



Authors Guild v. Google Inc:

The 2015 2nd U.S. Circuit Court of Appeals Decision and Its Effects on Mass Digitization Initiatives at The Metropolitan Museum of Art

A Capstone Project by Erin Cathleen Zysett



Authors Guild v. Google Inc: The 2015 2nd U.S. Circuit Court of Appeals Decision and Its Effects on Mass Digitization Initiatives at The Metropolitan Museum of Art

A Capstone Project by Erin Cathleen Zysset

Abstract

There is a growing trend among larger museums and heritage organizations to digitize their collections, both their artifact collections, and their collections of research publications and books. Some of these printed works are still under copyright, and some are in the public domain. Of those copyrighted works, some are owned by the organization, and some are not. This creates some significant legal hurdles for organizations wanting to digitize large collections on a massive scale. By looking at the recent legal case Google v. Authors' Guild and the mass digitization projects at the Metropolitan Museum of Art, this Master's degree capstone study examines the emerging issues of mass digitization projects, and the marketing potential of making collections available online.

KEYWORDS: ebook, Mass Digitization, Google Books, Museum Publishing, *Ecurators*

Table of Contents

Advisor Approval

Acknowledgements

SECTION 1 5

Introduction

- 1.2 Conceptual Framework
- 1.3 Research Questions
- 1.4 Definitions
- 1.5 Methodology
- 1.6 Limitations and Delimitations

SECTION 2 18

Google Books, Museum Publishing, and Mass Digitization of Libraries:

Understanding the Current Publishing Environment

- 2.1 The Google Print Program: Libraries and Publishers
- 2.2 Authors Guild v. Google Inc: The 2015 2nd U.S. Circuit Court of Appeals Decision (Google’s Legal Trouble: a handy timeline)
- 2.3 Questions that came out of the lawsuits

SECTION 3 25

The Digitization Initiatives at The Metropolitan Museum of Art

- 3.1 The Role of Nonprofit Publishing House vs. a For-profit Publishing House
- 3.2 The Watson
- 3.3 The Watson and Google
- 3.4 MetPublishing and Google

SECTION 4 34

Challenges and Benefits of the ebook

- 4.1 What About Big Brother: Privacy Issues and Google Books
- 4.2 Role of the print book in a digital world
- 4.3 Benefits of the ebook
- 4.4 The future of the ebook in Museum Publishing

SECTION 5 39

Resources for ebook creation: What can smaller organizations glean from what the Met is doing

- 5.1 Design Software
- 5.2 What staff resources are needed
- 5.3 Reflowable vs. Fixed Text: when to use them

- 5.3.1 The good, the bad, and the unintelligible
- 5.4 Marketing and Distribution Platforms
- 5.4.2 Distribution
- 5.5 What are the potential difficulties
- 5.5.1 Securing content
- 5.6 What are the potential benefits

SECTION 6 53

In Conclusion

APPENDIX ONE: Ebook Creation Resources 54

- Self-Publishing and Print Options (all encompassing)
- Ebook Only Publishing Platforms
 - Kindle Direct Publishing
 - Nook Press
- “Borrowing” Platforms
 - Scribd
- How To Design Tutorials
 - Adobe Help
 - Lynda.com
 - Terry White’s Tech Blog
- Securing documents
 - Inkwell Editorial

REFERENCES 58

Advisor Approval

Approved by:

A handwritten signature in black ink, appearing to read 'Patricia Lambert', written over a horizontal line.

Dr. Patricia Lambert
Arts Administration Program

June 1, 2016

Date

Acknowledgements

I'd like to thank my advisors Dr. Patricia D. Lambert and Dr. John Fenn for their support and patience as I sort out my path through graduate school, and for their constant reassurance that I will, in fact, "figure it out".

I'd also like to thank Dr. Kim Sheehan for her Spring 2015 *History and Theory of Communications* class that helped me find a theory framework for my thinking; Matt Bean, J.D., for his Fall 2015 *Copyright Law* class, which was geared at non-law students, and helped me wrap my brain around a very complex subject; and Eric Schiff for his Spring 2016 *Comparative Technologies for Art Administrators* class that allowed me the freedom to explore the practical side of digital publishing. All of which contributed heavily to this final project. Each instructor created a safe space for curiosity and inquiry that helped me find my research path. Thank you.

And finally, I'd like to dedicate this paper to my grandmother who didn't quite make it to see me get my master's, but was always so thrilled that I got to live life on my terms in a way she wasn't able. I love you grams.

SECTION 1 Introduction

There is a growing trend among larger museums and heritage organizations to digitize their collections, both their artifact collections, and their back catalog of research publications and books (Wayne, 2013). During the past four years, the Metropolitan Museum of Art has made almost its entire back catalog of out of print books available for free to the public in *ebook* format via the Google Partners Program (Hauser, Housewright, & Polizzotti, 2012). The Google Partners Program is a digitization program for publishing houses that own the copyrights to current and back catalogs they wish to digitize. Through this program, the Met publishing house, and other publishers, are able to offer researchers *ebooks*, but also print on demand options, meaning a person can request the book be printed, bound, and mailed to them at cost. This program, which makes previously unavailable titles publicly accessible to everyone, is not only in line with the public education goals of most museums and libraries, but also serves as a viable marketing tool that raises the brand profile on the international stage (Wayne, 2013).

While on the surface this may seem like a low cost way to further the organization's mission and raise its profile, there are some very real costs in human capital, software, as well as legal concerns around copyright permissions (Kubis, 2010). *Mass digitization* projects face some daunting issues around gaining permissions for

copyrighted works; both in the sheer number of individual rights holders, and with the issue of *orphan works*, or works that fall under copyright, but whose holder is not available or easily found (U.S. Copyright Office, 2015). Also, *mass digitization* requires special equipment and database know-how that average museums and research libraries may not have available within their existing staff.

As a result, some libraries have chosen to join the Google Library Project (a separate program from the Google Partners Program), wherein Google does the scanning and database creation for the organization, then adds the book to the Google database. Google also makes a full digital copy available to the library to use as the organization sees fit. For many libraries and archives wishing to engage in a *mass digitization* project without paying the upfront costs in staff, software, and time, this project may seem the ideal solution. However, the Google Library Project is not without controversy or detractors, and was the subject of a 2014 lawsuit, which is discussed in Section 2.2 of this paper. The Google Partner Program (where publishers and copyright holders partner with Google to make their catalogs available) has no copyright controversy, because the participants (partners) are the rightful holders of the copyrighted works (Manuel, 2009).

Google has stated that their ultimate goal with both the Libraries and Partner programs is to “work with publishers and libraries to create a comprehensive, searchable, virtual card catalog of all books in all languages that helps users discover new books and publishers discover new readers” (Google, “Perspectives Page,” 2016). A library to rival the legendary library at Alexandria: this is conceived as a place that

anyone with curiosity and an internet connect can learn and view the world's treasures. It's an ambitious vision that few organizations have the resources to create without the aid of a company like Google. In the case of the Digitization Initiatives program at The Metropolitan Museum of Art, a *mass digitization* project is underway in both the research library, and the publishing house, but only one is partnering with Google.

1.2 Conceptual Framework

Three courses were selected to deepen the author's understanding of the issues around the issues facing museums wishing to engage in *mass digitization* of their printed work collections. The work in three classes contributed to the final paper, including: "Copyright Law" and "History and Theory of Communications" through the University of Oregon School of Journalism and Communications (SOJC), and "Comparative Technologies for Arts Administrators" through the School of Allied Arts and Architecture (AA&A).

The curricula in History and Theory of Communications covered the issues around emerging technologies and media content management. Rooted in communication theory, the readings explored how advances in digital technology have affected the sharing of copyrighted material. The Copyright Law class was geared at non-law students and focused on the history of copyright law in the United States and the process courts go through to make decisions in copyright law cases. The goal being to help people entering into the media and content creation industry develop a firmer understanding the role of copyright law in the media industry.

Comparative Technologies offered an opportunity to explore the technical end of *ebook* creation, and allowed for the development of the last section of this paper, which explores the software, marketing, and distribution platforms available to organizations wishing to create a digital catalog of their printed collections.

CAPSTONE

The methodological paradigm of this research is of the interpretivist and constructivist perspective based in subjectivist epistemology.

courses were selected to deepen the author's understanding of the role of copyright law in creating eBooks from existing works, and the resources needed to complete such projects, while contextualizing these topics in terms of mass digitization projects.

“Copyright Law” through the UO School of Journalism and Communications (SOJC)

- case study
- document analysis
- literature review

“History and Theory of Communications” through the UO School of Journalism and Communications (SOJC)

MAIN QUESTION How and where has the Metropolitan Museum of art successfully used Google's programs as part of its Digitization Initiative?

SUB-QUESTIONS how can small, local and regional organizations, leverage advances in digital publishing and the Google Partner Program in the same way at the Met? What are the barriers and benefits of diving into the eBook and print on demand marketplaces?

- website analysis
- document analysis
- literature review
- Software analysis

- website analysis
- video review
- document analysis
- literature review

Comparative Technologies for Arts Administration through the UO School of Allied Arts and Architecture (AAA)

1.3 Research Questions

Most of the research literature on this topic surrounds the Google Library copyright controversy, and the steps being taken by large, well resourced organizations to undertake *mass digitization*. There is very little research on how smaller organizations can benefit from technological advances, and what they might need in staff, time, and technology to enter into the digital world. The main research question for this study was: How and where has the Metropolitan Museum of art successfully used Google's programs as part of its Digitization Initiative? Sub-questions were: How can small, local and regional organizations, leverage advances in digital publishing and the Google Partner Program in the same way at the Met? What are the barriers and benefits for organizations with large print collections wishing to dive into the ebook and print on demand marketplaces?

1.4 Definitions

The following words may be unfamiliar to people who are not involved in the publishing and nonprofit industries. This section is aimed at helping readers have a clearer idea of industry jargon before exploring some of these ideas.

Orphan Work — a creative work with a copyright that has not expired, but for which an owner cannot be located. If a person or company wants to use an *orphan work*, they must document that every 'reasonable means' of locating the holder has been

exhausted. This documentation won't protect the user from the holder coming forward and demanding payment, but it can prevent the user from being saddled with additional fees and fines apart from the back licensing fee.

Mission Driven (variation: mission-driven) — an organization with a clearly stated mission that is the guiding philosophy of all decisions made for the organization by the leadership. Most commonly in the United States *mission driven* organizations are some form of not for profit. This term is used to create a distinction between an organization that exists to further an idea or philosophy and an organization that exists to create a profit for its shareholders or owners,

Loss leaders — a product sold at a loss (it costs more to create than is recouped by the seller) to attract customers.

Mass Digitization — when an individual or organization scans, photographs, types or otherwise creates a digital copy of analogue, physical items. This can be anything from scanning every page of a book, to creating 3D scans of artifacts, to taking extremely high resolution photographs of paintings that enable the viewer to zoom in and examine individual brush strokes. For the purposes of this paper, *mass digitization* refers mostly to the scanning of books and the conversion of them into *searchable text* for the purposes of creating searchable databases.

Transmediation — means to translate a work into a different medium or various media.

All the components of a *transmediation* or *transmediated* works are interlinked with each other to form the whole network or complete work. For instance the online presence of the Harry Potter universe, Pottermore, expands the world of the books and adds an added layer to J.K Rowling's storytelling. In a transmediated *ebook*, the Harry Potter *ebook* could contain moving photos (a key detail of the books) as well as links out to the various elements of Pottermore, or even video clips from the movies. Thus creating a multimedia, interactive book.

Ecurator (variation: e-curator, eCurator) — A curator is a person who cares for the collection of a museum or other types of archive, and typically selects which parts of that collection are to be displayed for the public and in what way. In this way, they are not only a steward, but frequently an interpreter of the organization's collection. To be an *ecurator* is simply to shift these responsibilities from the physical world to the digital world. *Ecurators* deal in *mass digitization* and the *transmediation* of an organization's collection.

.epub — standard document type for an *ebook*. In Adobe Indesign, a designer can export a book layout as a variety of document types, such as .pdf, .jpg, *.epub*, etc.

Searchable Text — digital text that isn't "flattened" or what amounts to a photograph that cannot be edited or changed without photo editing software. To have a body of text

be “searchable” means that search engine software is able to scan through the text to match up keywords from a user’s inquiry. Not all *ebooks* are searchable. If a book has been scanned on a flatbed scanner, and not run through conversion software to make it “readable” then it is not *searchable text*. Adobe Acrobat has the ability to convert a scanned image of text into readable text with remarkable accuracy as long as the scan is “clean” or without a lot of distortion. *Mass digitization* projects rely on conversion software to create “*searchable text*” *ebooks* from scans.

Ebook (variation: e-book, eBook, ebook) — an electronic version of a book (usually of a print book, but some books exist solely in digital form) that can be read on a computer or handheld device. *Ebooks* can be pieces of a transmedia project, or can stand alone as a digital work. *Ebooks* can contain *searchable text*, or be simple PDF scans of a physical book.

Flow Text (variation: reflowable) — a type of electronic document that can adapt its presentation to the device the user is reading it on. This means that an *ebook* with *flow text* will appear differently on a smartphone than it does on a tablet or a computer. This is the most common format for text heavy *ebooks*, because it is more important that the user be able to access the text in a readable way. What this generally means is text that will appear as one page on a computer may take two pages on a tablet or eight pages on a smartphone, but the content remains the same. *Flow text* also enables the user to

change the typeface being displayed, and the font size. This is an important accessibility feature for people with sight limitations.

Fixed Text — a type of electronic document that cannot adapt its presentation to the device the user is reading it on. This means no matter what device you are using the images and text remain exactly where the designer put them. The most common version of the *fixed text* document is the PDF. This form of *ebook* is used almost exclusively for image heavy books with very little text. Photo books and art books rely on this form. It should only be used for books where the layout and design are integral to the story telling, as *fixed text* tends to limit what devices the user can read the ebook on.

Extended Collective Licensing (ECL) — is a type of collective rights management that allows for freely negotiated copyright licensing of exclusive rights designed specifically for mass use, where negotiating directly with individual copyright holders is not possible because of their sheer volume. Under ECL laws, collecting societies (such as the Authors Guild or the American Society of Composers, Authors and Publishers) negotiate ECL agreements on behalf of their members, as well as non-members because ECL laws allow collecting societies to enter into ECL agreements on behalf of all rights owners of the category of rightsholder the collecting society represents. Once the collecting society and the user, such as a TV broadcaster or someone engaging in

mass digitization, have negotiated an ECL agreement, it comes into force and covers only the types of copyrighted works for uses specified in the ECL license.

Fair Use — According to the United States Copyright Office Website, *fair use* is a legal doctrine that promotes freedom of expression by permitting the unlicensed use of copyright-protected works in certain circumstances. “Section 107 of the Copyright Act provides the statutory framework for determining whether something is a *fair use* and identifies certain types of uses. . . criticism, comment, news reporting, teaching, scholarship, and research are examples of activities that may qualify as *fair use*” (United States Copyright Office Fair Use Index. More Information Page, 2016).

Transformative — Is a factor used in determining if the use of a copyright work falls under *fair use*, and therefore outside of the copyright violations. “*Transformative* uses are those that add something new, with a further purpose or different character, and do not substitute for the original use of the work” (United States Copyright Office Fair Use Index. More Information Page, 2016).

The Constitutional Provision Respecting Copyright — Section 107 calls for consideration of the following four factors in evaluating a question of *fair use* according to the United States Copyright Office Website:

1. Purpose and character of the use, including whether the use is of a commercial nature or is for nonprofit educational purposes

2. Nature of the copyrighted work
3. Amount and substantiality of the portion used in relation to the copyrighted work as a whole (is the process *transformative*)
4. Effect of the use upon the potential market for or value of the copyrighted work

1.5 Methodology

This capstone is the culmination of research conducted through three University of Oregon graduate courses, “Copyright Law” and “History and Theory of Communications” through the School of Journalism and Communications and Comparative Technologies for Arts Administration through the Arts Administration program. These courses were selected to deepen the author’s understanding of the role of copyright law in creating *ebooks* from existing works, and the resources needed to complete such projects, while contextualizing these topics in terms of *mass digitization* projects. The coursework provided an opportunity to explore the evolution of *ebook* publishing, copyright law, and *mass digitization* projects, while looking at how publishing operates in a *mission driven* environment. The methodological paradigm of this research is of the interpretivist and constructivist perspective based in subjectivist epistemology. The methods used as the foundation of this research were document analysis, website analysis, video review, software analysis, and literature review.

1.6 Limitations and Delimitations

For the purposes of this research I looked only at the Metropolitan Museum of Art's various digitization initiatives and, specifically, the role Google Books plays in those initiatives. I then attempted to extrapolate the methods being used at the Met to provide smaller organizations with recommendations should they chose to begin their own *mass digitization* project. I did not look at the Google Library Project or the Google Partner Program as a whole, nor did I examine other *mass digitization* projects underway at other museums. The Metropolitan case study not only provides a prime example of *mass digitization* enabled by Google, but also *mass digitization* conducted in-house without the aid of Google. In this case, both exist side by side in the same organization, and thus allow for a reasonable comparative analysis. Also, in the time allotted, given the resources available, and a lack of generalizability, a mass case study of the various types of *mass digitization* projects currently underway across the country would be beyond the reasonable scope of a master's capstone study.

SECTION 2

Google Books, Museum Publishing, and Mass Digitization of Libraries:

Understanding the Current Publishing Environment

2.1 The Google Print Program: Libraries and Publishers

In 1996, while working on a research project, Google co-founders Sergey Brin and Larry Page originally imagined the Google Print Program and Library Project. They envisioned a "web crawler" that could index the contents of books in digitized form and note the connections between each source to determine the relevance of any given book to a search query. The brainstorming around this idea eventually inspired Google's core search algorithms. Years later, in 2002, Google officially began the book project (Kubis, 2010), and in 2004, Google announced a new program called Google Print. This program initially consisted of fifteen publishers that agreed to make their products available online via Google's database. Two months later, in December 2004, the Google Library Project's first cohort of five prominent libraries signed on and agreed to make digital scans of their collections available worldwide (Yuan, 2011). Google initiated its Library Project by announcing agreements with five initial partners: Harvard University, the University of Michigan, the New York Public Library, Stanford University, and the University of Oxford (Kubis, 2010). Under the partnership agreements, the libraries allowed Google to digitize their printed collections. In turn, Google indexed the contents of the books, displayed "snippets" of the books in Google search results, and

provided the libraries with digital copies. If Google is successful in achieving Brin and Page's vision, Google Books will function as a one-of-a-kind digital resource. It will become "the world's largest search and browsing tool, library, bookstore, and book service combined" (Kubis, 2010, p. 219).

Google and its partners stated that since they didn't plan to make the full text of these books available online, they were not violating copyrights by making these "snippets" available to online searchers. They argued they were not only displaying small pieces of the work, but were also driving traffic to legitimate sellers of the complete work. Google claimed it was improving the profits of copyright holders, not infringing on their profits by making the contents of print books more accessible to searchers, who could potentially buy or borrow books after seeing excerpts of them among the search results. Google argued that by increasing access in this way, they and their partners were not infringing on the copyright holders' ability to profit from their works (which is key in copyright cases). Google also argued they intended to sell advertising "keyed" to results lists incorporating the digitized books that were for sale elsewhere (Manuel, 2009) and are therefore further assisting the copyright holders in profit making. Not all copyright holders agreed with Google's assessment of the situation. As a result, the Library Project led to two lawsuits in 2005 alleging copyright infringement against Google. One from the Authors Guild (which is discussed later in this paper) and the second from five high-profile publishers (Yuan, 2011). However, before we unpack the most recent court decision, there are a few more things to understand about Google's book controversy. While the Google Library Project is

currently the most notorious part of the Google Print Program, but it is not the only aspect of it.

Google's Library Project is part of a larger initiative, initially known as Google Print later renamed Google Book Search, and finally called Google Print. Along with the Google Partner Program the Google Library Project currently makes up Google Print. The Partner Program is far less controversial. It allows authors and publishers to submit copies of their books for indexing in Google's search engine and digital publishing. In the Partner Program, participants chose to have their copyrighted books digitized or indexed, and therefore it hasn't been subject to allegations of copyright infringement like those made against the Library Project (Manuel, 2009). "In the case of both the Book Partner Program content owners actively submit their material for inclusion in the searchable database. Only the Library Project does not seek authorization to copy from content owners. Hence, it is the only program that has been challenged in court" (Jewler, 2005, p. 2).

2.2 Authors Guild v. Google Inc: The 2015 2nd U.S. Circuit Court of Appeals Decision (Google's Legal Trouble: a handy timeline)

On September 20, 2005, several plaintiffs; the Authors Guild, Betty Miles, Jim Bouton, and Joseph Goulden individually and on behalf of all others, brought a putative class action on behalf of rights-owning authors and publishers who are facing the same copyright issues. After several years of negotiation, the parties reached a proposed settlement that would have resolved the claims on a class-wide basis in 2008. On

March 22, 2011 the district court rejected the proposed settlement as unfair to the class members who relied on the named plaintiffs to represent their interests (*Authors Guild v. Google Inc.*, 2015).

On October 14, 2011 the plaintiffs filed a fourth amended class action complaint which the district court certified (*Authors Guild v. Google Inc.*, 2015). On May 31, 2012 Google appealed the certification, and moved in the district court for summary judgment on its *fair use* defense. The plaintiffs cross-moved in the district court for summary judgment (*Authors Guild v. Google Inc.*, 2015).

On November 14, 2013, the district court granted Google's motion for summary judgment, concluding that:

The uses made by Google of copyrighted books were *fair uses*, protected by § 107. *Authors Guild*, 954 F. Supp. 2d at 284. Upon consideration of the four statutory factors of § 107, the district court found that Google's uses were *transformative*, that its display of copyrighted material was properly limited, and that the Google Books program did not impermissibly serve as a market substitute for the original works (*Authors Guild v. Google Inc.*, 2015).

On November 27, 2013 the court entered initial judgment. On December 10, 2013, the court entered amended judgment, dismissing the plaintiffs' claims with prejudice.

Plaintiffs filed timely notice of appeal. On December 3, 2014 *Authors Guild v. Google Inc.* was argued before the United States Court of Appeals for the 2nd Circuit (*Authors Guild v. Google Inc.*, 2015).

On October 16, 2015 the decision was rendered in favor of Google. Upholding the earlier judgement that Google's uses were *transformative*, meaning it altered the

original work enough to be considered something new and original (another key term in copyright law cases), that its display of copyrighted material was properly limited, and that the Google Books program did not impermissibly serve as a market substitute for the original works (Authors Guild v. Google Inc., 2015).

2.3 Questions that came out of the lawsuits

One of Google's most compelling arguments against acquiring copyright permission prior to scanning was the sheer number of books being digitized, but also the complexities of tracking down not only so many copyright holders, but absentee copyright holders. Some copyright holders are impossible to find, because they are dead, out of business, or unaware of their standing as holders (Burk, 2008). These copyrighted works without owners are called *orphans*, and the U.S. Copyright Office has said they are "a widespread and significant problem . . . that the uncertainty and the gridlock they produce do not serve the goals of the copyright system" (U.S. Copyright Office, 2015, pp. 2-3).

Google's approach for tackling the overwhelming project of finding copyright status for more than 20 million (plus) books (Kubis, 2010) is to establish a registry where scanned works are listed for potential holders to review. If a copyright holder finds their work listed on the registry they can contact Google. Then Google will either pay the appropriate copyright fees or remove the work. This "opt out" policy is the exact opposite way copyright has been traditionally handled, and is changing how the courts are thinking about how copyright laws should be applied (Kubis, 2010). It could enable

groups engaging in *mass digitization* to skip the time consuming task of tracking down a gaining individual permissions, or, in the case of *orphan works*, meticulously documenting their numerous failed attempts at finding the holder. The U.S. Copyright Office allows for limited remediation for copyright holders if a diligent search on the part of the violator can be proven “and mandates that the office keep current and easily available standards on what ‘diligent’ search means” (Yuan, 2011, p. 237). However, Google is extending the “opt out” to known copyright holders by shifting the burden of rights assertion to the holder, rather than asking for permission in the first place. The U.S Copyright Office is taking notice of how Google is dealing with the orphan book issues, and agrees that in the case of *mass digitization*, the issue is not so much a lack of information as it is a lack of efficiency in the licensing marketplace.

For a digitization project involving hundreds, thousands, or millions of copyrighted works, the costs of securing ex ante permissions from every rightsholder individually often will exceed the value of the use to the user. This would be true even if every relevant copyright owner could be identified and located. Thus, even where a library or other repository agrees that a use requires permission and would be willing to pay for a license . . . the burdens of rights clearance may effectively prevent it from doing so (U.S. Copyright Office, 2015, p. 4).

To encourage further dialogue among stakeholders, and to assist Congress, the Copyright Office has proposed a statutory framework known as *extended collective licensing*, which can be used to authorize projects on terms set forth by the parties under government supervision. *Extended collective licensing*, is a type of collective rights management that allows for freely negotiated copyright licensing of exclusive

rights designed specifically for mass use, where negotiating directly with individual copyright holders is not possible because of their sheer volume. This is the model used by ASCAP and BMI to license the copyrighted works of music to performance halls, clubs and radio stations. All rightsholders would have the right to opt out, and the framework would seek to eliminate the difficulties faced by organizations like Google that are attempting to digitize huge collections. The idea is to create a centralized, market-based mechanism for the clearance of rights and the compensation of copyright owners (U.S. Copyright Office, 2015). This sort of official government opt out program is good news not only for Google, but for non-profit museums (the Smithsonian) and research libraries (the Met's Watson Library) engaging in their on digitization initiatives independent of Google.

Authors Guild v. Google Inc has undoubtedly changed the face of online research. The fallout from it will not be apparent for years, but there are many aspects to the decision that will not only start to redefine *fair use*, but will also begin to redefine the rights of copyright holders in this country. "The copyright system that developed in a world of hard copy print is challenged both by the technology of digitization and by the construction of metadata indexes for digitized works. Rules of ownership developed to control access to atoms apply only uncertainly when used to control access to bits" (Burk, 2008, p. 712).

SECTION 3

The Digitization Initiatives at The Metropolitan Museum of Art

3.1 The Role of Nonprofit Publishing House vs. a For-profit Publishing House

Before discussing *ebooks* and Google Books at the Metropolitan, it is important to start with a basic understanding of the publishing house, run by Publisher and Editor-in-Chief Mark Polizzotti, that is operating as part of a much larger 501(c)3 nonprofit. The Metropolitan Museum of Art is a *mission driven* organization with a nearly \$535,000,000 annual budget; eighty-two percent of which is typically devoted to programming (GuideStar, 2015). It is focused, among other things, on “advancing the general knowledge of kindred subjects, and, to that end, of furnishing popular instruction” (The Met, “Terms and Conditions/Terms of Use,” 2016). The Metropolitan Museum of Art (The Met) was founded on April 13, 1870, in New York City, but its origins are even earlier. In 1866 in Paris, France, a group of Americans decided to create a "national institution and gallery of art" to bring art and art education to the American people. Since then it has amassed one of the largest collections of art and art history objects in the world. It is a preeminent research institution, with twelve research libraries, including their flagship research facility, the Thomas J. Watson Library. The Met also run seven conservation and scientific facilities, and seventeen curatorial facilities. All of these programs and facilities are devoted to scholarly endeavors first,

and are heavily supported by philanthropic activities and grants (The Met, “About, Mission Statement, History,” 2016).

The museum’s mission, more than profit margins, affects the project decisions made by Polizzotti and his staff. Serving a mission allows them more leeway in the areas of profit margins and *loss leaders* that are not afforded to for profit publishers (Polizzotti, Art Books and eBooks 4). The Metropolitan’s publishing house, under Polizzotti’s leadership, publishes 25 to 30 new scholarly books and exhibit catalogs each year, as well as any other printed publicity or information materials needed by the various departments within the museum. This publishing house also manages and preserves a massive back catalog of more than 5,000 titles (Polizzotti, 2012). In the commercial world, lending libraries and publishing houses hold very different priorities “while both need to focus on a core mission, the nonprofit organization generally has a broader time horizon. For profit companies must keep quarterly earnings and revenue projections in mind” (CLIR, 2001, pp. 13-14). Because the Metropolitan’s publishing house is part of the larger nonprofit, and is therefore beholden to the organization’s mission and the public good over making a profit, Polizzotti and his staff are less concerned with the financial implications of *ebooks* than commercial publishers. While Random House struggles to figure out how to make the brave new digital world profitable, and concerns itself with *loss leaders* and battling Amazon’s potential for price fixing (Emerald Publishing Group, 2012), Polizzotti is free to explore this world with free downloads and print on demand through Google Books.

MetPublishing (the digital arm of the Met's publishing house) launched in 2014, and within its initial content were 409 free, searchable PDF downloads from the Metropolitan's back catalog that ranged from 1964 to the present. Polizzotti and his staff continue to add to this free bookstore as they progress through the rest of the out of print books ranging from the museum's founding in 1870 up to 1964. MetPublishing is able to make all these books available for free, with a few exceptions due to copyright issues, because the books have lived their natural lifespan (Polizzotti, 2012), and no longer have a large enough demand to warrant making print runs. But Polizzotti understands that these books may still have scholarly value to someone, and it is within his mission to make information available to anyone who needs it. The fact that Polizzotti first concentrated on digitizing and making available the museum's back catalog (a huge cost in time and labor) with very little, if any, immediate monetary payoff demonstrates a clear difference in priorities from a commercial operation. Polizzotti's first priority was providing the public with information, and turning a profit second. "I don't have as rigid constraints let's say, as I might, if I were an editor at Random House . . . I would have to worry about making every penny that I could possibly make on every book. You know, if we're losing a little bit of money for this, but we are actually opening the book up to a wider audience, in that sense our job is done" (Polizzotti 2012, Art Books and eBooks 4, 16:36).

Here is where things start to intersect with Google and its copyright struggles. While Polizzotti may be less concerned with profit, he is still beholden to rule of law, and copyright laws. He isn't allowed to simply give away books that are still in print,

belonging to other copyright holders, and for sale elsewhere. Fortunately, for MetPublishing, they own most of the copyrights in their collections. The back catalog they are digitizing belongs to them, as do all the images (with a few exceptions that can be redacted before publishing). However, his counterparts at the Watson are not as free to play. The Watson Library, under the eye of Librarian and Collection Development Specialist, Erika Hauser, contains approximately 900,000 volumes, including monographs and exhibition catalogs; over 11,000 periodical titles; and more than 125,000 auction and sale catalogs. The Library includes a reference collection, auction and sale catalogs, a rare book collection, manuscript items, and vertical file collections. The Library is accessible to anyone 18 years of age or older simply by registering online and providing a valid photo ID (The Met, “About, Mission Statement, History,” 2016). Most of the works under the management of Hauser and her staff are either *orphan works*, in the public domain, or owned by other copyright holders not affiliated with the museum.

3.2 The Watson

The mission of the Digitization Initiative at The Metropolitan Museum of Art is “to expand access to the Libraries’ rare and unique materials by developing, supporting, and promoting a distinctive digital collection of these items” (The Watson, “About,” 2016). The initiative operates in tandem with the Watson Library and MetPublications. The Watson’s mission is to “to support the research activities of the Museum staff; in addition, it serves an international community of researchers. The Watson Library’s

collection of books and periodicals relating to the history of art is one of the most comprehensive in the world (The Watson, "About," 2016). MetPublishing (the digital publishing branch at the Metropolitan Publishing Department) is a collaboration between the digital content department and the publishing department at the Metropolitan; between those who are used to working in publishing and print, in a slightly more traditional sense, and those who are used to working in more cutting edge, and emerging technologies (Polizzotti, 2012). "These two arms of the Metropolitan Museum of Art [the Watson and MetPublishing] have the common goal of creating a comprehensive digital collection to support the scholarly endeavors of the Museum staff, and will be accessible to an international community of researchers" (The Watson, "About the Digitization Initiative," 2016).

The primary mission of the Watson Library Digitization Initiative is "to make the museum's rare research materials available to as many people as possible by developing, supporting, and promoting the digitization of texts and artifacts" (The Watson, "About the Digitization Initiative," 2016). The Watson's goal is to make their materials accessible to support the scholarly endeavors of The Metropolitan Museum of Art's staff as well as an international community of researchers. MetPublishing's goal is to expand the scope of the Met's reach and communicate with the broadest possible audience, thus fulfilling its mission to "increase public awareness of and appreciation for art, presenting insightful scholarly discussions and diverse museum voices on works of art, art history, and especially the Museum's collections and exhibitions" (MetPublications, "About," 2016, April). This subtle difference in focus (researchers

versus broad-based public outreach) has affected the distribution and platform decisions of both branches of the organization. The Watson's program specifically targets materials outside the parameters of other major digitization efforts, such as Google Books or the Internet Archive (The Watson, "About the Digitization Initiative," 2016).

3.3 The Watson and Google

While much of the Watson's materials may fall outside the parameters of Google Books, and therefore under the digitizing responsibilities of the Digitization Initiative staff, the Met is already benefiting from the legal outcomes of Google's various legal battles over *fair use* and copyright infringement. The *Authors Guild v. Google Inc* case forced the courts to clarify the issues surrounding *mass digitization* projects. For many years it was the goal and desire of large museums to digitize their collections and archives (Clough, 2013). However, one of the main stumbling blocks, besides manpower and money, has been the issue of *orphan works*, and the daunting task of tracking down individual copyright holders. The act of digitizing in and of itself is an enormous undertaking, "putting research collections online demands new skills: selection, interpretation (such as the creation of descriptive metadata that can be indistinguishable from exhibition captions), and presentation" (CLIR, 2001, p. 18). Add to that the complexity of determining a book's copyright status and you have an insurmountable obstacle for most organizations. Copyright may have been transferred to an entity other than the author or publisher; publishers may have gone out of

business; authors may be deceased; the heirs of authors, who may or may not have inherited the copyright to the work, may be difficult or impossible to locate.

Consequently, the cost of simply locating the copyright holders of many books, in order to obtain permission for their works to be scanned into the database, is potentially enormous (Burk, 2008).

However, the decision to uphold Google's claims to *fair use* and its settlement agreement to create a Book Rights Registry that will collect payments, and hold them for potential copyright owners to claim (Kubis, 2010) has set a precedent for other digitization projects, and is already being reflected in the policies and procedures of the Met's digitization project:

These digitized collections are accessible for research purposes. We've indicated what we know about copyright and rights of privacy, publicity, or trademark. Due to the nature of archival collections, we are not always able to identify this information. We are eager to hear from any rights owners, so that we may provide accurate information. When a rights issue needs to be addressed, upon request we'll remove the material from public view, while we look into it (The Watson, "About the Digitization Initiative", 2016).

In essence, the Digitization Initiative is taking the same "opt out" stance that Google has taken and that the U.S. Copyright Office has all but endorsed in writing in its 2015 recommendation to Congress on how to handle *orphan works* (U.S. Copyright Office, 2015). *Authors Guild v. Google Inc* has changed the playing field for digital libraries and museums, and while much of what the Watson and other museum research libraries own may fall outside the scope of Google Books, "one of the great

opportunities for collaboration in the digital realm is between libraries and museums in the emerging new paradigm of ‘*ecuratorship*,’ [the selecting and displaying of objects and writings in a digital format, typically online], and such cooperation would be an ideal way to provide cross-fertilization between subject specialists” (CLIR, 2001, p. 18).

Meaning, even if Google isn’t interested in adding materials to its database, there is a push from the world of research and public education to digitize. The growing expectation from the public is ease of access, and digital versions of works and artifacts at their fingertips, without having to arrange inter library or museum loans. Digital natives, or those born after the advent of the internet, simply expect to be able to find what they are looking for online, and even those born before the internet, have come to rely on on it as a “first line” for research. As a result, organizations are fielding more and more requests for robust websites that serve as an education tool as well as a marketing tool. For smaller organizations like local historical preservation groups and museums with large back catalog of published works they own the copyrights to, the Google Partner program could be a goldmine for marketing and distributing their materials. Google has proven eager to expand its database, and covers most costs for organizations with compelling bodies of work to digitize, as is the case with MetPublishing.

3.4 MetPublishing and Google

Unlike the Watson, which is mostly digitizing letters, photos, and *orphan works* belonging to other copyright holders or public domain works, MetPublishing is the

digital manifestation of the Met's inhouse publishing company. As such, the works they are making part of Google's database are books that have been wasting away on the back catalog list. For MetPublishing and Google, the collaboration is mutually beneficial. The Met's back catalog is enormous, and adds to the depth of information available in Google searches; while Google brings thousands and thousands of people to the Met's digital doorstep. Online collections and catalogs also raise the organization's profile and can be a useful public relations tool. Digital collections can be used as leverage with benefactors and funders by demonstrating an institutional commitment to education, access, and scholarship (Hughes, 2004). The hope at the Metropolitan is that if readers can gain easy access a book, that access will create ongoing interest in other areas of the collection. "We're not doing it as a marketing tool per se, but in terms of publishing economics...I don't think it's going to have a serious negative impact. In fact, in anything, it may end up having a positive one" (Polizzotti, *Art Books and eBooks* 4, 16:25).

By putting collections online, free of charge, for anyone to borrow or download, museums are creating an online repository that anyone with a curious mind and internet connection can access. This was unthinkable even 15 years ago (Clough, 2015). As the world slowly levels the technology gap, the hope is that online catalogs will also narrow the education gap. "Everyday, people should be able to access its treasures through a wider range of electronic and mobile devices" (Clough, 2015, p. 5). If nothing else, it takes the Metropolitan one step closer to fulfilling its mission to "advance the general knowledge of kindred subjects, and, to that end, of furnishing popular

instruction” (The Met, “About, Mission Statement, History,” 2016). One of the greatest benefits of digital technology is the opportunity to support K–12 education. “Museums in particular can bridge the gap between formal and informal learning with lesson plans, online summits, real-time connections to experts, and better credentialing of informal digital learning” (Clough, 2013, p. 8)

SECTION 4

Challenges and Benefits of the ebook

4.1 What About Big Brother: Privacy Issues and Google Books

For all of the excitement around a global library to rival Alexandria (Marina, 2012) there are concerns. Google, for all its talk of “not being evil,” is a for profit organization with shareholders. While the current leadership has proven to be willing to carefully navigate the thorny issues of privacy and come out fairly consistently on the side of individuals, that may not always be the case. Unlike a *mission driven* organization that exists for further the mission and vision of its board of directors, a profit driven organization is ultimately accountable only to its bottom line and shareholders (Kubris, 2010). “Librarians consider themselves ‘trusted guardians of patron privacy,’ and protecting patron privacy and confidentiality has long been a core mission. Indeed, the American Library Association (ALA) has recognized a right to privacy since 1939” (Kubis, 2010, pp. 245-246). According to a survey, 89% of respondents believe that their Internet searches are kept private, and 77% believe that Google searches do not reveal their personal identities (Kubis, 2010).

The reality is, what books you read, the things you research, and the things you download tell a great deal about who you are. These tidbits of information, collected over time, can reveal your deepest secrets, and Google is creating a record of every last click. This is the fiscal trade off. Google isn't providing scanning and distribution services to organizations like MetPublishing for free. Rather, MetPublishing is paying for those services with the personal information of its patrons. The individual is by far and away Google's biggest commodity.

Google Books is sidestepping privacy safeguards long recognized by traditional libraries despite the many commonalities between them. Indeed, the only real differences are that Google has moved the browsing from a building to the Internet, and that the Google service is private, rather than public. While public libraries prioritize patrons' privacy, Google Books, which began with the digitization of library books, does not adhere to similarly strict standards. This technological change-without corresponding protection-is cause for concern (Kubis, 2010, p. 245).

One can only speculate that part of the reason the Watson Library is not using the Google Library Program has more to do with privacy and commodifying concerns than a lack of interest on Google's part to digitize some of the Watson's more obscure materials. The truth is, while the Watson and MetPublishing are arms of one organization they are run by two different groups, with two different sets of professional ethical standards. The Watson is run by librarians, MetPublishing is run by marketers and publishers. One group is focused on the distribution of unfiltered knowledge, and the other is focused on collecting information on the customer to better serve his or her needs, and to better understand when and how to present new products. While

Polizzotti is not under the same economic pressures as the head publisher at Random House, he is tasked with marketing the museum and distributing and selling publications. The more he knows about the people downloading books from his website, the better he can do his job. Of the two, Polizzotti's position is more nuanced and complex. He is operating as a capitalist in a nonprofit world. This diametric pull makes for fascinating further study.

4.2 Role of the print book in a digital world

The Metropolitan is an interesting microcosm of the publishing world. The benefits and challenges presented by the digital evolution has, in many ways, turned on its head an industry that hasn't seen major change in 200 years (Øiestad & Bugge, 2013). But books as objects are not dead. Their uses in the physical world are still varied and currently unparalleled by binary code (Housewright & Polizzotti, *Art Books and eBooks 4*, 2012). The *ebook* may never replace the print book, but will rather evolve into something completely new. While Amazon currently sells more books in *ebook* format than paper (Øiestad & Bugge, 2013) the printed page still has a place in the world of scholarship, at least.

What we are looking at, and what I think makes it interesting, is that the technology has arrived at a level of maturity that we can start envisioning other ways of publishing certain kinds of scholarship that actually make more sense. So, in other words, if a book has a little bit more of a narrative arch that you want to sort of sit with and ruminate on, that is something that I think is still best served by a traditional medium, but if it is something that is more research driven that you want to be able to

access and are trying to gain information with it that's a case where electronic or digital publishing can actually be a little more advantageous, because it allows you to zero in very quickly on the piece of information that you want and get rid of the rest (Polizzotti, *Art Books and eBooks 4*, 18:14).

There is also a concern in some museum and library circles that a surrogate digital copy of the original can never be a satisfactory substitute for the artifact itself (Hughes, 2004), and that the tactile experience of holding a printed book and smelling the ink is a notably different experience than viewing the same book, especially a book with fine art prints, on a iPad. The two mediums instead, and increasingly, are serving vastly different roles in the publishing world (Polizzotti, *Art Books and eBooks 1*). “The lesson to be learned is that there is a place for both the physical and the digital, with one complementing and leveraging the other” (Clough, 2013, p. 3).

4.3 Benefits of the ebook

“Documents once hidden in stacks or in storage facilities are suddenly becoming accessible on demand” (Clough, 2013, p. 3). The two benefits that *ebooks* offer the Metropolitan are: ease in searchability and cross-referencing, and the ease in storage and retrieval. Not only can researchers do keyword searches across catalogs in seconds, pulling up dozens of resources that would have at one time taken hours to locate, those resources are now immediately available. There is no longer a need to wait three days for a book to find its way from the overflow storage house (Hauser, *Art Books and eBooks 4*). Libraries are also able to digitize rarely used titles and in some

cases dispose of the print version all together. This is a huge benefit for large institutions that frequently pay thousands and thousands of dollars per year for offsite storage of books that may be requested once every ten years (Carroll, *Art Books and eBooks 4*). While preservation is a concern in the world of *ebooks*, it is also a benefit. By developing a digital surrogate of a rare or fragile original, objects can provide access to users while preventing the original from damage by handling or display (Hughes, 2004).

4.4 The future of the ebook in Museum Publishing

In the same way web 1.0 was text dumped into cyberspace that didn't stray far from established models, but eventually evolved into something infinitely more complex, non-linear, and multifaceted with the advent of web 2.0 and *transmediation* (the telling of a story over multiple different platforms and media), so too are *ebooks* already evolving into something different from their print predecessors (Carroll, *Art Books and eBooks 4*, 2012). "People are slowly moving away from the model of what reading means, which has been predicated upon the structure of a book. Little by little we are beginning to unlearn that, but having completely unlearned that and moved on to something else, is still something we have yet to see" (Polizzotti, *Art Books and eBooks 4*, 30:40). Ross Housewright, senior analyst with Ithaka S+R, says there are some things that the print book is really designed to do, that it's very good at. Ithaka S+R is a research and consulting service that helps academic, cultural, and publishing communities in making the transition to the digital environment. However, Housewright

believes there are other things publishers use print books for that print books are no longer the ideal medium for. Polizzotti, believes that the future of the *ebook* is a more curated experience that mimics walking through a gallery. However, he says, no one has yet come up with an elegant model for this types of interactive, transmedia experience. That said, prototypes presented by researchers at MIT that enable a more multimedia, web-like environment show a future where *ebooks* are no longer attempts at print replication, but their own medium with content specifically designed to showcase their strengths (Carroll, *Art Books and eBooks* 4). When this happens, Polizzotti says he and his staff look forward to creating content that is unique to the *ebook*, as a separate, but equal partner to the printed page.

SECTION 5

Resources for ebook creation:

What can smaller organizations glean from what the Met is doing

The Metropolitan Museum of Art is a world class, well funded institution with some of the best trained and smartest staff members in the field. They are breaking new ground for other organizations that possess a wealth of printed knowledge looking for its audience. With the maturing of desktop publishing software, and the proliferation of self publishing houses, the idea that an organization needs a huge staff or professional publishers and designers is obsolete. There is no reason why an organization with a good product and a little bravery can't jump into the digital world and swim.

5.1 Design Software

You get what you pay for with design software. There are open source (read free) versions of Adobe's Creative Suite out there, but they lack the level of support, training, and standards that Adobe has cultivated in its almost four decades of design software development. In the days before "the cloud", third party brokers like TechSoup would offer nonprofits at cost Adobe Software (they still do for Quickbooks, Microsoft Office, etc.). However, two years ago Adobe moved its software into the cloud. This means the software updates and source files exist on the Adobe server, and people or businesses login with their user ID to access the latest Adobe updates and the software applications. This user ID is based on a subscription that is paid monthly, and is based on whether the user is accessing one, two, or multiple programs, and how many people have rights on the account. There are single user plans, or multiple user plans, and each plan can be tailored to the account holder's needs.

It is advisable for any organization that does any amount of print and digital design to invest in Adobe software. In the long run, the cloud will probably prove to be better for users, because it offers automatic software upgrades, and Adobe is notorious for its constant iterative software changes that quickly make earlier software generations obsolete. Adobe does offer educator and student prices to nonprofits that can demonstrate their 501c3 status. Meaning, single user plan with all the suite applications can be purchased for as little as \$34 a month. It's easy enough to designate one person as the staff designer, or even hire a communications specialist with design skills if the organization is at that point. This person, and the Adobe Cloud

Suite, can save the organization a great deal in time and money working with outside contractors. Further, the Adobe Creative Suite is the basis for just about any kind of print or digital design an organization will need.

Once the software is sorted out, Adobe, YouTube, and other online video hosts have a vast wealth of free video tutorials. Just about anything an organization might want to learn about Adobe's software can be learned online for free. Design sensibility is another story, so it's important that the designated staff designer have at least a working knowledge of good design theory and visual communication.

That said, the reality is many organizations find it far more cost effective to budget for a contracted graphic designer. For a variety of reasons, an organization may not have the staff resources, or even have paid staff who are adept enough at computer-based design, to become proficient in-house designers. In these cases, outsourcing the design, and managing of Adobe software, and publishing platforms, makes the most sense.

Design contracting, much like other professional services, comes in two varieties: one time jobs, and retainer contracts. This means that an organization can hire for a specific project (laying out a book for print and ebook), and agree on the terms of the contract, and the professional relationship ends with the project. Or, an organization can pay a monthly fee to retain the services of a contracted designer on a retainer contract for a one or two year cycle. This keeps the organization from being permanently tethered to one person, which is part of the draw of having a contract worker, and also keeps the organization from having to pay the salary and benefits of a full-time

employee. If the organization becomes dissatisfied with the contractor's work or decides to go in a different direction with their visual communications, the option exists to not renew the contract and hire a new designer.

The big message in all of this is: don't rely on volunteers, even talented ones. They can and will walk away at any time, and are under no obligation to meet an organization's needs in a timely fashion. There are great volunteers out there, but not only do organizations disrespect the profession by expecting designers to work for free, they are placing themselves in a position to get subpar work by an amateur, or someone who simply isn't motivated to give one hundred percent. It is also unnecessary. There are dozens of qualified, affordable freelancers who can meet all of an organization's graphic design, and book design needs.

5.2 What staff resources are needed

As was covered in section 5.1, there are a few options depending on the size of the organization in question, and how much graphic material the organization produces. If the organization already has a staff member charged with designing fliers and booklets, it isn't huge stretch to have them design a book, or scan an existing printed book and convert it to a PDF or .epub. There are countless quick, free online tutorials that can get an already knowledgeable staff member up to speed on how to create and market epublications. However, if the organization is small with no in-house designer the best option is to contract the work out.

If contracting is an organization's preference, the first question to be answered is what kind of book is needed? Is this an image heavy book with very little text (like many art museum catalogs), or is this a text heavy book with only a few images (if any) throughout? The answer to this very basic question will determine the organization's next step.

5.3 Reflowable vs. Fixed Text: when to use them

There are two types of *ebooks*: *reflowable text* books, and *fixed text* books. *Reflowable text* is exactly what it sounds like, the text adjusts to fill whatever device screen it is being viewed on; it's rather like water that way. *Fixed text* is fixed in place; rather like a mountain. Books that are text heavy with only a few images and captions exist best as *reflowable text* and require very little graphic design. In fact, the only thing an organization would need a designer for is the cover design (which is still important in *ebooks*), and photo or graphics editing. A designer is also useful if image captions are needed as part of the images. Designers can also prepare the images for screen viewing so they use as little storage memory as possible. If the book to be created is a text heavy book, then an organization only needs a good editor (either on staff or contracted), and some form of word processing software (like Microsoft Word). The rest can be done by whatever distribution platform the organization chooses to use (see section 5.4).

However, if the book in question is mostly composed of images, and the storytelling is heavily reliant on layout and look, a designer is definitely required. If the

organization is choosing to contract out design, it is worth it to look for a book designer who knows how to create *ebooks* as well as print books. If the organization is planning to have the in-house graphic designer create the *ebook*, that staff member should educate themselves through online tutorials (some are listed in the appendix section) before any design is done (Harmer, 2015). The reality is, fixed design (*fixed text*) *ebooks* have a lot to offer, but also have some major drawbacks.

5.3.1 The good, the bad, and the unintelligible

Reflowable text is the basic standard for *ebook* publication, and depending on how the organization chooses to convert the final product into an .epub document (via in-house design or an online company), it is the easiest type of book to create and distribute. It also fails miserably to harness the full power of digital publication (Harmer, 2015). The good thing about *reflowable text* is it adapts easily to any device. Also, the user can change the font size and style, and it is easily read by text to audio software, both of which are important for ADA accessibility. Also, *reflowable text* uses very little storage, and is equally accessible on the fancy multimedia ereaders like Kindle Fire, and the grey-faced text readers like the original Kindles. The drawback: *reflowable text* moving all over the place means all the images and graphics shift around too, which can make a book that is image dependent completely unintelligible as the *ebook* flows from one device to the next. *Reflowable text* also doesn't allow for the multimedia incorporation of video and animation (White, 2016).

Fixed text, on the other hand, operates like a completely self-contained website. It not only allows image heavy *ebooks* to appear exactly the same regardless of what device they are accessed on. *Fixed text* layouts also allow for embedding video files, audio files, animation, and slide shows. The .epub can become a multimedia work in its own right, one that is readable on multimedia ereaders (Kindle Fire), tablets (iPad), smartphones, or desktop computers. The drawbacks, however, speak to why reflowable, image-light *ebooks* are still the norm. A multimedia *ebook* with a ton of animation, and/or video can eat up storage memory very quickly. This is an issue that hasn't been completely solved yet, however it is one that can be worked around if the final product is something the audience is willing to sacrifice memory to. There is also a limit on what type of device the user can access the *ebook*, and Adobe Indesign requires the designer to choose what kind of device they are designing for from the beginning, ie. iPad, Kindle Fire, iPhone, etc. (Harmer, 2015). However, an *ebook* designed for an iPad, looks just as good on a smartphone, as it does on a desktop. There may be some viewing issues on a Kindle Fire, but that is easily tested within the design software. MetPublishing is just starting to experiment with interactive, *fixed text* epublications for its digital art catalogs. In the case of art publications, the limits of *fixed text* far outweigh the drawbacks. However, if an organization is going to attempt *fixed text* ebooks, they need to invest in good design, or the entire effort will be a waste. Also, *reflowable*, text-heavy *ebooks* allow for hyperlinks within the text, which take the reader out to various websites or videos stored on YouTube or the organization's website. However, relying on hyperlinks in text-heavy books that may be viewed on

older ereaders not designed for multimedia content is a bad idea. The grey-screened ereaders are great for one thing, reading text. Many of them don't have wifi download options, and none of them have any kind of web browser. This is why it so important for an organization to have a clear idea from the beginning of what kind of book it wants, and who its audience is. In the end, it all comes down to market research and meeting your audience members where they are.

5.4 Marketing and Distribution Platforms

One of the biggest benefits of online book distribution, hard-copy or *ebook*, is the increased potential audience. Simply by having an online store via the organization's website increases the number of potential customers exponentially. However, to get the most out of an online presence, it is advisable to piggyback off existing distribution platforms (websites that distribute things) with built in audiences, such as Amazon. While allowing customers to order copies of organizational books, or download *ebooks*, directly from the group's website is a good start, the customer pool is limited to people who are already seeking out the organization, and are possibly predisposed to purchasing, but they are most likely people who already exist in the organization's audience pool. By creating an online store with a company like Amazon, the organization will not only be selling to its existing audience, but will also have access to all of Amazon 300 million potential customers. This not only increases the chance of discovery, but also may serve to drive customers to an organization's website or in person events, if the customer lives near-by. A presence on Amazon may also

improve a website's Google search placement, which is affected in part by the amount of traffic a website receives. Another benefit of Amazon's publishing and store platform is the anti-pirating measures that can be easily put in place with the check of an online box. There is more about that in section 5.5.

Amazon, of course, is not the only method of book selling, and it may not be in the organization's plan to sell its *ebooks* at all. The Metropolitan Museum of Art is providing its old art books and catalogs as free PDF downloads or print on demand (meaning the purchaser pays for printing and shipping of the book) through Google Books. This option not only leverages Google's database, but allows groups to save on upfront printing costs. The idea behind print on demand, is just what it sounds like. The book in question doesn't get printed until someone orders a hard copy, and only then is the customer charged. So, it doesn't cost the Metropolitan anything other than the time and labor of scanning 400 plus books into the Google database. This is something Google is more than happy to help with if it means more eyeballs on their search engine.

Google and Amazon are not the only ways organizations can leverage print on demand or *ebooks*, they are simply the most ubiquitous. Most small, independent publishing companies will provide print on demand and *ebook* production as part of their services. A quick online search will turn up dozens of small companies who specialize in helping authors and organizations self-publish. Some even offer placement on dozens of different *ebook* sales platforms, like Amazon. It's up to the organization how they want to price, distribute, and package their book once it's written and designed.

Some organizations find it's best to offer their older books as free *ebook* downloads, but charge if a customer wants the hard-copy. Some, like the Metropolitan, offer all their *ebooks* free, and only charge the customer the cost of shipping and printing; thereby making no profit, but also not costing the organization much either. Still other organizations only offer hardcopies of their books, but give the customer a free *ebook* download with the purchase. This is essentially selling the *ebook* as well, because the customer is forced to buy a hard-copy in order to get the *ebook*. There is also the option of selling the hard-copy and the *ebook* at the same price, or offering the *ebook* at a reduced price. It really depends on what the organization's goals are; making money? Increasing readership and patronage? Both? *Ebooks* have become the publishing industry's *loss leaders*, and are mostly used to draw people in and leverage their interest to make more sales. This is why today it is common to see *ebooks* offered for free, or at very low prices. There is no reason why a small organization that is already creating books can't use the same tools and ideas to drive more traffic to their websites and increase consumership of their books. If the book has already been designed for print, it's just as easy to export it as an .epub or PDF and offer a digital download to the customer as well.

In marketing, perception is everything. How something looks sends a very clear message to the customer, who is bombarded on a regular basis by well-designed advertising and publicity pieces. The "average number of advertisement and brand exposures per day per person: 5,000+; Average number of "ads only" exposures per day: 362; Average number of "ads only" noted per day: 153; Average number of "ads

only” that we have some awareness of per day: 86; Average number of “ads only” that made an impression (engagement): 12” (Johnson, 2014, sjinsights.net). Good cover design that is pleasurable to look and engage the reader by clearly communicating the book's content in an attractive way can help cut through that clutter. This is why hiring a professional designer for at least the book cover is important. Readers still rely heavily on the images of book covers when they are perusing the thousands of listings offered by online dealers. Not to mention the look and feel of the print book. Even if an organization has skilled editors and writers, it's important to design a cover that looks professional and has impact. Like it or not, people are more likely to trust something that looks professionally designed, and the average American consumer is way more adept at spotting poor design than even they may realize. A lot of design is based in emotion and feeling, as is consumerism. Good design can break through the noise and make the consumer feel instantly safe with an organization's book, which will engage them and make them far more likely to buy.. “Bottom-line, increasing engagement will help advertising break through the clutter of not only the hundreds of ad exposures per day, but of the thousands of ad and brand exposures per day” (Johnson, 2014, sjinsights.net).

5.4.2 Distribution

As was discussed in the previous sub-section, Amazon offers a service to self publishers. There is no upfront cost, Amazon simply takes a small percentage of each sale. Sellers are allowed to set their own prices and decide if they will be selling

ebooks, *ebooks* and hardcopies, or just hardcopies and in what combination. Sellers are also allowed to change their prices at anytime and offer discounts when they choose. The drawback of course, is if you use Amazon, you only use Amazon, and Amazon isn't the only game in town. Unless an organization has small sales goals, or has a staff member devoted to managing the books publicity, and acting as the books agent, it's probably best to work with an organization like IngramSpark.

IngramSpark not only offers organizations a wide variety of print or digital book options, their staff also handle all of the book placements at "Independent bookstores, more than 50 Online stores, Big chains, Little chains, e-book retailers, Local niche retailers, Libraries, Schools, Universities, And just about anyone on the planet who sells (or is even thinking about selling) a book in any format" (IngramSpark, "Distribution" 2016). There is a service fee involved, and tiered pricing depending on what services the organization chooses, but it offloads all of the marketing and retail management onto a contractor, thus freeing staff to focus on other things, like developing the next book.

There are other companies like IngramSpark, but few offer the same outreach to retailers, both on and offline. The bottom line again: you get what you pay for. If an organization simply wants to sell their hard-copy book, and are only interested in giving away a PDF version to the handful of people they think will be interested, then it's as simple as having the book design export the design as a pdf and having the web manager post it online. However, if an organization hopes to reach more readers and a

wider audience, it would be beneficial to at least post the book where the potential audience hangs out online.

There is an additional benefit to working with IngramSpark and Amazon, as both services enable organizations to easily secure their digital books from piracy— at least as much as is possible. This, of course, may be only important for groups who are selling. Organizations giving their books away may have fewer concerns about piracy. However, it's always good to have options, and using a third party distribution and publishing program makes copyright protection easy.

5.5 What are the potential difficulties

If an organization is already writing and editing publications, the hardest part is over. The rest is just a matter of budget and motivation. However, there are some challenges to consider before diving into the world of digital publishing.

5.5.1 Securing content

Ebook encryption: the written word is converted into an algorithm and thus unreadable without the correct software to decrypt it. This in essence is what an ereader does, whether the Kindle app on an iPad or a Barnes and Noble Nook. Ereaders decrypt encrypted *ebooks*. They also read unencrypted *ebooks*, but that is less important to this discussion. This encryption can also prevent people from sharing an *ebook* by sending copies to friends, which is a copyright violation should an author

choose to enforce their rights. When an author uploads their *ebook* to Amazon, there is an option to secure the book from sharing and copying. Some people advocate for this, some people argue against it, stating that by sharing the books, they are reaching more people who may come back to the source author for the latest work that person is producing. It's a personal choice, and one an organization should be clear on from the beginning. Is there goal to spread the information far and wide, or is there goal to make a profit, and then decide how to go about that.

The other issue around copying is plagiarism. The reality is, if someone is going to plagiarize a work, they can just as easily type it up from a hard-copy version as they can cut and paste. However, encryption does make plagiarism more time and labor intensive, and most plagiarizers are inherently lazy and will look elsewhere if there are too many roadblocks to stealing. There are many schools of thought on whether users should be able to share *ebooks* at will; most everyone agrees that plagiarists should be stopped in every way possible. Again, if an organization is concerned with plagiarism or willful sharing, it's best to market *ebooks* through a third party who will make the encryption and security process as easy as possible. However, there are some good self-publishing blogs that explain how to best secure an *ebook*, should an organization choose to go it alone. Some are listed in the resource appendix.

5.6 What are the potential benefits

To reiterate, if an organization is already publishing books, the benefits of taking the next step and producing a digital version alongside the print version is mainly one of

marketing. It will raise an organization's brand and profile, as well as make the hard work and hard-won final product available to a wider number of customers. *Ebooks* are more adaptable for people with sight limitations, and are easier and cheaper to produce and disseminate than they're hard-copy siblings. By taking the write steps, an *ebook* is no easier to pirate or plagiarize than a hard-copy book, and it is the preferred method of reading for some customers. This is not to advocate jettisoning print books, rather it is to suggest providing a richer offering and experience to a wider variety of customers. The point of writing a book is to have people read it; therefore reaching as many audience members as possible can only be a benefit.

SECTION 6 In Conclusion

The digital revolution in publishing can feel daunting and, at times, completely out of control. In many ways the industry, which hasn't had to alter its business model much in nearly 500 years, has failed to adapt to advances in technology and market expectations. After all, 500 years is a lot of history to upend and, unfortunately, copyright laws around the world have done little to keep pace with the day to day issues faced by publishers, writers, libraries, copyright holders, and anyone connected in any way with copyrighted works and publishing. The lack of proactive leadership has led to case by case, piecemeal policy and reactive "band-aid" legislation like the 1996 Digital Millennium Copyright Act, which criminalizes the production and dissemination of

technology, devices, or services intended to circumvent digital rights management (DRM) measures that control access to copyrighted works. It has been largely ineffectual at stopping piracy and plagiarism at the legislative level.

All of that being said, while copyright laws in this country and around the world have a long way to go before they can claim to be abreast of the latest technology, publishers, especially small ones with less to lose if the old ways fall, are finally adapting to shifts in the marketplace, and are using ebooks in much the same way they once used mass market paperbacks. The industry still fears copyright violations, but is beginning to realize that *ebooks* are really no more vulnerable to violations than printed works are, a lesson being driven harshly home by Google's blatant scanning of printed books. So, what does this mean for small organizations? Mainly, fearing the digital world isn't going to make it go away, and digital books, in the end, pose no more risk than printed books, but offer a huge opportunity to reach a larger audience, which in the end is the whole point of writing a book in the first place.

APPENDIX ONE: Ebook Creation Resources

This is by no means a comprehensive list of resources. There are far too many to include, and more are coming into existence every day. These are just a few recommendations of trusted, and established companies that have built good reputations and are reasonably priced.

Self-Publishing and Print Options (all encompassing)

This is just two of the top self-publishers out there. Just Googling “self-publishing” will pull up countless others. The things to look for are: do they offer print on demand, where do they distribute to and to how many places, both online and in real life, do they offer menu pricing, and what services are offered.

IngramSpark

<http://www.ingramspark.com/>

BookBaby

<https://www.bookbaby.com>

SmashWords

<https://www.smashwords.com/>

INscribe Digital

<http://www.inscribedigital.com/>

Draft2Digital

<https://www.draft2digital.com/>

Ebook Only Publishing Platforms

These tend to be less expensive, but more labor intensive, as the author is responsible for maintaining and setting prices, and communicating with customers. Also, an *ebook* published on these platforms only exists on these platforms. This means that, an organization would have to pick one (Kindle being by far the most ubiquitous) or

commit to publishing and maintain accounts on multiple platforms. The ones listed here are only the biggest, and doesn't include the dozens of less well know platforms where potential customers shop. If an organization is interested in saving some money, and just wants to get an *ebook* out there, Kindle is probably the best bet. Most of the customers will be there. Followed by Barnes and Noble's Nook.

Kindle Direct Publishing

https://kdp.amazon.com/signin?ref_=kdp_gp_p_us_psg_bt_ad1

Nook Press

<https://www.nookpress.com/>

“Borrowing” Platforms

This is another way to distribute a book. Scribd calls themselves “Netflix” for books. Basically, customers pay a monthly subscription fee and can consume as many books as they want, or at least as many as are allowed by their subscription level. It is possible to work directly with them to have an independently published *ebook* place in their libraries, however, again, paying a group like IngramSpark or BookBaby, will usually get a book automatically placed with these sorts of companies. Scribd pays fees based of number of reads to the publisher, or self publisher which is disseminated down to the writer. There was another company, Oyster, that offered a similar service, but they went out of business in 2015. Scribd seems to be going strong, however.

Scribd

<https://www.scribd.com/publishers>

How To Design Tutorials

These are just a few videos on how to use InDesign for *eBook* publishing. There are dozens more. Adobe itself offers hundreds of hours of free training.

Adobe Help

<https://helpx.adobe.com/indesign/how-to/ebook-fixed-layout.html>

Lynda.com

Offers a free trial, but likely will cost for the full courses

<https://www.lynda.com/InDesign-tutorials/InDesign-CC-2015-EPUB-Fundamentals/374187-2.html>

Terry White's Tech Blog

<http://terrywhite.com/create-fixed-layout-ebook-adobe-indesign-cc/>

Securing documents

This is only relevant for organizations choosing to make books available from their own websites and want to completely bypass companies like BookBaby or even Kindle Publishing.

Inkwell Editorial

<http://inkwelleditorial.com/prevent-ebook-theft>

REFERENCES

Adobe Creative Cloud Website. Payment Plans Page. (2016, April 26).

Retrieved from https://creative.adobe.com/plans?plan=edu_inst&promoid=KTRQB

Antonelli, P.. (2014). "The Museum of Modern Art embraces all aspects of digital culture and design." *Arts, Culture & Media*, Public Radio International (PRI), USA, December 15. Consulted January 25, 2015. Retrieved from <http://www.pri.org/stories/2014-12-15/museum-modern-art-embraces-all-aspects-digital-culture-and-design>

Brennan, Timothy J. (October, 2009). The Proposed Google Book Settlement: Assessing Exclusionary Effects. *Global Competition Policy: Antitrust Chronicle*, (2), 1-9. Retrieved from <http://ssrn.com/abstract=1503414>

Burk, D. (2007-2008). Mereology of Digital Copyright, The. *Fordham Intellectual Property, Media Entertainment Law Journal* 18(3), 711-740. Retrieved from http://heinonline.org/HOL/Page?handle=hein.journals/frdipm18&div=25&g_sent=1&collection=journals

Carroll, J., Hauser, E., Housewright, R., Polizzotti, M. (2012). "Video (panel discussion): Art Books and eBooks 4" *MetMedia*, Metropolitan Museum of Art Online, USA, September 28, 2012. Retrieved from <http://www.metmuseum.org/metmedia/video/lectures/art-books-and-ebooks-4>

CLIR (2001). *Building and Sustaining Digital Collections: Models for Libraries and Museums*. Washington, D.C.: Council on Library and Information Resources.

Clough, G. Wayne (2013). *Best of both worlds: museums, libraries, archives in a digital age*. Washington, DC: Smithsonian Institute. ISBN: 978-0-9819500-1-3 (epub)

Emerald Publishing Group (2012). An electronic shock: How e-books are transforming the traditional world of publishing. *Strategic Direction*, 28(7), 23-26. DOI 10.1108/02580541211234769. Retrieved October 13, 2014.

GuideStar (2015). Nonprofit Report for the Metropolitan Museum of Art. Retrieved from <http://www.guidestar.org/organizations/13-1624086/metropolitan-museum-art.aspx>

Google Books Website. Perspectives Page. (2016, November 28). Retrieved from <https://www.google.com/googlebooks/perspectives/>

Hauser, E. (2012). "Video: Art Books and eBooks 2" *MetMedia*, Metropolitan Museum of Art Online, USA, September 28, 2012. Retrieved from <http://www.metmuseum.org/metmedia/video/lectures/art-books-and-ebooks-2>

Harmer, Tony (Oct. 13, 2015). EPUB: To Go With The Flow, Or Not?. *Adobe Blogs*. Retrieved from <http://blogs.adobe.com/tonyharmer/2015/10/13/epub-to-go-with-the-flow-or-not/>

Hughes, L. (2004). *Digitizing Collections: Strategic issues for the information manager*. London, England: Facet Publishing.

IngramSpark Website. Blog. (2016, April 26). Retrieved from <http://www.ingramspark.com/blog>

Jewler, Robin (December 28, 2005). The Google Book Search Project: Is Online Indexing a Fair Use Under Copyright Law?. Congressional Research Service, The Library of Congress, Report for Congress. Retrieved from http://www.ipmall.info/hosted_resources/crs/RS22356_051228.pdf

Ji, Y. (2011). Why the Google Book Search Settlement Should Be Approved: Response to Antitrust Concerns and Suggestions for Regulation. *Albany Law Journal of Science Technology* 21(1), 231-278. Retrieved from http://heinonline.org/HOL/Page?handle=hein.journals/albnyst21&div=8&start_page=231&collection=journals&set_as_cursor=0&men_tab=srchresults#

Johnson, Sheree (2014, September 29) New Research Sheds Light on Daily Ad Exposures, SJ Insights. Retrieved from <https://sjinsights.net/2014/09/29/new-research-sheds-light-on-daily-ad-exposures/>

Kubis, K. (2010-2011). Google Books: Page by Page, Click by Click, Users are Reading Away Privacy Rights. *Vanderbilt Journal of Entertainment and Technology Law* 13(1), 217-254. Retrieved from http://heinonline.org/HOL/Page?handle=hein.journals/vanep13&div=9&start_page=217&collection=journals&set_as_cursor=0&men_tab=srchresults#

Lakin, C.S. (July 6, 2015). The Only Self-Publishing Platforms You'll Ever Need. *Live Write Thrive Blog*. Retrieved from <http://www.livewritethrive.com/2015/07/06/the-only-self-publishing-platforms-youll-ever-need/>

Lao, Marina (2012). The Perfect Is The Enemy Of The Good: The Antitrust Objections To The Google Books Settlement. *Antitrust Law Journal*. Retrieved from <http://www.jstor.org/stable/43486985>

Manuel, Kate M. (February 5, 2009). The Google Library Project: Is Digitization for Purposes of Online Indexing Fair Use Under Copyright Law?. Congressional Research Service, The Library of Congress, Report for Congress. Retrieved from <https://www.fas.org/sgp/crs/misc/R40194.pdf>

Metropolitan Museum of Art Website. About, Mission Statement, History. (2016, April). Retrieved from <http://www.metmuseum.org/about-the-met>

Metropolitan Museum of Art Website. Terms and Conditions/Terms of Use (2016, November 28). Retrieved from <http://www.metmuseum.org/information/terms-and-conditions>

MetPublications. About. (2016, April). Retrieved from <http://www.metmuseum.org/art/metpublications/about-metpublications>

Na, Nari (2007). Testing the Boundaries of Copyright Protection: The Google Books Library Project and the Fair Use Doctrine. *Cornell Journal of Law and Public Policy*, 16(2), 417-448. Retrieved from <http://scholarship.law.cornell.edu/cjlp/vol16/iss2/6>

Øiestad, S. & Bugge, M. (2013). Digitisation of publishing: Exploration based on existing business models. *Technological Forecasting & Social Change*. Vol. 80. 2014. pp. 54-65. Retrieved from <http://dx.doi.org/10.1016/j.techfore.2013.01.010>

Peritz, Rudolph J.R. and Miller, Marc, An Introduction to Competition Concerns in the Google Books Settlement. NYLS Legal Studies Research Paper No. 0/10 #23. Retrieved from SSRN: <http://ssrn.com/abstract=1564363> or <http://dx.doi.org/10.2139/ssrn.1564363>

Polizzotti, M. (2012). "Video: Art Books and eBooks 1" *MetMedia*, Metropolitan Museum of Art Online, USA, September 28, 2012. Retrieved from <http://www.metmuseum.org/metmedia/video/lectures/art-books-and-ebooks-1>

Samuelson, Pamela (February, 2010). Academic Author Objections to the Google Book Search Settlement. *Journal of Telecommunications and High Technology Law*, Forthcoming UC Berkeley Public Law Research. Paper No. 1553894. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1553894

Thomas J. Watson Library Website. About the Digitization Initiative (2016, November 28). Retrieved from <http://libmma.contentdm.oclc.org/cdm/about>

Thomas J. Watson Library Website About Page (2016, November 28). Retrieved from <http://libmma.org/portal/category/about-watson-library/>

Thomas J. Watson Library Website. Collection Development Policy (2016, November 28). Retrieved from <http://libmma.org/portal/watson-library-collection-development-policy>

Thomas J. Watson Library Website. Office of Digital Projects (2016, November 28). Retrieved from <https://sites.google.com/site/departmentofcircumlocution/>

Traditional Fine Arts Organization Website (2005). Retrieved from <http://www.tfaoi.com/aa/5aa/5aa60.htm>

U.S. Copyright Office (June 2015). Orphan Works and Mass Digitization. A Report of the Register of Copyrights. Retrieved from <http://copyright.gov/orphan/reports/orphan-works2015.pdf>

United States Copyright Office Fair Use Index. More Information Page. (2016, April 5). Retrieved from <http://www.copyright.gov/fair-use/more-info.html>

White, Terry (April, 2016). How To Create a Fixed Layout eBook with Adobe InDesign CC. *Terry White's Tech Blog*. Retrieved from <http://terrywhite.com/create-fixed-layout-ebook-adobe-indesign-cc/>

