of the digitized copyrighted materials.

Additionally, users who rely on screen readers or other aids to access online materials may not be able to interact successfully with the digitized content on Google Books. In certain cases, such as when a book is within public domain (e.g., Shakespeare's *Hamlet*), there is the option on Google Books to "View as plain text," which converts the digital images to a plain text document. A screen reader could then assist a visually impaired or partially sighted person in accessing this text (Kerscher, 2007). However, this conversion option is not available for copyrighted material, either in the fuller preview or the limited "snippet" view, which drastically reduces the number of books accessible to users with visual impairments. Google Books also neglects to include external links to accessible high-quality versions of eBooks, such as from a

library like Recording for the Blind and Dyslexic (RFB&D) or another library within the Digital Accessible Information System Consortium, which serves people with visual impairments (Kerscher, 2007; The DAISY Consortium, 2024).

There have also been criticisms made of Google Books' metadata, referred to first as "The Metadata Mess," and then later as "A Metadata Train Wreck" (Dougherty, 2010, p. 87). Metadata is commonly known as 'data about data,' such as the description of a book in a library catalogue (e.g., the ISBN, the author, the date of publication, etc.). In the case of Google Books, metadata describes what type of digitized book a user is looking at and provides a back-end description of the book, which helps link the resulting record to a user's search keywords. However, mistakes in metadata can lead to negative consequences for the user experience. For example, problems seen with Google Books' metadata include lack of quality control on scanning and OCR (Optical Character Recognition) processing, misdating of publications, as well as classification errors (Dougherty, 2010). In a study by James and Weiss (2012), 400 book samples randomly extracted from Google Books displayed a 36% rate of error in scan and metadata quality, with the average number of errors per book at 1.97 and some titles showing as many as four errors. This "abysmal" (Jones & Janes, 2010, p. 46) state of metadata negatively impacts a user's ability to search for titles, authors, and dates effectively, and may foster mistrust in the reliability of the Google Books platform.

Privacy Concerns

Users put themselves at a certain degree of risk when searching on Google Books versus searching their own public library catalogue. When users of online information resources (like Google Books) go searching, potentially identifying information is consistently being tracked, such as IP addresses. Google has firmly established itself as a "leader in search" using predictive analytic techniques in user searching (West, 2019, p. 34). According to West (2019), Google then sought to "…monetize the data it collected through search by establishing its dominance in online advertising… [as] it monopolizes so many corners of the web and can leverage a broad pool of user data" (p. 34). Stanford graduates Brin and Page initially presented Google Books to Eric Schmidt, former CEO of Google, not only as a mass digitization project for information resources, but also a business project, "since the vast amounts of

information Google could extract from books would improve Google's ability to deliver information… eventually also result[ing] in an increase in traffic and clicks on ads" (Thylstrup, 2019, p. 40). Once user data has been collected (via the process of browsing through books or searching for texts on Google Books, for example), the data is "often shared or sold to third-parties for advertising revenue" (Weiss, 2021, p. 1789). Using mass digitization as a business technique aligns well within the concept of "big data," by capitalizing on user information-seeking behavior through data-mining techniques (Thylstrup, 2019; Zuboff, 2019).

In contrast, when users of a public library check out physical or eBooks from the library's holdings, this information is protected and remains private (Jones & Janes, 2010). One may draw a "contextual integrity" comparison between the privacy (or lack thereof) of an online user searching the Google Books interface, and that of an in-person or online user searching the holdings of a public library (Jones & Janes, 2010). Contextually, within a public library, patron records are kept private, confidential, and are not shared. Circulation records may be "…deleted and purged from the system" once a book has been returned (Jones & Janes, 2010, p. 54). As Google Books is offering a type of mirrored framework in its online search index, it should also contextually mirror privacy protections for users. However, this is not the case.

Librarians are bound by a code of ethics to protect the privacy of their patrons. According to the American Library Association, librarians must "…protect each library user's right to privacy and confidentiality with respect to information sought or received and resources consulted, borrowed, acquired or transmitted" (Jones & Janes, 2010, p. 54). Since Google Books has claimed to create a "universal digital library" (Jones & Janes, 2010, p. 51), it should uphold the privacy and rights of a traditional public library. The information transmitted by users on Google Books is considerably detailed; as Jones & Janes (2010) describe, "…information is revealed about what a given reader is either reading or thinking about reading… potentially at the level of pages or even words… much more closely tracked than that which they might transmit in public libraries" (p. 57). Nearly all user behavior – searching, selecting, viewing, and downloading books – is tracked by Google (Jones & Janes, 2010). The default settings on Google will track the information-seeking behaviors of users, including searches and

clicks, and use the data to "target advertisements and refine search results" (Vaidhyanathan, 2012, p. 89). As outlined by Weiss (2021), Google Books can easily track what eBooks users search for when using the platform, "eroding the user privacy that remains a central tenet of libraries" (p. 1783). Whereas librarians are bound by professional ethics regarding patron confidentiality, there are no such ethics governing this mass corporation (Jones & Janes, 2010). This becomes problematic, especially since users of Google Books may not even realize their every search is being tracked.

Current Implications for Public Libraries

Interlibrary Loans

Google Books is clearly a massive resource for open access to electronic books, and should be factored into consideration by public librarians, particularly regarding interlibrary loans, reference services, and information resources management. Interlibrary loans (ILL) are a way in which libraries share access to resources, not only locally, but also globally. The convenience of ILL makes it possible for patrons to borrow resources from another library without ever having to leave their home branch. However, as Google Books provides an index of searchable digitized publications online, does the need for ILL diminish? Based on statistics gathered by approximately 112 members of the Association of Research Libraries in the United States, the answer appears to be no (Willett, 2009). This study, focusing on the years 1986 through 2006, showed a continued increase in ILL statistics. Interestingly, requested books were found to be publications from 2000 or later, which are copyrighted material in Google Books (Willett, 2009). Unless users happen to find the perfect quote within a limited two-to-three-page copyrighted publication on Google Books, as seen in the limited preview or "snippet" view, the need for ILL remains (Willett, 2009).

Additionally, Google Books has integrated a search option into their platform for users to find a book in a local library (Lackie, 2008). The Google Books 'Advanced Search' page allows users to limit their search so that results indicate whether their local public library has access to a physical or online copy of the book (Lackie, 2008). In a study on the functionalities of Google Books, Chen (2012) discovered the success rate of the "Find in a library" link to a local library catalogue – in Chen's case study, the

Mirlyn Catalogue at the University of Michigan – works 74.8% of the time. Although this success rate showcases room for improvement, embedding the "Find in a library" feature within the Google Books platform increases visibility of the holdings of public libraries and encourages users to seek full access to copyrighted material at their local library. In turn, this increases the chance of either a direct request by the patron to place a hold on the material, or, to request an ILL.

Reference Services

Reference services are the ways in which library professionals interact with patrons and are considered one of the cornerstones of quality library services. Reference can include user guidance regarding selection of information sources, instruction of information literacy skills, and either "…finding the required information on behalf of users, or assisting users in finding information" (Weiss, 2016, p. 286). Patrons have a certain level of expectation that a librarian can accurately and efficiently assist them with their information needs. This type of interaction is not available on mass digitization platforms such as Google Books. As an index, Google Books is not equipped with any kind of reference service. Users must navigate the platform on their own, without the guidance or expertise of reference librarians. Keyword searching becomes critically important, which is problematic when incorrect metadata are associated with the online records.

With digital innovations comes the necessity for public libraries to innovate their reference services. Public libraries have created ways to emphasize the usefulness of reference service, even in a digital era. These innovations include 24/7 live chat reference service, developing "roving digitally equipped reference librarians" (Weiss, 2016, p. 287), as well as providing online instructional reference through virtual platforms. Public library reference service has been largely influenced by Ranganathan's five laws of librarianship (Weiss, 2016). In our age of technology, librarians need to focus a little more on Ranganathan's fourth law: "Save the time of the user" (Weiss, 2016, p. 290). In a fast-paced, digital world, patrons value efficiency and convenience. Library professionals still provide their expertise in finding the right book for the patron, that is, Ranganathan's "Every reader his book" and "Every book his reader"; however, in an era of digitization, adapting to online reference services is key

for library professionals to remain relevant and continue to prove their usefulness (Weiss, 2016).

When providing services regarding digitized publications, such as those found in Google Books, it is important to note several areas of concern for reference librarians. Due to the poor quality of metadata, as well as the "…high rate of blurred, turned, slanted, and obscured pages" (Weiss, 2016, p. 300), there is a high potential for negative experiences for library patrons. Patrons rely on information professionals to consistently provide recommendations that are of high quality. If librarians recommend a resource on Google Books that contains errors, such as, "…pictures of scanner's fingers, [or] lost text from books held incorrectly" (Johnson, 2018, p. 218), this runs the risk of damaging the trustworthiness of reference service interactions (Weiss, 2016). Public libraries, and reference librarians in particular, need to ensure they are recommending quality resources to library patrons.

Information Resources Management Decisions

Google Books has allowed online users to search for books that may have previously been out of reach, due to the rarity or fragility of the book, the language of the book, or the book's physical geographic location (Weiss, 2016). Online digitized copies of books are now readily available at the click of a search. This digitization trend has impacted public libraries' consideration of collection development and holdings (Weiss, 2016). Public libraries can take advantage of the Google Books platform by providing online 24/7 access to digitized books, as well as accessing rare book collections for their patron's use (Weiss, 2016). Google Books is, without a doubt, a relevant information resource for public libraries. However, there are several factors that information resources managers or collections management librarians must consider when utilizing Google Books.

The digitized holdings within Google Books are linked with WorldCat, a worldwide online catalogue of almost 560,000,000 public, academic, special, and research library bibliographic records (OCLC, 2024). WorldCat is a powerful search tool and can serve as a starting place for finding digitized resources, including eBooks (Chen, 2012). Collections management librarians can use this connection to their advantage when considering whether to purchase an eBook title for their public library

holdings. If a title is freely available through Google Books, as seen via WorldCat, collections management librarians may then choose to avoid purchasing a digitized copy for their collection. However, it is important for collections development librarians to keep in mind that Google Books has been known to upload blurred, unreadable, or fragmented versions of digitized books; this is something to consider when contemplating whether to purchase a title or rely on the copy in Google Books (Weiss, 2016). As Weiss (2016) states: "[r]elying purely on these [Google Books eBooks] for the sake of saving money or for opening up space in tightening physical [or virtual] spaces would be *compromising* at best and *disastrous* at worst" (p. 300, emphasis added). Just because Google Books has digitized a book does not mean that book is of a quality acceptable to the standards of a public library. It is important that collections management librarians take the time to review a resource on Google Books and check for quality control *before* deciding whether it may serve as a freely available alternative to a purchased eBook title for the library catalogue.

Additionally, digital collections have been known to be ephemeral and have less permanence than physical collections. If Google Books were to decide to restrict access or to delete digitized works, and public libraries have relied on those as a supplement to their collections, this could be disastrous (Evans & Saponaro, 2019; Weiss, 2016). Relying solely on a digitized index has massive implications for public library holdings, particularly regarding perpetual access (Johnson, 2018). When making collections management decisions it is wise to remember that "[a] copy on a library shelf… will be fully visible to a reader or lendable through an interlibrary loan agreement" (Weiss, 2016, p. 303). In a comparison study between coverage of resources on Google Books and WorldCat, Chen (2012) discovered that less than 10% of digitized Google Books had free full-text views, and around 15% had limited previews or "snippets" available. This is especially important when considering publications under copyright, as Google Books will only show "snippets" of copyrighted texts, whereas library patrons will usually require the full text (Willett, 2009). As Jones (2017) states, "[m]aking a book *digital* does not necessarily mean making it fully *available*" (p. 251, original emphasis). When considering library holdings, collections managers must put the needs of their local patrons first. Deciding not to purchase a book, or removing a book from the collection,

just because the published work has been digitized, may not serve in the best interests of library patrons (Weiss, 2016).

Looking Ahead: Artificial Intelligence and Google Books

When deciding whether to rely on eBooks indexed in Google Books as a supplement to a public library's virtual collection, collections management librarians may also wish to consider the impact of artificial intelligence on the (un)reliability or (dis)trustworthiness of Google Books. For example, there is a newfound dilemma occurring between the Google Books platform and the artificial intelligence (AI) tool ChatGPT. A detailed explanation of ChatGPT is beyond the scope of this paper, but a brief background will help contextually situate ChatGPT in relationship with the Google Books platform. ChatGPT is a Large Language Model (LLM) which generates coherent, grammatically correct text on almost any topic (Lametti, 2023). It has been pre-trained for conversation and adopts a very human-like speech (or written script) in its responses to prompts. ChatGPT was freely released to the public in November of 2022, but that version (GPT-3.5) does not perform to the same standards as GPT-4 (launched in March 2023), which is currently available only on a paid subscription (OpenAI, 2024).

Despite the pre-trained nature of ChatGPT as a human-like, conversational generative tool, it does fall short in its current capabilities in generating creative text. For example, Johansson (2023) compared human-written and ChatGPT-generated texts and concluded that while ChatGPT can produce appropriate context-based texts (in this case study, an academic essay for an English literature course), it lacks the "…depth, specificity, and accurate source referencing present in human-generated text" (Johansson, 2023, p. 4). ChatGPT has also been known to "hallucinate" when prompted to list citations, generating unreliable, false citations that reference texts which do not exist (Lametti, 2023). In its current capabilities, ChatGPT lacks the critical thinking skills and the nuanced characteristics of human-authored texts when generating written content. There are some 'dead giveaways' that flag whether a low-quality text has been written using ChatGPT, such as: lengthy introductions, generic thoughts or opinions, and even signature phrases like "as of my last knowledge update" (Cook, 2023; David, 2024).

A plethora of low-quality texts produced by ChatGPT have flooded the internet, appearing on blog posts, websites, online marketplaces, and even on Google Books. As mentioned previously, Google Books focuses on quantity over quality, and does not appear to have any type of selection criteria as to which books are indexed and digitized. According to 404Media, Google Books is now indexing eBooks which appear to have been written by AI tools (David, 2024). When searching on Google Books using the phrase "as of my last knowledge update," 404Media reporters found dozens of low-quality books written by AI tools, like ChatGPT or Google's Gemini, that appeared to trawl Wikipedia for generic information, such as the text *Bears, Bulls, and Wolves: Stock Trading for the Twenty-Year-Old* (David, 2024). The fact that Google Books does not currently differentiate between human-created and AI-generated eBooks indexed on the platform is alarming. The Director of Research at the Distributed AI Research Institute, Alex Hanna, has concerns that Google Books may become an unstable and untrustworthy platform as a result of indexing AI-generated eBooks: "AI-generated content will be pulled into Google Books, and Google will use that content to train new AI models; it's like an Ouroboros…" (MetaNews, 2024, n.p.). Users of Google Books, then, must not only be wary of privacy concerns, quality control, metadata, and accessibility issues when browsing the platform for eBooks; they must now also consider whether a seemingly reputable text has been authored by AI tools like ChatGPT.

Conclusion and Future Considerations

Google Books is an influential information resource, reaching worldwide and impacting informational settings such as public libraries. Not only does it help obscure, rare books find new audiences, Google Books also "…go[es] a considerable distance toward making access to… information more egalitarian worldwide" (Jones & Janes, 2010, p. 45). Online availability to tens of millions of digitized books is an exemplar of worldwide open access. However, public librarians, especially those in charge of handling collections development decisions, must consider the weaknesses of mass digitization. Google Books can – and should – be used as a *complementary* resource to the eBook collection but should not be used in *place* of a library's own holdings. Reference librarians must ensure that, if referring publications from Google Books to

their local library patrons, each text has been checked for any errors in metadata or scanning. This way, Google Books may be used as a supplemental addition to the library's holdings, if publications contain accurate metadata with no loss of quality to the scanned pages.

A consideration for future research is how Google's newest artificial intelligence, Gemini (previously known as Bard), may shape the Google Books interface and change the way users interact with the platform. For example, Google has already dabbled in experimental artificial intelligence, such as the GoogleAI Semantic Experiences "Talk to Books" initiative, which was broadcast in 2018 as "…an experiment… in using a neural language model to search through a large collection of books… delight[ing] millions of users who found that you really could converse with books in a way never imagined" (GoogleAI, n.d., n.p.). The "Talk to Books" experiment successfully "proved the effectiveness of language models," going on to shape other Google software products (GoogleAI, n.d., n.p.). Investigating the potential effects of artificial intelligence on mass digitization projects, and the implications for end users and library patrons, may be an intriguing area for further research.

References

Byrne, A. (2012, August 6). Authors sue Google over book digitization project. *Inside Counsel Breaking News.*

Chen, X. (2012). Google Books and WorldCat: A comparison of their content. *Online Information Review, 36*(4), 507-516. https://doi.org/10.1108/14684521211254031

Cook, J. (2023, December 6). *6 giveaway signs of ChatGPT-generated content.* Forbes. https://www.forbes.com/sites/jodiecook/2023/12/06/6-giveaway-signs-of-chatgpt-generated-content/

Coyle, K. (2006). Mass digitization of books. *The Journal of Academic Librarianship, 32*(6), 641-645. https://doi.org/10.1016/j.acalib.2006.08.002

Dames, K. M. (2006, March-April). Library organizations should support Google Book Search. *Online Magazine*, 18-19.

David, E. (2024, April 5). *Google Books reportedly indexing bad AI-written works.* The Verge. https://www.theverge.com/2024/4/5/24122077/google-books-ai-indexing-ngram

Dougherty, W. C. (2010). The Google Books project: Will it makes libraries obsolete?
The Journal of Academic Librarianship, 36(1), 86-89.
https://doi.org/10.1016/j.acalib.2009.12.002

GoogleAI. (n.d.). *Talk to books.* Semantic Experiences.
https://books.google.com/talktobooks/

James, R., & Weiss, A. (2012). An assessment of Google Books' metadata. *Journal of Library Metadata, 12*(1), 15-22. https://doi.org/10.1080/19386389.2012.652566

Johnson, P. (2018). *Fundamentals of collection development and management* (Fourth edition.). ALA Editions.

Johansson, I-R. (2023). *A tale of two texts, a robot, and authorship: A comparison between human-written and a ChatGPT-generated text* [Unpublished bachelor's thesis]. Malmö University. https://www.diva-portal.org/smash/get/diva2:1767661/FULLTEXT03

Jones, E. (2017). The public library movement, the digital library movement, and the large-scale digitization initiative: Assumptions, intentions, and the role of the public. *Information & Culture, 52*(2), 229-263. https://doi.org/10.7560/IC52205

Jones, E. A., & Janes, J. W. (2010). Anonymity in a world of digital books: Google Books, privacy, and the freedom to read. *Policy & Internet, 2*(4), 43-75. https://doi.org/10.2202/1944-2866.1072

Kerscher, G. (2007, November). *First step in adding accessibility to Google Books: Was it enough?* Braille Monitor. https://nfb.org/images/nfb/publications/bm/bm07/bm0710/bm071003.htm

Koehler, W. (2008, June). Good and evil in the garden of digitization: Google and fair use. *Searcher, The Magazine for Database Professionals,* 24-57.

Lackie, R. J. (2008). From *Google Print* to *Google Book Search*: The controversial initiative and its impact on other remarkable digitization projects. *The Reference Librarian, 49*(1), 35-53. https://doi.org/10.1080/02763870802103654

Lametti, D. [Maple League of Universities]. (2023, January 31). *AI and academia: The end of the essay?* [Video]. YouTube. https://www.youtube.com/watch?v=geKsTy8QKhY

Lee, H. (2019, October 17). *15 years of Google Books.* Google: The Keyword.
https://blog.google/products/search/15-years-google-books/

MetaNews. (2024, April 7). *Dozens of AI-generated books found on Google Books.*
https://metanews.com/dozens-of-ai-generated-books-found-on-google-books/

Nagaraj, A., & Reimers, I. (2023). Digitization and the market for physical works:
Evidence from the Google Books Project. *American Economic Journal: Economic
Policy, 15*(4), 428-458. https://doi.org/10.1257/pol.20210702

OCLC. (2024). *Inside WorldCat.* WorldCat. https://www.oclc.org/en/worldcat/inside-worldcat.html

OpenAI. (2024). *Pricing.* ChatGPT. https://openai.com/chatgpt/pricing/

Saponaro, M. Z., & Evans, G. E. (2019). *Collection management basics* (Seventh
edition.). Libraries Unlimited, an Imprint of ABC-CLIO, LLC.

The DAISY Consortium. (2024). *Global reach.* The DAISY Consortium: Creating the
Best Ways to Read and Publish. https://daisy.org/

The Economist Newspaper Ltd. (2015, October 20). Google Books wins a court battle:
Words for the masses. *The Economist (Online).*

Thylstrup, N. B. (2019). The trials, tribulations, and transformations of Google Books.
In *The politics of mass digitization* (pp. 38-55). The MIT Press.
https://doi.org/10.7551/mitpress/11404.003.0005

Vaidhyanathan, S. (2012). The googlization of us: Universal surveillance and
infrastructural imperialism. In *The Googlization of Everything* (1st ed., pp. 82-114). University of California Press.

Weiss, A. (2016). Examining massive digital libraries (MDLs) and their impact on
reference services. *The Reference Librarian, 57*(4), 286-306.
https://doi.org/10.1080/02763877.2016.1145614

Weiss, A. (2021). Massive digital libraries (MDLs) and the impact of mass-digitized book
collections. In M. Khosrow-Pour (Ed.), *Encyclopedia of information science and
technology* (5th ed., pp. 1782-1795). IGI Global. https://doi.org/10.4018/978-1-7998-3479-3.ch123

West, S. M. (2019). Data capitalism: Redefining the logics of surveillance and privacy.
Business & Society, 58(1), 20-41. https://doi.org/10.1177/0007650317718185

Willett, P. (2009). Mass digitization and its impact on interlending and document supply. *Interlending & Document Supply, 37*(3), 143-148. https://doi.org/10.1108/02641610910985620

Zuboff, S. (2019). Surveillance capitalism and the challenge of collective action. *New Labor Forum, 28*(1), 10-29. https://doi.org/10.1177/1095796018819461