Libraries in the Digital Renaissance

Law and Policy for Book Digitization

Argyri Panezi

Thesis submitted for assessment with a view to obtaining
the degree of Doctor of Laws of the European University Institute

Florence, 21 December 2017

European University Institute
Department of Law

Libraries in the Digital Renaissance: Law and Policy for Book Digitization

Argyri Panezi

Thesis submitted for assessment with a view to obtaining
the degree of Doctor of Laws of the European University Institute

Examining Board

Professor Lionel Bently, University of Cambridge
Professor Giorgio Monti, European University Institute
Professor Pamela Samuelson, University of California, Berkeley
Professor Giovanni Sartor, European University Institute (Supervisor)

Researcher declaration to accompany the submission of written work
Department of Law – LL.M. and Ph.D. Programmes

I Argyri Panezi certify that I am the author of the work *Libraries in the Digital Renaissance: Law and Policy for Book Digitization* I have presented for examination for the Ph.D. at the European University Institute. I also certify that this is solely my own original work, other than where I have clearly indicated, in this declaration and in the thesis, that it is the work of others.

I warrant that I have obtained all the permissions required for using any material from other copyrighted publications.

I certify that this work complies with the Code of Ethics in Academic Research issued by the European University Institute (IUE 332/2/10 (CA 297).

I declare that this work consists of 104,044 words.

Statement of inclusion of previous work (if applicable):

I confirm that chapter Six draws upon an earlier article I published The Role of Judges in Deciding the Future of Digital Libraries, Global Jurist. ISSN (Online) 1934- 2640, ISSN (Print) 2194- 5675, DOI: http://doi.org/10.1515/gj-2015-0025, January 2017

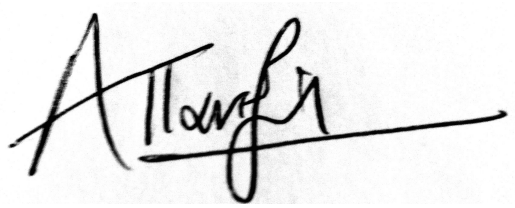
Statement of language correction (if applicable):

This thesis has been corrected for linguistic and stylistic errors.

I certify that I have checked and approved all language corrections, and that these have not affected the content of this work.

Signature and date:

August 10, 2017



ABSTRACT

This dissertation examines the phenomenon of book digitization and the legal challenges for the creation of digital libraries. It explores whether the current copyright framework supports libraries in the digital era as they strive to remain the institutional guardians of our literary treasures. The thesis hypothesis is that the current framework is unsuccessful or inadequate in supporting libraries. To test this hypothesis and identify the legal inefficacies the dissertation examines context, i.e. the history and rationale for applying copyright rules to books and libraries, the copyright rules as applied to libraries before and after digitization was possible, various digitization projects and the practical challenges that digitization brought about, and, the relevant litigation which has started on both sides of the Atlantic covering over a decade now. Furthermore, the thesis explores normative directions of copyright rules, including their exceptions, applicable to libraries in the digital era. It ultimately proposes that among various players claiming this role, libraries, given their institutional functions and capabilities, should be attributed rights to digitization and a favorable legal framework for providing access to digital material.

ACKNOWLEDGEMENTS

I am grateful for the privilege of joining the academic community of the European University Institute, and in particular for the continuous support and guidance of my supervisor, Giovanni Sartor, and my second reader Giorgio Monti. I am indebted for their constant supply of academic direction. I also thank my external examiners Lionel Bently and Pamela Samuelson for their intriguing critique and valuable suggestions for the thesis. I feel honored to have such distinguished academics on my examination board. A sincere thanks to John Palfrey, Michael Klarman, Jason Schultz, James Grimmelmann, Carlotta Alpigiano, Martine Daalder, Thomas Bourke, Serge Noiret, and Melissa Levine for their mentorship, generous sharing of ideas and fruitful discussions. The usual disclaimers apply: all errors and omissions are mine.

I thank the Alexander Onassis Public Benefit Foundation for the continuous support in my education starting from the sponsoring of my legal studies at Harvard Law School in 2009, the States Scholarship Foundation together with the EUI for the sponsorship of my doctoral studies and the research stay at Berkeley Law, and the Foundation for Education and European Culture for generously contributing to my research fellowship at NYU. The NYU Global Fellows program and the academic house of the Jean Monnet Center have contributed to the final shape of the thesis.

The following are conditions *sine qua non* to writing this thesis.

I am grateful for the love and support of my family. My parents for providing home and vitamins, Maria and Doug for providing more white boards. It is thanks to Maria that I ever read law or anything else. Special thanks to Andreas, very close to becoming one year old, because he is my favorite person, and he is currently supervising my writing of the acknowledgements.

I am grateful for the support of my friends and colleagues whom I met during the wonderful years writing the thesis. Naming and shaming only a few: I thank Emma for the Band and our shopping lists, François for the Thole and our travels, Despina for her providing inspiration and gelato, the unbeatable team: Thea, Stavros and Stavros, Giulio for infrastructure, Matteo for avocados, Dónal for brilliance in music and history, Fede for neuro-scientific analyses, Jirka and Karolina and Jofre for chocolate surprises in the library, Gabi, Monika and Marcin, Elias, Chloé, Maria, George, Anna, Miguel and Nele.

Last but not least, I am grateful for having been able to use the distinctive services of many different libraries and academic institutions. First and foremost I thank the incredible team of the sunny EUI Library, the sunlight of which met significant competition only from the reading room and terrace at Berkeley law; the librarians at the imposing fourth floor with the favorite beanbags of Langdell Hall; the

people behind the inspiring New York Public Library, whose lions Patience and Fortitude have kindly provided for titles in slides-presentations as well as jokes for researchers' motivation. There are more libraries that I should mention, including the Biblioteca Nationale Centrale di Firenze, which, despite its unfriendly, barred entrance, offers an unbeatable location next to the Arno and other Dante legacies, the vibrant Biblioteca delle Oblate, and surely more that I now fail to mention. Within library premises, I have enjoyed reading digital versions of books as much as I have enjoyed reading paper books chez moi.

> On peut imaginer le télescope électrique, permettant de lire de chez soi des livres exposés dans la salle 'teleg' des grandes bibliothèques, aux pages demandées d'avance. aceCe sera le livre téléphoné. [Paul Otlet (1868- 1944)]

TABLE OF CONTENTS

Part II: Digitization Rush

INTRODUCTION

0.1 SUBJECT OF STUDY: DIGITIZATION RUSH AND THE CREATION OF DIGITAL LIBRARIES

Libraries amidst a digitization rush

This dissertation examines the phenomenon of book digitization and the legal challenges for the creation of digital libraries. It further explores the normative directions of copyright rules, including their exceptions, which are applicable to libraries in the digital era. The main premise of the dissertation is that the digitization of books has created new value distinct from the one which already exists *vis-à-vis* traditional paper books. Moreover, a number of stakeholders have engaged in legal battles to capture this additional value. The environment has become competitive since the newly-found resource is still 'shining': technology companies, authors, publishers, libraries and other non-profit organizations as well as public administrations are all actors that have participated in a gold rush-like phenomenon. In a context of litigation that has unfolded parallel to efforts for legislative reforms, mainly in the United States and Europe, the thesis concentrates on the position of libraries. Ultimately the thesis argues for the need to rethink the role of libraries as gatekeepers of books, to reflect the evolution of digitized books.

When libraries were approached by various technology companies interested in mass digitization of library collections, they were already aware that these collections are valuable in two ways. First, book collections are valuable because of their rich physical corpus of content; some of it is old, historical or rare. Second, librarians understood that "a lot of money is to be made in digitizing books" and that "when the content moves from physical to digital, its value jumps enormously."[1] This was the reason why companies like Google, Microsoft and Yahoo approached libraries offering the budget that they lacked to digitize on a mass scale.

Within the past three decades there have been a number of important digitization initiatives aimed at digital library projects. Some of them have direct links with 'brick' libraries and others are indirect.[2] First was Michael Hart's Project Gutenberg that began as early as 1971,

[1] Hahn, T.B., 2011. Mass digitization. *Library Resources & Technical Services*, *52*(1), pp.18-26, p. 21.
[2] The term 'brick libraries' will be used interchangeably with the term traditional libraries to point to the physical institution as opposed to the virtual institution of a digital library.

much before the advent of the Internet. Project Gutenberg can be considered the pioneering digital library project. Subsequently, there have been several projects, most of them ongoing. In the United States these projects include the Internet Archive, the Open Content Alliance, the Stanford Digital Library Project, Google Books, which is perhaps the most renowned, HathiTrust and the Digital Public Library of America (DPLA). Europe has a central portal; Europeana, which is supported by the European Union (EU). Europeana connects different national initiatives, such as the pre-existing Gallica of the Bibliothèque Nationale de France – the national library of France, the forming national digital library of the Koninklijke Bibliotheek - the national library of the Netherlands, and other national and regional digital libraries. The thesis includes an extensive timeline of the various digitization initiatives from both sides of the Atlantic, in order to provide a clear picture of the phenomenon, especially its historical context and the most relevant stakeholders.[3]

Even though brick libraries continued to play a central role in digitization, given that they have preserved for centuries the physical corpus of books that is now undergoing gradual digitization, their position is not necessarily strong. More specifically, the institutional position of libraries has not been particularly favored by the copyright rules that apply to them. Indeed, the starting point of this research is to comparatively analyze the role of libraries in the digitization process. Focusing on the libraries' position amidst this digitization 'fever', or 'rush', is extremely interesting especially because most digitization initiatives employ the term 'digital library' when referring to their projects. Not all of these, however, are meant to provide libraries with the full output of digitization. Also in non-library projects, some involvement of libraries is usually necessary but can be indirect. Libraries, for example, provided the books for the Google Books project; however, the digital corpus that was created belongs in its entirety to Google and not to the libraries. According to the terms of their separate contracts with the company, each library participating in the project keeps a digital copy of the corpus of books they provided but, as to be expected, they do not keep records of the books contributed by other libraries.

In sum, libraries find themselves caught in the middle of the digitization rush. The technology companies, as well as other stakeholders, have been interested in capturing the new value of digitized books for different motives. Google, for example, understood the computational value that the corpus has and how it can benefit its search engine and other

[3] *Infra*, chapter six of the thesis.

algorithms. The normative claim of the thesis is that libraries deserve special attention in view of their distinctive roles in societies; roles that range from equalizing access to knowledge, to public service, to providing a vital infrastructure within our social and cultural environment. In other words, the claim is that we need to appraise the benefits of retaining these roles and all the more enhance them in the digital context. Libraries' traditional mandate of curating, preserving, and maintaining access to a cultural record fits well with the possibilities that digitization has opened and with the realization of the dream for a universal digital library.

Defining the concept of digital libraries

To appreciate the concept and the potential of digital libraries we must extend the definition of a library, so that it refers not only to the traditional 'brick' library but also to digital collections. Technology has already proven its potential to radically transform libraries into robust and valuable digital institutions. The concept of a digital library, however, is still not clear and has not been consistently defined. Quite the contrary, there is a tendency to use the term broadly to include any kind of digital collections, usually available online in open or restricted networks, coming from very different institutional structures and operating under very different models. This dissertation will adopt a more restricted notion of a digital library which better captures what legislators actually intend to regulate when they attempt to modify existing rules applicable to libraries or create new rules in order to cover digitization. Therefore, for the purposes of this study, the concept of a digital library mirrors the characteristics of a library in the traditional sense and does not include every digitization project that self-identifies as a digital library project. This is not to say that the initiatives of Google or other non-librarian stakeholders are not relevant. On the contrary, they are relevant and valuable to study in relation to the involvement of libraries in the same phenomenon.

Digital libraries do not need to be separated from the 'old' institutions. Digital collections can exist within or outside of 'old' institution of brick libraries. What exactly defines a digital library, however, can be traced back to its purpose as well as certain specific characteristics, which are both symbolic and functional. Absent these symbolic and functional traits, not every digital collection is automatically a digital library. It is necessary to insist on the scope of the definition especially if we are to make claims about applicable regulation.

Needless to say, libraries fully understand their ongoing relevance in the digital age and want to be able to play a central role in their activities. The institution is undergoing a conscious transformation in order to have a strong digital presence. The process is familiar to libraries, which have not been a static institution but has evolved over time to encompass multiple roles in society in different historical contexts. On this basis, we can say that the digital library does not forecast the end of the traditional one since the physical and digital components do not compete but, rather, coexist.

0.2 The Research Question, Hypothesis and Claim

The Question and Hypothesis

The research question is: whether the current copyright framework sustains libraries in their role as the institutional safeguard of our literary treasures in the digital era. The hypothesis is that it does not. Ample litigation on both sides of the Atlantic since over ten years, together with an emerging body of scholarship in the field, point in this direction. If the current legal framework is unsuccessful or inadequate in supporting libraries, the next question is: where do the main inefficacies lie and what are possible normative directions to remedy identified problems?

The thesis engages in a legal and policy analysis in order to detect these inefficacies. The analysis is mostly doctrinal and consists of two parts. The first part focuses on legal context, i.e., the existing copyright framework applicable to books and libraries. It begins with a historical background and then moves to the normative justifications and goals of copyright law. The second part examines the phenomenon of digitization and the creation of digital libraries, mapping the relevant actors, interests and relationships between them. Creating a timeline of the relevant digitization initiatives this part of the thesis also focuses on the litigation that has unfolded around some of the main digitization projects on both sides of the Atlantic and the issues that the legal battles uncovered. A number of legal questions, mostly related to copyright rules, have been already debated by courts, policy-makers, legislators and scholars. The thesis explores the different views that have been presented on these legal questions, assesses the different proposals to address such questions, and develops original suggestions to policy makers. The last part of this Introduction gives a more elaborate overview of the thesis structure.

Normative claim

The normative goal of the thesis is to critically assess the position of libraries before the eyes of the law, and to suggest to policy-makers some possible improvements, on the basis of the current debate. It argues that with respect to the future of libraries, and more specifically digital libraries, policy-makers need to address two basic questions. First, among the many actors that claim this role they need to determine, which actors, given their institutional functions and capabilities, should be attributed rights to digitization and a favorable legal framework for providing access to digital material. Provided that their answer rests with the libraries, as this thesis proposes, the second question is whether the current copyright rules offer an appropriate framework for them to play a primary role in the ambitious task of digitizing the world's literary treasures. The main normative claim that the dissertation puts forward is that the copyright framework should provide better support to digital libraries as the institutional safeguard of our literary treasures. Libraries should be less burdened by copyright obstacles when it comes, first, to digitization and, subsequently, to providing electronic access to books. Thus, my central thesis is that the position of libraries needs to be strengthened in the digital era both from a legal (copyright) and from a broader policy (or normative) perspective. If entrusted with such mandate, and provided with the support that is needed to accomplish it, libraries can play the role of the institutional safeguard of knowledge also in the digital environment.

Policy choices that actively promote libraries as gatekeepers of the digitized cultural record must be made on the basis of some normative benchmark, i.e. some normative goal or purpose which libraries are most apt to satisfy than other institutions. The benchmark must explain why libraries are to be preferred over other actors, including authors, publishers and their associations, or private companies, and in particular why they are to be given a more favorable treatment by copyright law. To find such benchmark the dissertation starts with the history of copyright and the theory on the rationale for creating a market for books (thus going back to the Statute of Anne in the early 1800's and the inception of copyright) and for reserving a special status for libraries and what they can do with books. Indeed, while library privileges go as far back as the inception of copyright, in each jurisdiction they evolved into copyright exceptions or limitations guaranteed by statutes, doctrines, such as the fair use, and international rules in copyright treaties. Looking into historical and theoretical explanations which are also predominantly enshrined in the copyright discourse, as illustrated in the preambles of

international copyright treaties or in national provisions, one can identify a commonly expressed theme: the goal of encouraging learning and advancing knowledge. If the normative benchmark is to encourage learning and to advance knowledge, in the broadest sense of the meaning, then for a number of reasons libraries are most fit for the job, also the digital space. In fact, the primary mandate of libraries consists in preserving, organizing, and providing access to literary works, all conditions *sine qua non* to encourage learning and succeed in the advancement of knowledge. For libraries to continue in this role in the digital age, at the very least it is important to grant them the right to digitize and to store their collections, which helps counter-balance the rise of private information superpowers and avoid the creation of bottlenecks. Bottlenecks pose a particular risk with regard to access to knowledge.

To remain relevant, libraries today find themselves competing with players from the private and public sector even though their mandates are separate and distinct. Instead of struggling to compete, I argue that they need to be given enough breathing space to be able to keep building digital collections in cost-efficient ways, to have access to their patrons under clear rules of electronic access as well as to e-lend and further collaborate with active authors and publishers in a rational and fair manner. Of utmost importance is to preserve their gatekeeping role. As we will see when we study the most relevant copyright litigation, the *Authors Guild v. Google* case, Google now possesses a mass digital corpus to use and profit from.[4] This was the result of its legal victory following a long battle over its digitization project. If Google is allowed to possess this record, then arguably libraries as institutions dedicated to the preservation and dissemination of knowledge for centuries, if not millennia, should be able at the very least to digitize and safeguard their own cultural record. This argument, *a maiore ad minus*, in support of digital libraries as institutional safeguards to our literary treasures, aims to complement the main normative claim that the role of the libraries should be strengthened. In sum, supporting laws that facilitate rather than inhibit the building of digital collections by libraries, and subsequently the creation of robust digital libraries, will most likely prove beneficial to society and most certainly successful for widespread *access* to knowledge. Libraries have proven their dedication in

[4] *Authors Guild v. Google* is a complex case that lists three main judicial decisions: (1.) *Authors Guild et al. v. Google Inc*., 770 F. Supp. 2d 666 (S.D.N.Y. 2011) [hereinafter: *Authors Guild etal. v. Google Inc*. – 2011 Opinion Rejecting the Settlement], (2.) *Authors Guild, Inc. v. Google Inc*., 954 F. Supp. 2d 282 (S.D.N.Y. 2013), District Court's Summary Judgment on Google's fair use defense [hereinafter: *Authors Guild, Inc. v. Google Inc*. – 2013 Summary Judgment] & (3.) *Authors Guild et al. v. Google, Inc*., No. 13-4829 (2d Cir. 2015) [hereinafter: *Authors Guild, Inc. v. Google Inc*. – 2015 Second Circuit Opinion]. While the mentions throughout the thesis are frequent, the case is mainly analyzed within the timeline of chapter six.

providing such access and have long enjoyed and, at least in theory, still enjoy the respect of society and, to a certain extent, of the law.

0.3 Relevance and The State of Research

Dreams of a 'New Renaissance'

The research for this dissertation started in September 2012. At the time there were at least three major developments taking place in the United States and Europe. First, the above-cited litigation against Google had begun. A parallel case also started against HathiTrust, the consortium of libraries that collaborated with Google for the Google Books mass digitization project.[5] Second, the Digital Public Library of America, a non-profit initiative to create a national digital library in the US, had been announced and was in full swing, with the official launch in 2013. And third, Europeana, the equivalent European digital library project, was launched in 2008, after extensive recommendations by the policy-makers on the EU level. By 2012 the EU had also started to introduce legislation to facilitate digitization and digital library projects, with the Orphan Works Directive, for example, being the first such concrete set of rules.[6]

By 2017, when the dissertation concluded, there have been many more developments. A rich body of case-law has emerged. Besides the American cases already mentioned, this body includes three additional cases litigated before the Court of Justice of the European Union, all of which are relevant to digitization or to libraries' digital services.[7] In chronological order these are: the German case, *Technische Universität Darmstadt v. Eugen Ulmer KG*,[8] the Dutch case, *Vereniging Openbare Bibliotheken v. Stichting Leenrecht*,[9] and the French case *Marc Soulier &*

[5] There are two judicial decisions in the case against Hathitrust: one at first instance and one at the appellate level: (1.) *Authors Guild, Inc. v. HathiTrust, 902 F. Supp. 2d 445, 448 (S.D.N.Y. 2012)* [hereinafter: *Authors Guild, Inc. v. HathiTrust– 2012 District Court Opinion*] & (2.) *Authors Guild, Inc. v. HathiTrust, 755 F.3d 87 (2d Cir. 2014)* [hereinafter: *Authors Guild, Inc. v. HathiTrust– 2014 Second Circuit Opinion*]. Again our case-law analysis is in chapter six.

[6] Directive 2012/28/EU of the European Parliament and of the Council of 25 October 2012 on Certain Permitted Uses of Orphan Works, OJ L 299, 27.10.2012, p. 5–12.

[7] Hereinafter CJEU.

[8] *Technische Universität Darmstadt v. Eugen Ulmer KG*, (Case C-117/13), ECLI:EU:C:2014:2196.

[9] *Vereniging Openbare Bibliotheken v. Stichting Leenrecht*, (Case C-174/15), ECLI:EU:C:2016:856.

Sara Doke v. Premier ministre & Ministre de la Culture et de la Communication.[10] Furthermore, legislators, both at the international and national level, are working towards adapting legal rules to the digital age, particularly those affecting libraries and access to knowledge. In this context, the WIPO Marrakesh Treaty represents a concrete milestone, being the first international treaty providing a clear copyright exception for access to published works for persons who are blind, visually-impaired or otherwise print disabled. It signals the critical momentum to address the question of access to knowledge on an international level.[11] Moreover, the EU has engaged in extended reform efforts of its copyright rules.[12] The articulated objectives of the EU copyright proposals have been, first, more cross-border access to content online, second, the wider and easier use of copyrighted material in education, research, and by cultural heritage institutions, and, third, a better functioning copyright marketplace.[13] The focus on cultural heritage

[10] *Marc Soulier & Sara Doke v. Premier ministre & Ministre de la Culture et de la Communication*, (Case C-301/15), ECLI:EU:C:2016:878.

[11] Marrakesh Treaty to Facilitate Access to Published Works for Persons Who Are Blind, Visually Impaired or Otherwise Print Disabled, adopted on June 27, 2013 and part of the body of international treaties administered by WIPO. The treaty entered into force on June 30, 2016 after Canada's notification (twentieth notification). The contracting parties that have acceded or ratified until today are twenty-seven countries (latest accession is on March 29, 2017) with many ratifications pending including those of the European Union and its Member States and of the United States. The text of the treaty is available at http://www.wipo.int/treaties/en/ip/marrakesh/.

[12] On the EU's efforts to modernize its copyright rules, see the proposals for a new regulation and a new Directive on copyright as well as the proposals to implement the Marrakesh treaty are all available at the Commission's Digital Single Market webpage: https://ec.europa.eu/digital-single-market/en/modernisation-eu-copyright-rules. The proposals include:

i. Proposal for a Regulation of the European Parliament and of the Council Laying Down Rules on The Exercise of Copyright and Related Rights Applicable to Certain Online Transmissions of Broadcasting Organisations and Retransmissions of Television and Radio Programmes, COM/2016/0594 final - 2016/0284 (COD) (14.9.2016), text available at http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016PC0594.

ii. Proposal for a Directive of the European Parliament and of the Council on Copyright in the Digital Single Market, COM/2016/0593 final - 2016/0280 (COD) (14.9.2016), text available at http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016PC0593.

iii. Proposal for a Directive of the European Parliament and of the Council on Certain Permitted Uses of Works and Other Subject-matter Protected by Copyright and Related Rights for the Benefit of Persons Who are Blind, Visually Impaired or Otherwise Print Disabled and Amending Directive 2001/29/EC on the Harmonisation of Certain Aspects of Copyright and Related Rights in the Information Society, COM/2016/0596 final - 2016/0278 (COD) (14.9.2016), text available at http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016PC0596 and is the proposed implementation of the Marrakesh treaty combined with one final proposal for a regulation:

Proposal for a Regulation of the European Parliament and of the Council on the Cross-border Exchange between the Union and Third Countries of Accessible Format Copies of Certain Works and Other Subject-matter Protected by Copyright and Related Rights for the Benefit of Persons Who are Blind, Visually Impaired or Otherwise Print Disabled, COM/2016/0595 final - 2016/0279 (COD) (14.9.2016), text available at http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016PC0595.

[13] See, again, the Commission's Digital Single Market webpage: https://ec.europa.eu/digital-single-market/en/modernisation-eu-copyright-rules

institutions is of particular relevance to this dissertation. On the other side of the Atlantic, the US Copyright Office has also been active in exploring new rules and methods to deal with similar issues. At the beginning of June 2016, the Copyright Office published a Notice of Inquiry inviting interested parties to discuss the revision of the library and archives exceptions in the Copyright Act (17 U.S.C. Section 108).[14] Prior to this, in June 2015, the Office had published another Notice of Inquiry requesting comments on the draft regulation for mass digitization.[15] In the current political climate, however, progress in these legislative initiatives has become dubious, given the more pressing issues of federal funding that libraries suddenly face.[16]

The digital library projects that commenced in the US and Europe are now online and are steadily developing a solid digital infrastructure while having to address various legal and budgetary issues. Google, after being cleared of any copyright liability for its digitization project, as we will see extensively in our case-law analysis, is currently in possession of a large digitized corpus of books but cannot allow users to read most of it. As observed by an article that appeared in April 2017 in the Atlantic under the provocative title "Torching the modern-day library of Alexandria", "somewhere at Google there is a database containing twenty-five million books and nobody is allowed to read them."[17]

In sum, rethinking the rules for access to literary and cultural content online is a topic that enjoys incredible momentum. In light of the developments, there is a strong feeling that both sides of the Atlantic have been ambitiously designing a *Digital Renaissance*, a term borrowed from the Comité des Sages, an expert group appointed by the European Commission to make recommendations on the digitization of European cultural heritage, which named its report "The New Renaissance."[18] This renaissance is spearheaded by a generation who are not digital natives,

---

[14] Library of Congress & Copyright Office Notice of Inquiry, Section 108: Draft Revision of the Library and Archives Exceptions in U.S. Copyright Law, Docket No. 2016 – 4, available at https://www.gpo.gov/fdsys/pkg/FR-2016-06-07/pdf/2016-13426.pdf.

[15] Library of Congress & Copyright Office Notice of Inquiry, Mass Digitization Pilot Program; Request for Comments, Docket No. 2015–3, available at https://www.copyright.gov/fedreg/2015/80fr32614.pdf.

[16] See Templeton J.W., Trump drops the mother of all bombs on libraries, *The Hill*, May 01, 2017 [online], available at http://thehill.com/blogs/pundits-blog/economy-budget/331362-trump-drops-the-mother-of-all-bombs-on-libraries.

[17] Somers J., Torching the modern-day library of Alexandria, *The Atlantic*, April 20, 2017 [online], available at https://www.theatlantic.com/technology/archive/2017/04/the-tragedy-of-google-books/523320/.

[18] *The New Renaissance* - Report of the 'Comité des Sages' reflection group on bringing Europe's cultural heritage online (January 10, 2011), available at https://ec.europa.eu/digital-single-market/en/news/timeline-digitisation-and-online-accessibility-cultural-heritage. The committee comprised Maurice Lévy (Chairman and Chief Executive Officer of advertising and communications company Publicis), Elisabeth Niggemann (Director General of the German National Library and chair of the Europeana Foundation) and Jacques De Decker (author and Permanent Secretary of Belgium's Royal Academy of French language and literature).

but it concerns most significantly the generations born into the digital age. It is within the context of the digital era that we have gradually began to care more about the advancement of learning for the society of digital natives and their access to resources. For digital generations, access to digital material and learning are becoming inseparable; for a digital native, if the material cannot be accessed online, then it basically does not exist. [19] This brings us to one of the central themes of the dissertation: the nexus of digitization and access. On the one hand, the binary is essential for analytical purposes and thus is followed throughout the thesis. On the other hand, there is a strong argument that ultimately the two concepts are inseparable. In the words of Jeremy York, librarian of HathiTrust speaking on behalf of the library consortium in 2012, there is a commonly shared "philosophical belief" among librarians that "the value of preservation is gained through access."[20] In any event, it is important to remember that the ongoing developments will at some point become relevant to digital native generations only who might, by default, opt to access materials in digital forms. Thus, digitization becomes a necessary precondition to access for the digital native generations and for libraries to remain socially valuable.

Last but not least in the series of developments, there has been a notable revival of interest in the future of libraries which extends to their digital and physical presence.[21] The technological advancements that have allowed the digitization of books at a mass and cost-efficient scale have brought the idea of a universally accessible digital library to the fore. What was once only a utopian dream is now technologically possible today. The legal challenges of book digitization and the position of libraries in the digital era is a topic of discussion that has thus far only occurred among a select number of scholars and there remains a lot of ground to cover. The transition that libraries as institutions, together with the book sector in general, undergo has been dramatic. The same is true for digital citizens and their evolving expectations regarding access to material online. Therefore, these are crucial questions to thoroughly study in this timely context.

An emerging body of legal scholarship

[19] See Palfrey, J.G. and Gasser, U., 2013. Born digital: Understanding the first generation of digital natives. Basic Books.
[20] York J., A Preservation Infrastructure Built to Last – Preservation, Community, and HathiTrust, *in* Proceedings of the Memory of the World in the Digital Age: Digitization and Preservation. An international conference on permanent access to digital documentary heritage, 26-28 September 2012, Vancouver, British Columbia, Canada, edited by Luciana Duranti and Elizabeth Shaffer (UNESCO 2013).
[21] See *infra*, chapter seven, under section 7.3, discussing how libraries are far from dead in the digital era.

In the fields of information science and computer science, a relatively rich body of scholarship already exists on the topic of digital libraries, which mainly explores technical issues that arise when building digital collections.[22] Some of the most relevant publications are part of the MIT Press collection on digital libraries and electronic publishing. The benefit of this collection is that it is accessible to a broader audience than only computer science experts. It includes the works of computer scientist William Arms,[23] information scientist Christine Borgman,[24] and information scientists Ann Peterson-Kemp and Nancy Van House together with Barbara Buttenfield[25]. The analyses carried out by these scholars are not legal. Copyright is mentioned in their list of concerns, however.

Older legal scholarship on the question is relatively rare. Notably, it includes a somewhat prophetic article by Pamela Samuelson published in 1995 under the title *Copyright and Digital Libraries*.[26] Samuelson, who wrote much more on the topic in later years, is perhaps the most cited scholar in this field. In the article, Samuelson was the first legal scholar to observe that the many digital library projects underway at the time were raising more complex legal issues compared to legal questions concerning traditional libraries. Notably, Samuelson urged the technical community involved in these projects to participate in the public policy debates that would predominantly involve copyright at that 'early' stage.[27]

The phenomenon of digitization caught the attention of the legal world almost a decade later, around the time when Google announced its ambitious plan to digitize books on a mass scale. That was in 2004.[28] The project immediately sparked litigation which started in 2005 when

---

[22] See for example: Fox, E.A. and Marchionini, G., 1998. Toward a worldwide digital library. *Communications of the ACM*, *41*(4), pp.29-32; Chowdhury, G. and Chowdhury, S., 2002. *Introduction to digital libraries*. Facet publishing; Fox, E.A. and Sornil, O., 2003. Digital libraries.

[23] Arms, W.Y., 2000. *Digital libraries*. MIT Press.

[24] Borgman, C.L., 2000. *From Gutenberg to the global information infrastructure: access to information in the networked world*. MIT Press.

[25] Bishop, A.P., Van House, N.A. and Buttenfield, B.P., 2003. *Digital library use: Social practice in design and evaluation*. MIT Press.

[26] Samuelson, P., 1995. Copyright and digital libraries. *Communications of the ACM*, *38*(4), pp.15-ff.

[27] Id. p.16. As a matter of fact the same legal scholar together with Robert Glushko, information scientist, had published an article already a couple of years earlier (in 1993, with the first version of the article found already in 1991) in the Harvard Journal of Law and Technology discussing Xanadu and the Intellectual Property Rights for Digital Library and Hypertext Publishing Systems: Samuelson, P. and Giushko, R.J., 1993. Intellectual property rights for digital library and hypertext publishing systems. Harv. J. L & Tech., 6, p.237.

[28] Google started the project, under the code name "Project Ocean" in 2002. The company announced Google Books -first called Google Print Service- in 2004 at the Frankfurt Book fair. The company had initially partnered with a number of libraries to digitize their rich collections and was digitizing both public domain and books under copyright. Subsequently, the company was quickly sued by an authors' association, the Authors Guild, and by a number of publishers. The related copyright litigation starts officially in 2005.

the issue received increased public attention. This public response is demonstrated primarily in the press and the blogosphere.[29] The attention of the legal world, which is of most relevance to us, emerges in a series of publications by legal scholars commenting on the phenomenon and also taking positions on the legal issues raised in the litigation. There were many and diverse responses both in support and against the Google Books project and the question of its legality. A good number of articles were written during the extended litigation process, which was a decade long legal battle involving two main cases that were litigated separately and reached the appellate level in the American courts. We will examine these cases in more detail in chapter six, where most of the relevant scholarship also forms part of the analysis. It suffices to note here that one of the central questions under the academic discussion has been the attempted Google Books Settlement. This question gave rise to a number of journal articles.[30]

Generating a series of reactions from authors, publishers, readers, competing technology companies, policy-makers, legislators and judges and, of course, scholars, Google's mass digitization project can be seen as a critical starting point for a series of legal discussions revolving around various fields of law. The discussions mainly centered on copyright law but

---

See Google's December 2004 announcement of the library project: "Google checks out library books," December 14, 2004, available at http://googlepress.blogspot.gr/2004/12/google-checks-out-library-books.html.
See also Wyatt E., New Google Service May Strain Old Ties in Bookselling, The New York Times, 20 October 2004 [online], available at http://www.nytimes.com/2004/10/08/technology/08book.html.

[29] See particularly Markoff J. & Wyatt E., Google is Adding Major Libraries to its Database, The New York Times, 14 December 2004 [online], available at http://www.nytimes.com/2004/12/14/technology/google-is-adding-major-libraries-to-its-database.html; Kelly K., Scan this Book!, The New York Times, 14 May 2006 [online], available at http://www.nytimes.com/2006/05/14/magazine/14publishing.html?_r=0.

[30] To mention only a few of these pieces: Band, J., 2009. The long and winding road to the Google Books Settlement. *J. Marshall Rev. Intell. Prop. L.*, *9*, p.i; Ginsburg C.J., 2015, Google Books and Fair Use: From Implausible to Inevitable? The Media Institute, October 19, 2015 [online], available at http://www.mediainstitute.org/IPI/2015/101915.php; Grimmelmann J. 2012. The Orphan Wars, EDUCAUSE Review, 47(1) [online], available at http://www.educause.edu/ero/article/orphan-wars; Hanratty, E., 2005. Google Library: Beyond Fair Use?. *Duke L. & Tech. Rev.*, *2005*, pp.10-26; Na, N., 2006. Testing the Boundaries of Copyright Protection: The Google Books Library Project and the Fair Use Doctrine. *Cornell JL & Pub. Pol'y*, *16*, p.417; Picker, R.C., 2009. The Google Book Search Settlement: A New Orphan-Works Monopoly?. *Journal of Competition Law and Economics*, *5*(3), pp.383-409; Proskine, E.A., 2006. Google's Technicolor Dreamcoat: A Copyright Analysis of the Google Book Search Library Project. *Berkeley Technology Law Journal*, pp.213-239; Rogers, D.L., 2007. Increasing Access to Knowledge Through Fair Use-Analyzing the Google Litigation to Unleash Developing Countries. *Tul. J. Tech. & Intell. Prop.*, *10*, p.1; Samuelson, P., 2009. Google Book Search and the future of Books in Cyberspace. *Minn. L. Rev.*, *94*, p.1308; Samuelson, P., Legislative Alternatives to the Google Book Settlement'(2010). *Columbia Journal of Law & the Arts*, *34*, pp.697-705; Shah, M., 2006. Fair Use and the Google Book Search Project: The Case for Creating Digital Libraries. *CommLaw Conspectus*, *15*, p.569; Travis, H., 2005. Building universal digital libraries: An agenda for copyright reform. *Pepp. L. Rev.*, *33*, p.761; Travis, H., 2006. Google Book Search and Fair Use: iTunes for Authors, or Napster for Books. *U. Miami L. Rev.*, *61*, p.87; Wilhelm, T.E., 2006. Google Book Search: Fair Use or Fairly Useful Infringement. *Rutgers Computer & Tech. LJ*, *33*, p.107.

also antitrust law and, more generally, information law and policy entered the debates.[31] With the proliferation of stakeholders involved in the Google Books case and the other digitization cases that followed, the pile of legal issues has become more complex. This legal complexity also arises in the corresponding scholarship. Some authors have focused on the fair use doctrine. Some have written about the legal issue around orphan works or are preoccupied with works that are out of print, also known as out of commerce works. Others have focused on the specific issues of licensing and e-lending.[32]

Situating the thesis within the narrower focus on the legal challenges for digital libraries

The existing literature together with a rich body of case-law form a strong basis to begin a broader discussion which is not limited to the legal implications of particular projects of digitization, such as the one of Google and HathiTrust or of the European libraries. The phenomenon of digitization has matured and can be seen from enough different angles. This enables us to shift our focus to its institutional dimension: the library as an institution and its transformation into the digital library. This dissertation endeavors to, first, understand the position of the library as an institution in the midst of this phenomenon, and, second, to conduct a more comprehensive analysis of the notion of the digital library, as a structural product of digitization. Placing the library at the center of the analysis, the thesis reviews the legislative and

[31] On the broad scope of the field of Information Law, defined under US law, see James Grimmelman's recent course-pack project: Patterns of Information Law - Intellectual Property Done Right, Fall 2016 Course pack project [online], available at http://james.grimmelmann.net/ipbook/

[32] See the cited works under note 30, supra.

Also: selected scholarship commenting on the European digitization cases: Afori, O.F., 2013. The Battle Over Public E-Libraries–Taking Stock and Moving Ahead. *IIC-International Review of Intellectual Property and Competition Law*, *44*(4), pp.392-417; Carstens, A.M., 2015. Technische Universität Darmstadt v. Eugen Ulmer KG. *American Journal of International Law*, *109*(1), pp.161-167.; Linklater, E., 2015. Make me an offer I won't regret: Offers to license works on acceptable terms cannot block libraries'"right" to digitize for access on dedicated terminals: Technische Universität Darmstadt. *Common Market Law Review*, *52*(3), pp.813-823; Loewenheim, U., 2015. Boundaries of use of electronic terminals. *Journal of Intellectual Property Law & Practice*, *10*(5), pp.384-387; Matulionyte, R., 2016. 10 years for Google Books and Europeana: copyright law lessons that the EU could learn from the USA. *International Journal of Law and Information Technology*, p.18.

Selected works on the topics of licensing and e-lending: Chiarizio, M., 2013. An american tragedy: e-books, licenses, and the end of public lending libraries. *Vand. L. Rev.*, *66*, p.615; Davies, P., 2013. Access v contract: competing freedoms in the context of copyright limitations and exceptions for libraries. *EIPR*, *35*(7), pp.402-414; Dusollier S., 2014, A manifesto for an e-lending limitation in copyright. *Journal of Intellectual Property, Information Technology and E-Commerce Law* 3; O'Brien, D., Gasser, U. and Palfrey, J.G., 2012. E-books in libraries: A briefing document developed in preparation for a workshop on e-lending in libraries. Berkman Center Research Publication No. 2012-15, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2111396##; Matulionyte, R., 2015. E-Lending and a Public Lending Right: Is it Really a Time for an Update? *EIPR* 38(3), p.132; Müller, H., 2012. Legal aspects of e-books and interlibrary loan. *Interlending & Document Supply*, *40*(3), pp.150-155.

policy framework within which it operates and also discusses the notion of a digital library, its relationship with the 'old' institution of the 'brick' library and its future potential.

Thus far, few legal scholars, who have studied the phenomenon of digitization and the relevant copyright challenges, have focused on libraries. This early literature has played an important role in setting in motion the discussion at hand. We will start by making a brief review of the state of research in this narrower domain and, subsequently, we will situate the present dissertation within the emerging literature on digitization and on libraries. Initially, the study will focus on works that are dedicated to a legal analysis of the phenomenon of digitization, primarily from a copyright perspective. Then it shifts to works that think more broadly about the role of libraries in the digital era from a law and policy perspective.

One of the earliest works to discuss the difficult relationship between copyright law and digitization was the edited volume *Copyright and cultural heritage: preservation and access to works in a digital world*.[33] Edited by Estelle Derclaye, this volume consists of interventions by information technology and copyright law scholars from a number of jurisdictions, including the UK, the Netherlands, Belgium, the US, and Argentina. The book was the result of a conference on the topic in March 2009 and was published in 2010 while certain legislative reforms were underway in Europe.[34] The focus of the publication is on the promises of digitization for the preservation of cultural heritage materials, which is not restricted to books but includes photographs, artistic works, sound recordings and broadcasts. It offers a rich analysis of copyright exceptions for preservation purposes that benefit cultural institutions, such as libraries, museums, archives. The EU and UK norms are its primary preoccupation. Furthermore, the volume provides crucial insight into the practical difficulties that professionals working in cultural institutions face. For instance, these professionals and their patrons exercise self-restraint due to fear of legal infringement, thereby suggesting the need to revise existing law to cater to this reality.

In 2013 Maurizio Borghi and Stavroula Karapapa published *Copyright and Mass Digitization*.[35] Borghi and Karapapa offer a comprehensive overview of the relevant legal

[33] Derclaye, E. ed., 2010. *Copyright and Cultural Heritage: Preservation and access to works in a digital world*. Edward Elgar Publishing.
[34] For example, the book proposes a registry for orphan works similar to the one actually introduced by Directive 2012/28/EU of the European Parliament and of the Council of 25 October 2012 on Certain Permitted Uses of Orphan Works, supra note 6.
[35] Borghi, M. and Karapapa, S., 2013. *Copyright and mass digitization*. Oxford University Press.

questions and technical issues, namely the meaning of transformative use, the treatment of text and data mining as a 'technological transformative use', the problem of orphan works, and derivative rights and entitlements. One of the main objectives of this study, as the authors explain, has been "to reconcile copyright norms and automated text processing."[36] The authors are innovative in their focus on data mining and text processing as well as in their normative conclusions. They begin their conclusions on a critical note, arguing: "This convergence of works and data in mass digital repositories begs for a regulatory framework that transcends the boundaries of established copyright norms." This is a very important observation: blind application of copyright rules do not suffice to tackle the issues that the phenomenon of digitization brings about. In line with their normative suggestions, they propose a complex combination of property rights on personal data and the analogies that can be drawn when applied to mass digitized content as data. For digital libraries in particular they make an analogy to genetic databases. Their analogy goes as follows: "digital repositories of cultural heritage become silos of data to 'mine' in order to extract information or […] a 'cultural genome' of information about virtually all the world's languages and cultures."[37] Subsequently, they study genetic databases, and their regulation, which is mainly based on data protection regulation. Their model is the protection of genome databases in the UK. Thus, they mostly cite the relevant national and EU framework.[38] Through a complex analogy the authors propose a framework of permitted uses of data from a mass digitized corpus and then of uses requiring consent by employing the proportionality principle set down in Article 6 of the Data Protection Directive.[39]

Perhaps the most relevant work on the future of libraries has been written by John Palfrey in 2015. Under the revealing title, *BiblioTech: Why Libraries Matter More Than Ever in the Age of Google*, Palfrey offers a strong and compelling account in defense of libraries in the digital age.[40] In order to make his case, he relies on data proving the importance of libraries in the everyday lives of Americans in the 21st century while simultaneously incorporating a deeper level of understanding on the issue from the library's perspective. His argument is also strongly

[36] Ibid, Preface.
[37] Ibid, p. 149.
[38] Ibid, p.151, where the authors go further into the analogy examining the *Source Informatics* case in the UK.
[39] Ibid. p.155.
The reference is to Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, OJ L 281, 23/11/1995 P. 0031 – 0050.
[40] Palfrey, J., 2015. *BiblioTech: Why libraries matter more than ever in the age of Google*. Basic Books.

grounded on the social value of libraries in democratic societies. He masterfully merges the history of libraries in the United States with their digital future and insists on the importance of the institution in the educational sector. He also devotes his penultimate chapter to the legal issues of copyright and privacy. This book, however, is directed to a much broader audience than legal scholars. Thus, it opts for a less legalistic analysis. Being among the community of founders of the Digital Public Library of America, Palfrey is a central figure in the policy discussions. His views on the role of libraries in the "digitally networked era" are also reflected in a recent *Deadalus* issue on *The Internet* where he discusses design choices for libraries in what he calls the "digital-plus era."[41] Palfrey defines the "digital-plus era" as the era where "not everything must be digital, but materials tend to be born digital and, thereafter, take a variety of forms through which people access them." Part II of this study will focus in-depth on Palfrey's suggestions for critical design choices necessary in this era when we discuss the concept of the digital library.

A more recent work on the topic is *The End of Ownership* by Aaron Perzanowski and Jason Schultz. Published in 2016, this book examines the fluidity of the concept of ownership in an age of technological control of 'things', which include e-books and smart devices.[42] This book dedicates an entire chapter to the "promise and perils of digital libraries."[43] The authors, who have written extensively on digital exhaustion,[44] focus on the consequences that the absence of exhaustion has on the basic functions of the library. They are mainly concerned that libraries do not own their collections but venders retain control over electronic collections. In other words, licensing and e-lending practices are still problematic because of the many restrictions imposed by venders on libraries. Finally, they voice concern about privacy. In the digital age patron privacy is, indeed, a sensitive issue, which Perzanowski and Schultz stress must be taken seriously.

Last of all, relevant to this literature review are recent empirical studies on the topic. Several contributions have been published on the so-called 'shadow' or pirate libraries. Recent studies compiled by economists are pioneering because they introduce empirical data on the uses

[41] Palfrey J. 2016, Design Choices for Libraries in the Digital-Plus Era, *Daedalus* 145(1): 79–86.
[42] Perzanowski A. and Schultz J., 2016, *The End of Ownership: Personal Property in the Digital Economy*, MIT Press.
[43] Id. At p. 103 et. seq.
[44] See Perzanowski, A. and Schultz, J., 2010. Digital exhaustion. *UCLA L. Rev.*, *58*, p.889; Perzonowski, A. and Schultz, J., 2014. Legislating digital exhaustion. *Berkeley Tech. LJ*, *29*, p.1535.

of shadow libraries accessible on the web.[45] These studies link the phenomenon of book piracy to the inefficiencies of copyright laws and indirectly, I believe, strengthen the case for a better legislative framework applicable to libraries that wish to operate 'in the light', rather than in the shadows. The same applies to economic studies that look into the phenomenon of book unavailability due to copyright restraints.[46] Moreover, there is a useful number of papers and policy reports prepared by academic centers on specific topics related to digitization and digital libraries. These works, which fall between the academic and policy arena, have been extremely helpful to identify the most pressing legal issues. Notable are the following issues, presented in chronological order: (i) A briefing document compiled by John Palfrey, Urs Gasser, and David O'Brien in 2012 ahead of a workshop on e-lending in libraries at the Berkman Center for Internet and Society at Harvard University.[47] The authors present a comprehensive study on the *status quo* of licensing models for e-books in libraries; (ii) A report prepared by scholars at the University of Amsterdam in 2012 offering one of the first comprehensive legal and economic analyses of e-lending by public libraries[48]; (iii) Several contributions outputs of the Digital Library Copyright Project at Berkeley Law, many of which address the issue of orphan works (2012-2013);[49] (iv) Similarly, reports as part of the Mobile Collections Project at Cambridge

[45] Balázs, B., 2011. Coda: A short history of book piracy. *Media piracy in emerging economies*, p.399; Bodó, B., 2015. Libraries in the post-scarcity era. in H. Porsdam (Ed.), *Copyrighting creativity: creative values, cultural heritage institutions and systems of intellectual property* (pp. 75-92). Farnham: Ashgate. See also Karaganis J. 2017, *Digital Piracy. Handbook of Digital Media and Social Practice*. Springer; also the forthcoming work: Karaganis J. ed. (2018). *Shadow Libraries: Access to Educational Materials in the Digital Age*. MIT Press.

[46] Heald P. 2013, How Copyright Keeps Works Disappeared, Illinois Public Law Research Paper No. 13-54; Heald P., 2014. The Demand for Out-of-Print Works and Their (Un)Availability in Alternative Markets, Illinois Public Law Research Paper No. 14-31; Nagaraj, A., 2016. Does Copyright Affect Reuse? Evidence from the Google Books Digitization Project, [online] Available at SSRN: https://ssrn.com/abstract=2810761.

[47] O'Brien, Gasser, and Palfrey, supra note 32.

[48] R. van der Noll, K. Breemen, V. Breemen, B. Hugenholtz, M. Brom & J. Poort, Online uitlenen van e-books door bibliotheken. Verkenning juridische mogelijkheden en economische effecten, in opdracht van het Ministerie van Onderwijs, Cultuur en Wetenschap, Amsterdam, 2012, available at http://dare.uva.nl/search?identifier=743e4ef8-1ae8-45a7-becd-d123354f8ec0. On the same topic, there are more relevant publications from the Institute for Information Law (IVIR) at the University of Amsterdam at a later stage.

[49] Hansen D. 2011. Orphan Works: Definitional Issues, Berkeley Digital Library Copyright Project, White Paper No.1 [online], available at http://ssrn.com/abstract=1974614; Hansen D. 2012. Orphan works: mapping the possible solution spaces (Berkeley Digital Library Project, White Paper No 2 [online], available at http://ssrn.com/abstract=2019121; Hansen D. 2012. Orphan works: causes of the problem, Berkeley Digital Library Copyright Project, White Paper No 3 [online], available at http://ssrn.com/abstract=2038068; Urban J.M. 2012. How fair use can help solve the orphan works problem, 27 *Berkeley Tech. L.J.*; Berkeley Digital Library Copyright Project Comments in Response to the U.S. Copyright Office Notice Inquiry on Orphan Works and Mass Digitization (Feb. 4, 2013), available at https://www.law.berkeley.edu/files/Berkeley_Dig_Lib_Copyright_Proj__REPLY_Comments_to_Copyright_Office_OW_NOI_FINAL_with_appendix.pdf; Hansen D., Hinze G., and Urban J. 2013. Orphan Works and the Search for Rightsholders: Who Participates in a 'Diligent Search' Under Present and Proposed Regimes, Berkeley Digital

University, including Eleonora Rosati's *Copyright issues facing early stages of digitization projects in 2013,* which explain the issues libraries and museums in the UK have encountered when engaging in digitization, initially for the purposes of preservation.[50] The report places heavy emphasis on the problem of orphan works, the solutions considered by the UK legislators and the, then, 'fresh' Orphan Works Directive; (v) The most recent literature review by David Hansen deals with the digitization of orphan works. Published in 2016 as part of a Harvard University Library project (DASH – Digital Access to Scholarship at Harvard), the review explores legal strategies to reduce risks for open access to copyrighted orphan works.[51] Earlier publications of this project explored the range of open access options for digitization projects.[52] Last of all, we must mention the 2016 publication of the Copyright Review Management System team at the University of Michigan Library team, funded by the Institute of Museum and Library Services for a project called *Advancing Digital Resources*: (vi) Finding the Public Domain: Copyright Review Management System Toolkit.[53]

Thesis contribution

The goal of the study is to provide a comprehensive analysis of the legal issues pertaining to the involvement of libraries in the digitization of literary works and in the distribution of digitized content. It is written as an academic exercise, rather than a policy intervention. Although the geographic focus of the thesis is on Europe and the United States, mainly because of the sources used and the author's academic affiliations, it is not intended to concentrate solely on these

---

Library Copyright Project, White Paper No. 4 [online] available at http://ssrn.com/abstract=2208163; Hansen D., Hinze G., and Urban J. 2013 What Constitutes a Diligent Search Under Present and Proposed Orphan Works Regimes, Berkeley Digital Library Copyright Project, White Paper No. 5, 2013 [online] available at http://ssrn.com/abstract=2229021; Hansen D., Hashimoto K., Hinze G., Samuelson P., and Urban J. 2013. Solving the Orphan Works Problem for the United States, 37 *Columbia J. L & Arts* 1.

[50] Rosati E. 2013. Copyright issues facing early stages of digitization projects, Mobile Collections Project [online], available at http://www.digitalhumanities.cam.ac.uk/CopyrightissuesfacingearlystagesofdigitizationprojectsIRosati.pdf. More reports from the project at http://www.digitalhumanities.cam.ac.uk/mobilecollections.

[51] Hansen, D. 2016. Digitizing Orphan Works: Legal Strategies to Reduce Risks for Open Access to Copyrighted Orphan Works. Kyle K. Courtney and Peter Suber, eds., Harvard Library, available at http://nrs.harvard.edu/urn-3:HUL.InstRepos:27840430.

[52] See notably Suber, P. 2011. Open Access for Digitization Projects. In Going Digital: Evolutionary and Revolutionary Aspects of Digitization, ed. Karl Grandin, 70-93. Stockholm: Nobel Foundation, available at https://dash.harvard.edu/bitstream/handle/1/14976387/KVA_Going_Digital_webb%202.pdf?sequence=1.

[53] Finding the Public Domain: Copyright Review Management System Toolkit, Melissa Levine, Richard C. Adler, Justin Bonfiglio, Kristina Eden, and Brian S. Hall. DOI: http://dx.doi.org/10.3998/crmstoolkit.14616082.0001.001.

jurisdictions. Instead, its aim is to make general normative claims that are potentially applicable to different jurisdictions.

The thesis endeavors to fill specific gaps in the existing legal literature. First, and most importantly, so far limited attention has been paid to the legal position of libraries regarding digitization, which is curious when we consider the unique institutional challenges facing libraries in the digital age. Furthermore, there is still much to be said about the institutional role played by digital libraries in the networked environment, which from a normative basis urges us to ask whether the legislative environment can undergo reform to facilitate this reality. Placing itself among the emerging scholarship sketched already, the present dissertation intends to add depth to the discussion about the institutional role of digital libraries. Second, and on a more practical level, existing scholarship has been written during the litigation stages of the main digitization cases, thus pending the final decisions of these cases. This thesis will take into account also the final decisions and base upon them its normative claims. However, the topic remains very contemporary and, thus, every new court decision or other legal development shapes the entire picture. This is arguably a shortcoming of the present thesis, as with every thesis that is written while the sand in its field is still on the move.

It is worthwhile to say a few additional words about the novelty of the normative claim. The starting point is a doctrinal analysis of the phenomenon of digitization and the legal battles that quickly ensued. While the analysis is legal and mainly revolves around copyright norms, copyright history and theory add context. So far little attention has been devoted to the interpretation of the context in which libraries have been regulated and the difficulties to adjust to the digital age. Furthermore, the dissertation will attempt to propose normative direction on the basis of a specific benchmark: that is the encouragement of learning and, tied to that, the advancement of knowledge. This benchmark is chosen on the basis of the longstanding articulated rationale for copyright privileges that benefit libraries. The novelty of the thesis rests on the normative stance as well as the decision to assess the phenomenon of book digitization by placing the libraries center-stage. In addition to the centrality of libraries, books are given particular attention throughout the thesis. This is because the very essence of libraries' existence rests on books and they are the means by which our normative benchmark can fulfill its practical purpose. A final comment must be made with regard to the normative benchmark which sets the tone for the remainder of the thesis. A strong emphasis is placed on the public purpose served by

regulating libraries, both in their physical and digital presence. The same public purpose remains relevant in reading copyright rules as applied to books even if they are meant to have a commercial function. There is also criticism to be made to the one-size-fits-all system of copyright, which is characteristic of the national and international approach. For this reason, this dissertation does not claim to put forward a case that is applicable to intellectual content other than literary works as, for example, music or other performances. The study concentrates on libraries, and books stored in libraries. Thus normative claims about the market sector of books, or other markets, are beyond its scope of this study.

To sum up, the contribution that the present dissertation hopes to make is twofold. Firstly, at a descriptive level, it attempts to analyze an ongoing phenomenon of digitization of books in a manner that places the libraries in the center of the picture, while exploring actions and motives of other stakeholders. Secondly, drawing practical lessons that emerge from this phenomenon, it proposes to strengthen the position of libraries at a normative level.

0.4 METHOD AND SOURCE MATERIAL

The legal focus of the dissertation is on the copyright rules applicable to books and to libraries. While demonstrating an appreciation for historical context, the approach is largely doctrinal and takes a normative (*de lege ferenda*) view. Our concern is whether the legislative framework, while remaining both efficient and effective in reaching its targets, rises to the technological challenge and the technological potential. Indeed, the topics of digitization and access to digitized materials rest in the intersection of law and technology. Technology has shifted our understanding and experience of culture and, in particular, access to content. Furthermore, technology has introduced a distinct culture itself, separating generations born into the digital age and those who preceded.[54] At the same time, intellectual property laws have presented concrete problems if we are to realize the full potential of technology. We have already seen the battles in other sectors, such as the music industry or the film industry, and how they have had to adjust to their online access. Access to literary works is, I maintain, a distinct question and for this reason I decided to study it separately and focus on the institutional

[54] Palfrey and Gasser, supra note 19.

dimension of digital libraries. In sum, the objective is to reconcile the law in theory and practice and the potential of technology.

To contextualize my doctrinal analysis I conducted empirical research by engaging with libraries in three different contexts. These experiences have influenced the study as well as the decision to focus on certain issues; primarily the issue of digital exhaustion and e-lending, the orphan works problem as well as the focus on library collaboration, inter-operable platforms and shared metadata. First, I worked with the library of the European University Institute looking at the practical questions of e-licensing and subscriptions of an academic library. Second, I have closely followed the developments of the Digital Public Library of America since 2012 and participated in their discussions, especially with regard to the orphan works and international copyright rules.[55] Third, I participated in drafting a response to the public consultation on the review of the EU copyright rules submitted to the European Commission, DG Internal Market and Services by academics in collaboration with the Bocconi Library. The submission pointed to the need to reform copyright rules in order for libraries to benefit from new opportunities that the digital age offers. From that perspective, I have been following the European reform agenda for copyright, very much still an ongoing process. This is not to say that I employed empirical methods for my study, which remains doctrinal; it is only to say that the study is highly influenced by my practical experience and predominantly the multiple exchanges I had with librarians over the course of the research.

Methodologically the dissertation employs arguments from three fields: (i) copyright law, (ii) law and economics scholarship and (iii) law and technology scholarship. Furthermore, I pursue an interdisciplinary approach to the source materials. Besides the works of legal authors whether academics, judges and also legislators or policy-makers, it employs a number of works from other disciplines, mainly economics, history, information science and computer science. It also cites a large number of contemporary news reports to contextualize the primary and secondary sources. While the thesis cites secondary resources from all the aforementioned fields and disciplines, the primary sources are copyright law and judicial decisions. When it concerns litigation, some of these sources include the parties' briefs, *amicus curiae* briefs and public announcements of the litigants and other parties involved in the litigation.

[55] Digital Public Library of America Inaugural 2013 Fest (Boston MA, October 24-25) and 2016 Fest (Washington DC, April 14-15), details and material available at https://dp.la/info/get-involved/dplafest/.

The next, and final, part of the introduction presents the structure of the thesis, which is divided into two parts. Here, I couple each part of the thesis with its corresponding sources in order to give an overall idea of how each part employs a different type of source. The first part looks at materials on the history of copyright, copyright theory, including economic analysis that justifies the rules and their duration, current copyright laws applicable to books and libraries, including surrounding policy reports and legal scholarship, and finally, literature on the institutional dimensions of libraries combined with a law and economics view on the role of institutions in book de-commodification.

Part two focuses on the phenomenon of digitization. The sources include legal and economic scholarship dealing with copying and book availability in the market and the phenomenon of pirate libraries. This part also analyzes a number of digitization initiatives and creates a comprehensive timeline of these developments. The timeline is a diverse set of online sources and news reporting combined with information from case-law and legal scholarship. Within the timeline there is also a case-law analysis. It includes the judicial decisions as primary sources and secondary legal sources in order to contextualize and analyze the judicial texts. The case-law is crucial also for the normative argument in the concluding chapter of the thesis. In the conclusion, various sources are cited which belong to a broader law and technology field in order to make the argument for the role of libraries, and digital libraries, in a networked environment. Ultimately, the dissertation aims to reach conclusions that transcend the inherent limitations of copyright law and become part of a broader discussion on how the legal framework applicable to libraries and digital libraries in particular fails to match the technological potential. Arguably copyright law does not suffice in succeeding in fulfilling this goal. One needs to look at principles emerging from other related legal fields, such as copyright theory, competition, law and economics, and also law and technology, to enrich the monolithic copyright focus. This is the logic behind employing a diverse set of sources in each chapter, as outlined above.

0.5 THESIS STRUCTURE

The dissertation has two parts. The first *On Books, Libraries and the Needs of the Poor* is where most of the historical and theoretical analysis on copyright, and its application on books and libraries, is found. Its purpose is to understand the background of the legal framework that is

relevant to the dissertation topic: a copyright framework as was conceptualized for 'brick' libraries and for physical books.

Part two, *Digitization Rush*, is devoted to the phenomenon of digitization and the creation of digital libraries. The phenomenon of book digitization constitutes the backbone of the entire dissertation, and is therefore given considerable attention. This part also looks at how, as a result of digitization, books are being transformed and so is the institution of libraries, which were traditionally dedicated to the role of collection, preservation and systematization. It includes an extensive timeline of the most important digitization initiatives in the United States and Europe. The gold rush metaphor is employed to describe a competitive phenomenon, the rush to digitize and capture the value of a new kind of resource leading to long and complex legal battles. The battles took place among stakeholders that either engaged in digitization projects or claimed pre-existing property rights to oppose to what they perceived to be rivalry projects.

To provide greater detail, the structure appears as follows:

Part I: On Books, Libraries and the Needs of the Poor

Our legal systems have treated books predominantly as market goods, at least for a period of time considered worthy to do so. Afterwards, books are supposed to sit quiet and alone, end up living a far less exciting life, run out of print, and ultimately fall into an obscure state called public domain - most probably sitting in dusty library stacks. Now, how accurate is this sad portrayal of the legal framework of copyright as applied to books? Has it always been so and what role do libraries play in this predominantly market structure?

The dissertation starts with a historical inquiry into the first copyright laws aimed at creating a book market as well as the first laws guaranteeing the rights of libraries. Part I looks into the history to understand the theoretical justifications for creating exclusive privileges for books and similar privileges for libraries. The part is divided into four chapters as follows: (i) A marketplace of ideas and a marketplace for books; (ii) Justifications for copyright protection on books; (iii) Book de-commodification: libraries as reserved places for the legal distribution of books outside of markets and (iv) What can libraries do with their books?

Scholars who study the rationale behind the legal structure of copyright commonly make reference to the story of the first copyright legislation: the 1710 Statute of Anne in England. At the beginning of the first chapter, we will refer to the Statute, its express motivation for 'the

encouragement of learning' and the rights it provided not only to authors and right-holders in general but also to libraries. Subsequently we will move to modern justifications for copyright protection on books and see how market considerations and utilitarianism prevailed. It is noted that the focus shifted away from the encouragement of learning and therefore the rhetoric about libraries also shifted radically: in copyright law we no longer refer to library rights but to library exceptions. In England at the time of the Statute of Anne, very few libraries were accessible and those that were only to a small elite. In view of the Statute, these libraries enjoyed special rights. To these rights we owe the preservation of invaluable collections until today. The same rights were not at all self-evident and were not granted when the legislation for the first public libraries passed. These libraries were not accessible to the selected elite whose learning was encouraged by the Statute of Anne; they were the libraries for the poor and enjoyed far less privileges. Copyright laws until today carry this tension between elite and universal access to knowledge. It is helpful for us to always keep this elite dichotomy in mind.

Apart from the historical and theoretical analyses of copyright as applied to books, Part I examines how the role of libraries evolved into reserved places for the legal distribution of books outside of markets. It further includes a chapter investigating what the law allows libraries to do with their books. This chapter reviews contemporary rules that provide copyright exceptions and limitations in favor of libraries and touches upon the e-lending puzzle. It then proceeds to examine a flexible categorization of books in accordance to their copyright status: books under copyright include the two 'difficult' groups: the first group is out of print works, also referred to as out of commerce works and the second group concerns orphan works. The remainder of books belong to the public domain. The copyright status of books has always been relevant to their accessibility and has become most relevant today with the possibilities that digitization opens.

Overall, the first part seeks to set the stage for the following part which focuses on the phenomenon of digitization, the creation of digital libraries and the legal controversies that arise in the process. The title of Part I implies certain things with respect to the historical background and the corresponding discourse on books and libraries. First, that books are considered 'special' goods and historically have not been treated like meat, bread or potatoes. The comparison to meat, made in our very first chapter, is rooted in the Jurisprudence lectures of Adam Smith where he justifies exclusive privileges on the vending of new books as opposed to the selling of meat. Even though copyright was created so that books become goods traded in markets, they

have particular characteristics that have for better or for worse justified special legal treatment in many respects: their protection against censorship, their pricing, their taxation and of course their guardianship by institutions, such as libraries. As cultural products books can therefore be seen as merit goods. The legal and economic discussion on how we best treat merit goods ultimately leads to a political discussion on equal access to knowledge and the role of libraries in book de-commodification and the redistributive function that proves vital in unequal societies.

Part II: Digitization Rush

The second part focuses on the phenomenon of book digitization and the various digital library projects that started in America and Europe. This part is again systematically divided in four chapters, the fourth serving also as a the conclusion as well: (v) How the Internet did not kill the libraries; (vi) A timeline of digitization initiatives and the relevant case-law affecting digital libraries; (vii) Adverse consequences of digitization on books and on libraries and (viii) Conclusion – What's the rush?

The debate over the effects that digitization has on books, the book market and libraries involves economists, lawyers and policy-makers as much as librarians and all book-lovers. Indeed, there is a growing body of scholarship that deals with the future of books and the future of libraries. The subject allows viewpoints from very different disciplines. While there is an obvious risk leaving out some disciplines, I will mention only those that I found particularly helpful when thinking about this topic in the second part of the dissertation: law, economics, librarianship and information sciences and computer science.

This part starts with chapter five, which discusses how the Internet did not kill libraries as was perhaps first feared. On the contrary, the history of digitization and the prospect of creating digital libraries can actually be connected to the history of the development of the Internet. Indeed, digitization promises a new future for libraries that can actually evolve into digital libraries, without necessarily losing their physicality, and even develop a new infrastructure at the content layer of the Internet. The chapter includes an anecdote about how the vision of digital libraries is in fact connected to the development of the Internet. At the same time traditional libraries remain very relevant in the digital age. In addition, chapter five includes an analysis of the economics of copying. The economics of copying have always been relevant to copyright theory, at least since the time of the photocopying machines. Being an advanced and efficient

method of copying, digitization seems to disrupt the balances that copyright laws claim to maintain between access and incentives for literary creation.

Chapter six creates a timeline of various digital library projects and significant digitization efforts in the US and in Europe. The economics of digitization indicate that the value of books when (mass) digitized far exceeds the sum of the value of the original, which is why mass digitization has been more controversial than older, less advanced copying methods. Also, it is arguably the main reason behind public and private stakeholders, as seen in the timeline, engaging in digital library projects. Unsurprisingly, some of them are also engaged in fierce copyright battles over the digitization of books, which have lasted more than a decade now. A rush to capture the newly-found gold, the digitized book, was quickly followed by legal battles over the new resource. A legislative drive to deal with the numerous legal challenges that were first faced by courts quickly ensued. Thus, the timeline in this part includes a detailed examination of the various lawsuits and court decisions starting with (1) the long and complex Google Books case.[56]

Google triumphed over the Authors Guild, an association of authors, and managed to win a copyright case of mass digitization on the biggest scale. It won solely on the legal basis of fair use. Since then, Google is the most powerful player in the digitization arena and today holds a digital corpus of millions of books that it digitized from some of the biggest library collections in the world.[57] Sitting on its Californian goldmine, Google uses the digitized corpus to feed its search engine and perhaps will employ it for many more uses in the future.[58] Four more relevant cases follow: (2) the case that the Authors Guild brought against HathiTrust, the consortium of libraries that collaborated with Google in digitization;[59] (3) the case against the Technical University of Darmstadt in Germany, brought to the European court by the publisher Eugen

[56] *Authors Guild v. Google*, supra note 4.
[57] Under the 'Library Project' Google initially partnered with Harvard University, Stanford University, University of Michigan and Oxford University, as well as the New York public library to digitize their collections. Since then many more library partners have joined. Google says that it now works with over 40 libraries around the world to digitize their collections. On its website, it gives a sample of its partners: see https://books.google.com/googlebooks/library/partners.html.
[58] See Lewis-Krausdec, The Great A.I. Awakening How Google used artificial intelligence to transform Google Translate, one of its more popular services — and how machine learning is poised to reinvent computing itself, *The New York Times Magazine*, December 14, 2016 [online], available at https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html.
[59] *Authors Guild, Inc. v. HathiTrust*, supra note 5.

Ulmer;[60] (4) the case initiated by the Dutch association of public libraries requesting recognition of their right to e-lend[61] and (5) the most recent case, which the French Conseil d' État sent to the European court, concerns national legislation that allowed collecting societies to proceed with the digital exploitation of out of print works.[62] The chapter concludes with the first observations to be made of the phenomenon that the timeline as a whole portrays

Chapter seven continues with a better understanding of the consequences that the phenomenon of digitization has on books and on libraries. Although the consequences focused on are, on their face, negative, the chapter cautions against catastrophic scenarios or predictions. It first looks at the consequences that digitization has on books; how the medium is transformed. With digitization, it is not only copyright but also the very concept of the book that seems to be disrupted or even 'threatened'. Digital versions of books, now commonly referred to as e-books, lack physicality. A tension, therefore, arises in their legal treatment. Physical books are objects of property and at the same time their content is protected by a certain type of intellectual property, whose development we traced in the previous chapter. Along with many more goods, however, books in the digital era lose their *in rem* property features. While this transformation is taking place, our legal structures do not appear to be consistent with the treatment of the book: for example, it is not clear whether paper books and e-books are substitute goods; it is equally unclear whether exhaustion or else 'first sale' applies to digital books and under what circumstances. Having looked at the theories that justified the legal treatment of books and libraries in the analogue era, it seems that there is a need for new theories. We generally observe many challenges when we apply the same rules applicable to physical books or the same normative and economic rationales of physical books to their digital versions.

Subsequently this chapter focuses on a seemingly bizarre failure: in spite of the high expectations that come with digitization and the Internet, there are significant gaps in book availability. Thus, when books appear inaccessible, especially as a result of copyright, there is a visible demand for access expressed in various ways within the Internet society. These ways include online file sharing, the phenomenon of book piracy, and the creation of 'shadow libraries'. The chapter examines at the phenomenon of shadow libraries which links us back to the importance of maintaining robust libraries as spaces dedicated to giving access to knowledge

[60] *Technische Universität Darmstadt v. Eugen Ulmer KG,* supra note 8.
[61] *Vereniging Openbare Bibliotheken v. Stichting Leenrecht*, supra note 9.
[62] *Marc Soulier & Sara Doke v. Premier ministre & Ministre de la Culture et de la Communication*, supra note 10.

outside of the market. Digitization underscores a clear demand for book de-commodification and access. Ergo the need to focus on digital libraries, the legal challenges they face under the current copyright framework and normative directions for rules applicable to them.

Concluding Chapter

The eighth and final chapter evaluates the facts presented in Part II of the thesis, including the relevant litigation, as a sequel of 'digitization wars' part of the so-called 'copyright wars'. The employment of the 'war' metaphor goes hand-in-hand with the property metaphors that we often see in copyright law. Think, for example, of 'orphan works' as unclaimed or abandoned land. The chapter introduces yet another metaphor: that of the gold rush in California. The gold rush symbolizes not only a competition over resources but ultimately also over new lands. It is a 'war' over territory between the old world and the new world. It is:

> The discovery of gold in California coincided with the acquisition of that territory by the United States after its "manifest destiny" war of conquest over Mexico, and the previous Mexican elite suffered grievously under both headings. Under the old Mexican law, land titles did not confer mineral rights but the new US administration, wedded to notions of untrammeled capitalism, immediately struck out this provision. The result was an explosion of land speculators and shyster lawyers, swindling and threatening the previous Mexican elite out of their property. [63]

In our metaphor, digitization unveiled a new value: the new gold. Our Mexican elite perhaps consists of authors, publishers and other proponents of extensive copyright terms and restrictions to digitization. New explorers of gold, primarily the technology sector, rush to grab the new value and the previous elite feel tricked and threatened and basically sue in an effort to retain their old privileges. In this narration, the metaphor is an obvious oversimplification. The

[63] McLynn F. Review of The Age of Gold: The Story of an Obsession that Swept the World (by HW Brands), The Guardian, May 6, 2005 [online], available at https://www.theguardian.com/books/2005/may/07/featuresreviews.guardianreview11.

relations and tensions between the various stakeholders involved in a series of litigation and policy debates around digitization are manifold and complex.

In the context of this competitive phenomenon, that results in winners and losers both in the short run and in the long run, the concluding chapter reviews the thesis findings and then takes a normative stance: as regards the future of digital libraries, policy-makers need to rethink two basic considerations. First, they need to rethink whether the current copyright rules provide an appropriate framework to continue the ambitious task of digitization of the world's cultural heritage. Second, they need to reconsider which player is institutionally equipped to enjoy clear rights of digitization. This thesis argues that the library is the key player. If trusted with such a mandate, libraries can play the role of institutional guardian of knowledge in the digital environment. Setting a normative benchmark can help us appraise the suitability of the various institutions and actors that claim the same role. If that normative benchmark is to encourage and advance learning among societies of digital natives, then there are a number of reasons why libraries are institutionally suitable for the job. At the very least it is important to grant libraries the right to digitize and store their collections to counter-balance the rise of information superpowers and to avoid the creation of more bottlenecks, particularly when it comes to access to knowledge.

PART I

I.

ON BOOKS, LIBRARIES AND THE NEEDS OF THE POOR

Go to Rome and see! In the mansions of the great prelates there is no concern save for poetry and the oratorical art. Go thither and see! Thou shalt find them all with the books of the humanities in their hands and telling one another that they can guide men's souls by means of Virgil, Horace, and Cicero.... The prelates of former days had fewer gold miters and chalices, and what few they possessed were broken up and given to relieve the needs of the poor. But our prelates, for the sake of obtaining chalices, will rob the poor of their sole means of support. Dost thou not know what I would tell thee! What doest thou, O Lord! Arise, and come to deliver thy Church from the hands of devils, from the hands of tyrants, from the hands of iniquitous prelates […]

Girolamo Savonarola[64]

CHAPTERS' OVERVIEW

One cannot speak about libraries without speaking about books. Books are what make libraries valuable. This also remains the case with digital libraries. The title of the dissertation signals that books are a focal point, which, requires analysis. In order to make a normative claim with regard to the legal challenges for the creation of digital libraries, we will first devote our attention to

[64] From the volume: Borelli A., Pastore Passaro M., Beebe D., (ed.) 2006, *Selected writings of Girolamo Savonarola: religion and politics*, 1490-1498, Yale University Press.

studying legal and economic aspects of the book, which is the primary unit of analysis in this thesis. Books are the main resource that libraries hold and offer access to. Nowadays, libraries give access to content in many more mediums, including CDs, DVDs, or even old cassettes and videocassettes. Most libraries also keep records of newspaper articles and entire newspapers and journals. While all of these mediums are worth looking at separately, our focus remains on the book, both as a symbolic medium with historical weight and as the dominant medium when it comes to library operations. One can barely think of a library without thinking of books and *vice versa*.

The first chapter focuses on the special legal treatment of the physical book. Books, their trade as well as their preservation in libraries, have been regulated by the first copyright laws. Books have not always been protected by copyright. The exclusive rights, or privileges, that copyright laws afford to right-holders is a legal construction traced back to the 1700's and the celebrated Statute of Anne in England. Indeed, the first chapter, entitled "**A marketplace of ideas and a marketplace for books**", is to provide the historical context of copyright laws, which were first conceptualized to apply to literary works. The chapter aims at answering some basic questions: What is special about books and makes them a different 'product' to other products in the market? Why is the regulation of books necessarily connected to that of libraries? And what is the underlying rationale behind regulating both books and libraries? As discussed in the introduction, we will examine the history of copyright in order to discern a normative benchmark against which we can subsequently assess the evolution of copyright laws applicable to books and to libraries. We can detect our normative benchmark in the original copyright laws rationale which contemporary copyright laws have also in theory retained: the advancement of learning and the encouragement of knowledge. This, of course, applies to literary works when regulated by copyright and not necessarily to other types of content. For historical reasons this first chapter focuses on English sources and the particular and very interesting temporal context of the Statute of Anne: it was a few centuries after the invention of the press and a time when book trade was becoming widespread in Europe.

The second chapter of the thesis, under the title **"Justifications for copyright protection on books"**, explores the theories that have been employed to justify shorter or longer terms of copyright protection. The theories have moral and economic roots traceable to property theory and utilitarianism. Employing these theories as a prism, the chapter concludes with an overview

of the debate over the optimal copyright length; an issue that is still far from solved. Centuries after the first copyright statute, today's book market enjoys lengthy copyright terms. Concerns over censorship, which were the primary worry when copyright was first introduced, are fading. At the same time, concerns over the negative consequences of copyright, or at least the length of protection, have been on the rise. These concerns face a strong legislative structure that is backed by moral and economic rationales emanating from scholars. The theoretical justifications and the applicable copyright length are relevant to our discussion because they continue to determine the application of copyright rules on books in the digital era.

Chapter three on **"Book de-commodification: libraries as reserved places for legal distribution of books outside of markets"** brings the role of libraries to the forefront. Having discussed the market rationale for applying copyright to books, we will see how copyright exceptions, or privileges, allow libraries to become reserved spaces for the legal distribution of books outside of markets. The chapter explores the functional and the symbolic role of libraries from ancient to contemporary times. It describes libraries as publicly accessible spaces dedicated to public service. The most important message to glean from this chapter is that of the role played by libraries in book de-commodification. This connects with the normative benchmark that we have set to evaluate the applicable copyright rules: the encouragement of learning.

The fourth chapter poses the question **"What can libraries do with their books?"** It starts with an overview of copyright rules applicable to libraries. Hence, it looks at the current exceptions and limitations in international statutes and selected national legislation. As this thesis focuses on American and European case-law, the emphasis is placed on the corresponding jurisdictions. The chapter then proposes a flexible categorization of books, which is grounded on the effects of copyright on their availability. It explains the legal status of works in accordance with their copyright protection and looks at particular issues that arise with respect to books out of print, orphan works and the definition of the public domain. Two of these categories - the out of print and orphans - remain a contemporary conundrum despite digitization promises to solve their 'limbo' state. Our preoccupation lies on what libraries can do with such books in their collections. The chapter will review the relevant scholarship as well as various attempts to provide solutions provided by legislators, courts, policy-makers in collaboration with academics, as well as libraries themselves.

CHAPTER ONE - A MARKETPLACE OF IDEAS AND A MARKETPLACE FOR BOOKS

We will start with physical books and their treatment by the law in view of particular political and economic assumptions. This chapter is mainly historical and explains how copyright laws, vesting right-holders with certain privileges, aimed at creating a marketplace for books and at sustaining incentives for intellectual creation while putting an end to censorship. In spite of these initial intentions, the relationship between copyright laws and freedom of expression is idiosyncratic. Therefore, there is some kind of tension between creating a marketplace for books and simultaneously maintaining a marketplace of ideas. The first copyright legislation, the English Statute of Anne, was intended to create a marketplace for books and concurrently give specific rights to selected libraries. The position that libraries had enjoyed under the Statute of Anne was soon to be taken away when it concerned public libraries that were open to the poor. We will see how the first legislative act that recognized public libraries (again in England) has been significantly less generous. Yet, the Statute of Anne as emerged in its particular historical context is celebrated in almost every contemporary copyright textbook. I think it would be helpful to return to these times and place where the first discussions over copyright protection and over the rights or privileges of libraries happened, in order to understand the context within which this law emerged and spot the inconsistencies in contemporary copyright doctrine that rely on this history. We will do so in the order of five sub-chapters: (1.1) Comparing the selling of books to the selling of meat; (1.2) What is so special about books? (1.3) The Statute of Anne: "an act for the encouragement of learning"; (1.4) The forgotten original purpose of copyright; (1.5) Public libraries and 'the needs of the poor.'

1.1 COMPARING THE SELLING OF BOOKS TO THE SELLING OF MEAT

One of the oldest texts where we can find a normative justification for the exclusive rights that copyright affords books is in the *Lectures of Jurisprudence* by Adam Smith. [65] The

[65] Lectures On Jurisprudence, ed. R. L. Meek, D. D. Raphael and P.G. Stein, vol. V of the Glasgow Edition of the Works and Correspondence of Adam Smith (Indianapolis: Liberty Fund, 1982).

lectures were saved by one of his students at the University of Glasgow. To be accurate, Smith was referring to the exclusive rights afforded to booksellers. Smith discussed the concept of exclusive privileges, such as monopolies and privileges of corporations, as a category of 'real rights' in civil law - thus property rights as opposed to 'personal rights' where contract law applies. He did not support monopolies, quite the contrary. He argued that when the law assigns exclusive privileges and thus creates monopolies in the market, it runs the risk that the ones afforded such privileges will collude to raise prices. As an example, he mentioned the privilege of butchers to exclusively sell meat. He makes the argument in favor of competition and against cartelization of the market in a fashion similar to contemporary competition law: "When a number of butchers have the sole privilege of selling meat, they may agree to make the price what they please, and we must buy from them whether it be good or bad."[66] The harm caused by monopolies is ultimately against a public good: "But the great loss is to the public, to whom all things are rendered less comeatable, and all sorts of work worse done." Exclusive privileges to butchers to sell meat are, therefore, against the interests of the public that needs good quality meat and low meat prices.

Unlike meat, there are other goods that have an exclusive privilege, which according to Smith is *not so bad*. "The privilege of vending a new book or a new machine for fourteen years", says Smith, "has not so bad a tendency, it is a proper and adequate reward for merit."[67] Why is this the case? What makes the trading of books, and new machines, different to the trading of meat? The text of the lectures do not provide any further explanation but we can try to figure out the qualitative differences implied by Smith's comparison. Why can exclusive privileges be awarded *for merit* to vendors of books and not to butchers? By putting books and new machines together, Smith already gives us a clue as to why he differentiates them from meat. There is a creative process associated with the making of books and new machines, which is appreciated, but this is not the case when it involves the efforts of butchers when they cut and sell meat. There is possibly also a greater societal benefit to rewarding authors and inventors while rewarding butchers seems less so. Or maybe there are more complex economic reasons dictating such differential treatment; for example, butchers have adequate incentives to supply the market with meat but the same does not apply in the case of authorship or invention. To try to make the most

[66] Id.
[67] Id.

out of this weak argument found in Smith's lectures, let us focus on two more clues: first that the privilege is limited to fourteen years and, second, it is characterized as a proper and adequate reward for some sort of 'merit'.

It is worthwhile to pay closer attention to this text for several reasons. First of all, Adam Smith is cited as the first economist to touch upon the rationale for intellectual property protection, which is based on this brief passage. Interestingly, elaborate theories that justify intellectual property rights can be traced back to Smith's scant text. Secondly, his views on goods, which are now regulated by intellectual property, are important as they are viewed in the context of his overall political economy analysis that has been foundational for our (western) understanding of free markets. As will become apparent in the following analysis, even within our modern conceptions of free markets, books are not treated like meat. Arguably the reasons for that are not fundamentally different from those that Smith taught and wrote about.

Today books are goods that are traded in free markets. Copyright law ensures the marketability of these goods while assigning limited property rights to authors, publishers and their heirs. In the book market, regular laws of competition but also other laws, such as taxation, apply. Once we accept that copyright law helps ensure the creation of a market, the difference between books and meat is not substantial. Economists and copyright lawyers tend to focus on the so-called incentives and access tradeoff. Copyright is a monopoly that restricts access to intellectual creations but it is a *necessary evil*, or the price we need to pay to create this market in the first place.

Economists who are generally skeptical about monopolies have for a longtime rationalized this tradeoff as a necessary evil. [68] Copyright's anticompetitive effects are tolerated for a particular period of time and for specific reasons. As a statutory monopoly, copyright enables the right-holders of a book to charge a price above the marginal cost in order to cover the fixed costs of creation. Thus copyright laws ensure that there are enough monetary incentives for authors to take up the costs of creation in the first place. In this sense, being a statutory monopoly, copyright is understood as a system of public financing for intellectual production. In the words of Thomas Babington Macaulay at the House of Commons in England in 1841,

[68] For a classic analysis of the tradeoff and the difficult relationship between intellectual property and competition see for example Posner, R.A. and Landes, W.M., 2003. *The economic structure of intellectual property law*. Harvard University Press.
See also, the rich literature review of Towse, R., Handke, C. and Stepan, P., 2008. The economics of copyright law: A stocktake of the literature. *Review of Economic Research on Copyright Issues*, 5(1), pp. 1-22, at p.3.

copyright is "a tax on readers for the purpose of giving a bounty to writers."[69] In his view, copyright is a necessary tax on readers but should be carefully kept at such length and price levels that reward writers sufficiently.

It is important to note that this economic analysis places a heavy emphasis on the incentives of the tradeoff, rather than on the access. Most contemporary copyright theories, rooted in either economic or moral reasoning, I think, overemphasize the role of the law in maintaining the incentives. The theories are roughly speaking grouped into two categories that correspond to two schools of thought: the utilitarian school attributed to Anglo-Saxon world and a moral theories school attributed to continental Europe.[70] At the same time, the value generated by broad access to books is comparatively underdeveloped, at least in the existing copyright scholarship. This qualitative value of books is predominantly related to freedom of expression, and its mirroring freedom of information, which is important to keep in mind when examining how books are treated by the legal system. In other words, examining only copyright law or competition law, which are sets of rules designed for markets, cannot adequately explain why books are not treated like meat or other market products.

1.2 WHAT IS SO SPECIAL ABOUT BOOKS?

If market rationales do not sufficiently explain the different place of books in the eyes of legislators, perhaps other justifications can shed light on the question of what makes books so special. One such justification is the free flow of ideas as celebrated by modern democracies and other societies. Books are the primary medium to carry ideas, which find expression in writing. One of the most powerful rationales for a high-level protection of books can be found in the notion of the marketplace of ideas, particularly as it has developed in American case-law. Historically, the free flow of books in a marketplace of ideas is valuable for democracies because words are powerful and when they compete, valuable truths arise. Borrowing from the arguments of the Supreme Court Justice Oliver Holmes in his famous dissent in *Abrams v. United States*,

[69] See Volume 8 of The works of Lord Macaulay complete, Trevelyan, Hannah More Macualay, published in 1875 by Bombay, Longmans, Green, and co.

[70] We will discuss the justifications for copyright protection with a focus on books (literary production) in chapter two of this thesis. Brad Sherman and Lionel Bently have offered an exemplary historical account of the debates leading to the assignment of property to mental labor: Sherman, B. and Bently, L., 2003. The Making of Modern Intellectual Property Law: the British Experience, 1760 - 1911. Cambridge University Press. (see pp. 11-35).

rights that guarantee freedom of expression are based on a trust that in a world of "fighting faiths" competing for the truth - in other words a pluralistic world. Allowing ideas to compete in a free market is a good thing; or as Holmes put it, it is our only hope to reach some kind of imperfect truth.[71] The same rationale applies to written ideas as much as it does to ideas that are orally transmitted; what is more, ideas expressed in writing have the advantage that when preserved, they remain part of a cultural heritage. To a certain extent, copyright rules have been designed in a way that the free flow of ideas is not suppressed. The exclusivity it affords to right-holders is based on the expression and not on the ideas. Nonetheless, as long as access to the book is restricted or difficult due to copyright rules, it actually impacts the free flow of ideas.

The relationship between the two bodies of law, freedom of expression and copyright, is complex. While freedom of expression ensures a marketplace of ideas, copyright ensures a marketplace of books, where expressed ideas are protected with exclusive privileges. At first sight the marketplace of books and the marketplace of ideas are complementary concepts. There is an underlying paradox, however, which makes their relationship far more complicated. The primary focus in the marketplace of books is on production and, therefore, on the rights of authors and publishers. In the marketplace of ideas, however, the audience is as important as the speaker. The rights of readers, the audience in the marketplace of ideas, have arguably been neglected by copyright rules and theories.[72] As a legal monopoly afforded to authors and publishers, copyright restricts the dissemination of speech to a certain extent. Arguably it is less severe if the restriction is regulated by market rules that are not imposed by governments or other

[71] *Abrams v. United States*, 250 U.S. 616 (1919). See Holmes' dissent at pp.630-1:

> Persecution for the expression of opinions seems to me perfectly logical. If you have no doubt of your premises or your power, and want a certain result with all your heart, you naturally express your wishes in law, and sweep away all opposition. To allow opposition by speech seems to indicate that you think the speech impotent, as when a man says that he has squared the circle, or that you do not care wholeheartedly for the result, or that you doubt either your power or your premises. But when men have realized that time has upset many fighting faiths, they may come to believe even more than they believe the very foundations of their own conduct that the ultimate good desired is better reached by free trade in ideas -- that the best test of truth is the power of the thought to get itself accepted in the competition of the market, and that truth is the only ground upon which their wishes safely can be carried out. That, at any rate, is the theory of our Constitution. It is an experiment, as all life is an experiment. Every year, if not every day, we have to wager our salvation upon some prophecy based upon imperfect knowledge.

[72] See mostly Litman. 2011. J., Readers' Copyright. *Journal of the Copyright Society of the USA*, Vol. 58, p. 325. Readers are more frequently discussed today as part of a broader users community in the digital space: See most recently, Elkin-Koren N., Copyright in a Digital Ecosystem: A User Rights Approach, in Okediji, R.L. ed., 2017. *Copyright law in an age of limitations and exceptions*. Cambridge University Press.

sources of power, which in this situation would amount to censorship. But is this really the case? Below we will look at the history of the fist copyright legislation enacted in England in 1710. The Statute of Anne was introduced at a time when intellectuals of the time were arguing over the perils of censorship.

1.3 THE STATUTE OF ANNE: "AN ACT FOR THE ENCOURAGEMENT OF LEARNING"

Looking back at the history of the Statute of Anne, the first copyright act, and the intellectual battles against censorship taking place around the time that the Statute was enacted, we will see how copyright protection and freedom of expression have had a long and complex relationship. Even though the enactment of the first copyright law is celebrated as a triumph against censorship, a careful look at the history indicates that the line between censorship by government rules and censorship by market rules is thin. The history of the Statute of Anne is connected to a longer history over the battles against censorship. Before the Statute was passed, authorized printers from the Stationer's Company had the royal approval to print books. The so-called Licensing Act, or Printing Act, vested the Stationer's Company with a monopoly over publishing, which despite earlier opposition, was renewed at least for a period of time.[73] Moreover, the church was vested with powers to prohibit the printing of books that were contrary to the faith.[74] John Milton and John Locke were among those who loudly protested against the censoring practices that resulted from the Stationer's Company monopoly. Their respective texts are worth reading to understand the context in which the first English copyright laws were drafted to replace the existing monopoly.[75]

Printed in 1644, John Milton's Aeropagitica is a prose written in a form of a speech, which was directed at the Parliament of England and entitled "For the Liberty of unlicenc'd

[73] Hughes, J., 2009. Locke's 1694 Memorandum (and More Incomplete Copyright Historiographies). *Cardozo Arts & Ent. LJ*, *27*, p.555

[74] Ibid.

[75] On this topic see also: Feather, J., 1980. The book trade in politics: The making of the copyright act of 1710. *Publishing history*, *8*, p.19; Elkin-Koren, N., 1996. Cyberlaw and social change: A democratic approach to copyright law in cyberspace. *Cardozo Arts & Ent. LJ*, *14*, p.215; Rose, M., 2009. The Public Sphere and the Emergence of Copyright: Areopagitica, the Stationers' Company, and the Statute of Anne. *Tul. J. Tech. & Intell. Prop.*, *12*, p.123 (also in Deazley, R., Kretschmer, M. and Bently, L., 2010. *Privilege and property: essays on the history of copyright*. Open Book Publishers.)

Printing." Milton was expressly writing against the Licensing Order of 1644.[76] In his prose, Milton writes about the wisdom found in books and the evils of censorship. He talks about ancient Athens and Rome and how the absence of censoring practices left us a legacy of great works. In his text we can see how licensing was clearly associated with Catholicism and the censoring of texts was considered blasphemous. Milton makes a strong case against licensing, essentially defending free dissemination of truths and ideas, arguing:

> I cannot set so light by all the invention, the art, the wit, the grave and solid judgement which is in England, as that it can be comprehended in any twenty capacities how good soever, much lesse that it should not passe except their superintendence be over it, except it be sifted and strain'd with their strainers, that it should be uncurrant without their manuall stamp. Truth and understanding are not such wares as to be monopoliz'd and traded in by tickets and statutes, and standards.[77]

Much more than Milton's parliamentary speech, the views of John Locke against the same practices are cited as highly influential to the debates of the time. Locke's 1693 letter to Edward Clarke presents some of the philosopher's views. It was followed by a memorandum written around 1694 and submitted to the British Parliament. This memorandum, apparently shared with Parliamentary members, is considered to have influenced and helped end the monopoly of the Stationer's Company and contributed to the drafting of the Statute of Anne.[78] His memorandum reads as a commentary to what appears to be a version of the text of the Licensing (or Printing) Act, which vested the Stationer's Company with a printing monopoly. His comments portray his irritation with the state of the book market that resulted from the

---

[76] *June 1643: An Ordinance for the Regulating of Printing*, in Acts and Ordinances of the Interregnum, 1642-1660, ed. C H Firth and R S Rait (London, 1911), pp. 184-186. British History Online http://www.british-history.ac.uk/no-series/acts-ordinances-interregnum/pp184-186 [accessed 4 May 2017].
See in particular the following excerpt: *It is therefore Ordered by the Lords and Commons in Parliament, That no order or Declaration of both, or either House of Parliament shall be printed by any, but by order of one or both the said Houses: Nor other Book, Pamphlet, paper, nor part of any such Book, Pamphlet, or paper shall from henceforth be printed, bound, stitched or put to sale by any person or persons whatsoever, unless the same be first approved of and licensed under the hands of such person or persons as both, or either of the said Houses shall appoint for the licensing of the same, and entred in the Register Book of the Company of Stationers, according to ancient custom, and the Printer thereof to put his name thereto* […]
[77] John Milton, Areopagitica (1644) (Jebb ed.) [1644].
[78] See Hughes, supra note 73.

monopoly and, most notably, he compares the English market to the Dutch market at the time. Locke maintains:

> […] our printing is so very bad, and yet so very dear in England: they who are hereby privileged to the exclusion of others, working and setting the price as they please, whereby any advantage that might be made to the realm by this manufacture is wholly lost to England, and thrown into the hands of our neighbours; the sole manufacture of printing bringing into the Low Countries great sums every year. But our Ecclesiastical laws seldom favour trade, and he that reads this act with attention will find it upse [highly] ecclesiastical. The nation loses by this act, for our books are so dear, and ill printed, that they have very little vent among foreigners, unless now and then by truck for theirs, which yet shows how much those who buy the books printed here are imposed on, since a book printed at London may be bought cheaper at Amsterdam than in Paul's Church-yard, notwithstanding all the charge and hazard of transportation: for their printing being free and unrestrained, they sell their books at so much a cheaper rate than our booksellers do ours, that in truck, valuing ours proportionably to their own, or their own equally to ours, which is the same thing, they can afford books received from London upon such exchanges cheaper in Holland than our stationers sell them in England. *By this act England loses in general, scholars in particular are ground, and nobody gets, but a lazy, ignorant Company of Stationers, to say no worse of them* […][79]

In this passage, Locke hints at an understanding of the book market, which has strong normative undertones. One such normative position put forward by Locke is that the monopoly afforded to the Stationer's Company resulted in high prices. For Locke, this is obviously bad. Secondly, he attributes this market failure to the fact that the licensing of copies is not free. He points out that in the Netherlands, where the licensing of book copies is not restricted, there is a better market where even English books are cheaper. We can see how the arguments that Locke

[79] John Locke, Letter from John Locke to Edward Clarke, 2nd January 1692, in Benjamin Rand, ed. *The correspondence of John Locke and Edward Clarke* 366, (Harvard University Press, 1927), at 381 [emphasis added].

gives are based on primarily economic factors. Just like Adam Smith's views against monopolies that raise meat prices, Locke maintains that the same can happen to books. The economic arguments however are ultimately met with a broader rationale with regard to cheap availability of books and why is it vital for society. England loses in general, he argues. The country loses its scholars while those who win are a bunch of lazy and ignorant traders. Notably the loss of scholars is the comment that needs to further attention. There is a recognizable value implied here in the maintenance of an environment for fair and easy access to materials for those who wish to study them and further advance learning.

The text of the Statute of Anne, as adopted in 1710, seems to follow Locke's rationale. Its very title sends out a clear message with respect to its purpose: "An act for the encouragement of learning, by vesting the copies of printed books in the authors or purchasers of such copies, during the times therein mentioned." The first significant change is the prominent position of the author, who is vested with a limited right to control the printing and circulation of copies of his/her books. Paragraph II of the Statute specifies that the right lasts for fourteen years for books that were in not print at the time, and twenty-one years for books already in print. The remainder of paragraph II and a few additional paragraphs of the Statute provide further detail with regard to the purchasing of copies from printers and other issues related to the book market, such as foreign books. Paragraph V of the Statute is of utmost importance. It stipulates an obligation to deposit nine copies of each book to the Stationer's Company, "upon the best paper", to be available to the following list of libraries: "the royal library, the libraries of the universities of Oxford and Cambridge, the libraries of the four universities in Scotland, the library of Sion College in London, and the library commonly called the library belonging to the faculty of advocates at Edinburgh respectively." If not forthcoming, these libraries are entitled to demand their respective copy. The Statute attaches penalties to 'proprietors' who do not obey the requirement. In addition to a monetary sum for every copy that is not delivered, the proprietor who fails to deposit the book has to pay for the value of each copy.

The above provision about the copies of new books reserved for the libraries is, I contend, as important as the reform of the printing market, which was the original purpose behind the Statute. The Statute firstly achieves the transformation from a restrictive system of licensing and vesting a particular guild with printing monopolies to a system of limited tradable rights to the authors of books and thus helps foster a free market for books. As Locke argued,

this was a way to lower the prices of books but also end the debilitating censoring which was a result of the previous system. While transforming or rather liberating the market the Statute stays true to its primary purpose: the encouragement of learning. Not only were books to be traded freely and cheaper but also their availability in a number of important libraries around the country was guaranteed. All but the royal library are academic institutions. When speaking of the Statute of Anne as the first of copyright laws which influenced the spread of copyright laws in many more jurisdictions, [80] we must not forget the Statute's explicit rationale, or its revolutionary provisions: not only authors but also certain libraries were vested for the first time with special *rights* over copies of books.

1.4 THE FORGOTTEN ORIGINAL PURPOSE OF COPYRIGHT

Copyright evolved into a body of rules that promoted the economic rights of certain parties while it simultaneously moved away from its 'original' purpose to encourage learning. This is not to say that we must adhere to originalism. It is only necessary to rethink whether there are good reasons for maintaining the benchmark which the Statute of Anne expressly protected and which the Statute's advocates, like John Milton and John Locke, were also endorsing. In any case, we must first acknowledge the actual alienation from its original benchmark and assess the consequences that contemporary copyright laws have on books, their production, accessibility and use. Another observation is that the role of libraries, as clearly expressed in the Statute of Anne, has since lost its significance. Again, library rights, which the Statute affords, are aligned with the stated purpose. If libraries today enjoy lesser privileges, it is also the case that we have departed from the advancement of learning rationale.

Mark Rose explains how modern analyses of copyright readily draw on the Statute of Anne but tend to neglect how its broader rationale was based on a different inception. He explores the nature of the transformation that it aimed to bring about and observes how economic

[80] The American Copyright Act, for example, first introduced in 1790, was modeled after the Statute of Anne and is similarly titled "An Act for the Encouragement of Learning, by Securing the Copies of Maps, Charts, and Books to the Authors and Proprietors of Such Copies, During the Times Therein Mentioned.", Copyright Act of 1790 online available at https://copyright.gov/about/1790-copyright-act.html.
On the influence that British copyright law has had to more jurisdictions see also Crews K., WIPO Study on Copyright Limitations and Exceptions for Libraries and Archives, SCCR/17/2 (August 26, 2008), online available at http://www.wipo.int/edocs/mdocs/copyright/en/sccr_17/sccr_17_2.pdf, at p. 25 et. seq.

factors were only the means to a non-monetary cause: the emergence of a public sphere.[81] Those who protested against the old censoring regime wanted to see the authors at the center, rather than the monopolist printers.[82] As seen already, a public sphere is what Locke basically advocated for: a space where books can be available (cheaply) and where learning and scholarship is encouraged. In general, it was understood that the interests worth protecting are not those of the printers but of the authors, which serve to advance learning and result in benefits associated with the dissemination of knowledge. To quote Rose's analysis of Milton:

> Milton was concerned with liberty and the advancement of knowledge; the company was concerned with propriety and the maintenance of order. As the controversialist Henry Parker put it in The Humble Remonstrance of the Company of Stationers, published in 1643 as part of the campaign for the reinstitution of licensing, the issue as the company saw it was *not merely the advancement of knowledge but "the advancement of wholesome knowledge."*[83]

The advancement of knowledge and the advancement of "wholesome knowledge" are two different goals. Rose notes that while Milton was concerned with the advancement of knowledge as a result of a liberty afforded to authors, the Stationer's Company "was concerned with propriety and the maintenance of order."[84] The point in emphasizing the difference is that this analysis of Milton and Locke and, subsequently, the Statute of Anne reveals a particular benchmark against which we can evaluate the introduction and evolution of copyright laws. To repeat the wording of the Statute, the benchmark was "the encouragement of learning". The development of copyright laws particularly those applied to books has, in my judgment, alienated itself from this benchmark. This is a crucial observation when it comes to libraries, and their operation, within the copyright framework. The encouragement or advancement of learning is a normative goal. It arguably begs for regulation that supports institutions dedicated to that goal and prioritizes their needs. If we are to sustain this normative benchmark, libraries should

[81] Rose supra note 75.
Rose cites more scholars who have invoked the Habermasian notion of a public sphere with respect to copyright. See his first citation at p. 123.
[82] Ibid.
[83] Ibid, p. 133 citing Parker H., *The humble remonstrance of the company of stationers to the high court of parliament* (1643).
[84] Id. (p. 133).

always enjoy privileged treatment, since they are traditionally the institutions dedicated to the preserving and providing access to learning materials. Moreover, their immediate link to educational institutions, schools and academia, needs to place them at the center of the legislator's attention insofar as copyright norms apply to them.

The next point explores the role of libraries. As discussed above, according to the Statute of Anne, the role of the library is central. Libraries are granted a specific right to a copy of every new book that is printed not because it amounts to an exception to the rights of other stakeholders, like authors or publishers, but it is a right that they are entitled to in their own right. Moreover, as we saw there were penalties attached if the libraries' right to a copy was not respected; penalties that were enforced even against authors. Let us bear in mind the same Statute elevated the role of the authors to that of just copyright-holders. The new copyright legislation is, in a sense, the product of renegotiation of the relationship between the following parties: the State, the printers/booksellers, the authors and the reading public.[85] They all had their own interests in books even before the enactment of the Statute; however, the balance of interests is substantially different under the Statute of Anne and the preceding Licensing Act. The previous regime, in contrast, was a product of a two-way bargain, between the State and the printers/booksellers.[86] With the new copyright law the authors' interests got explicit recognition together with those of the reading public. Arguably the role of the State had changed radically, moving from the forefront to the background. With the passing of the Statute, the State willingly positioned itself in the background where it supervised a free market for books and the collection and preservation of books from libraries. The reading public benefited from both the market and the non-market arrangements of the new Statute: lower prices as a result of the book market and a cultural record as a result of the libraries' keeping a copy of each book printed.

If we look at the position of libraries in contemporary copyright legislation, we notice a difference between the rhetoric expressed in the Statute of Anne and the rhetoric in modern copyright. Libraries' expressed position is hardly central any longer. From concrete rights over copies of books, libraries now only enjoy special treatment in the form of copyright exceptions. Unlike in the old Statute, libraries do not have any explicit right to copies of books; however, they are in certain circumstances exempt from copyright infringement when they make copies. In

[85] Deazley, R., 2004. *On the origin of the right to copy: Charting the movement of copyright law in eighteenth century Britain (1695-1775)*. Bloomsbury Publishing.
[86] Rose, supra note 75, p. 139.

certain jurisdictions, there is still an obligation to deliver a copy of every book printed. According to the latest study on library exceptions published by the World Intellectual Property Organization, of 188 States members, 156 have at least one statutory library exception in their copyright legislation.[87] A general exception for copying is only found in 31 countries and the remainder have combinations of specific exceptions, specifying the acts where copying is permitted and by which actors: copies made by librarians mainly for preservation or replacement but also other reasons, such as interlibrary loans, and copies made by library users mostly for purposes of research and study. A number of conditions, including the number of copies allowed, are usually attached. To use the contemporary terminology in the UK legislation as an example, we witness the rhetorical shift: from library rights to exceptions available to public libraries. Under Articles 40A to 44B of the Copyright, Designs and Patents Act of 1988, there is a list of acts that a public library can engage in and not infringe copyright.[88]

How did the above shift in the rhetoric take place? Let's begin with the fact that at the time the Statute of Anne was enacted the listed libraries that benefited from the law were not publically accessible. The right to establish public libraries was granted to municipal boroughs in the UK over a century later when the Public Libraries Act of 1850 was enacted. This was preceded by a number of legislative reforms, including the Poor Law Amendment Act of 1834, the so-called *New Poor Law*.[89] It is impossible to understand the reason why the legislative rhetoric changed *vis-à-vis* libraries at a time when libraries would become more publically accessible unless we link this development to the social context: the division between elites granted library access under the Statute of Anne in 1710 and the rest of society, the poor.[90] As mentioned elsewhere, in 1710, when the Statute of Anne passed, it explicitly mentioned nine

[87] Crews K., WIPO Study on Copyright Limitations and Exceptions for Libraries and Archives: Updated and Revised, SCCR/30/3 (June 10, 2015), online available at http://www.wipo.int/edocs/mdocs/copyright/en/sccr_30/sccr_30_3.pdf.

[88] We will discuss further what modern libraries can and cannot do, under contemporary copyright rules in chapter four.

[89] On the Poor Law Reform see the documents online available at http://www.parliament.uk/about/living-heritage/transformingsociety/livinglearning/19thcentury/overview/poorlaw/.

[90] Black, A., 2000. Skeleton in the cupboard: social class and the public library in Britain through 150 years. *Library History*, *16*(1), pp.3-12.; McCook, K.D.L.P., 2001. Poverty, democracy and public libraries. *Libraries &Democracy: The Cornerstones of Liberty*, pp.28-46. For more contemporary discussions focused on American public libraries see also: Holt, L.E. and Holt, G.E., 2010. *Public library services for the poor: Doing all we can*. American Library Association; Pawley, C., 2010. *Reading places: Literacy, democracy, and the public library in Cold War America*. Univ. of Massachusetts Press.

libraries as the beneficiaries.[91] None of these libraries could be accessed by the poor. It would only be much later in history when public libraries were thought of as a medium to remedy social exclusion.[92] It is worth noting that raising taxes for public libraries was controversial and the financial support of libraries predominantly came in the form of charity. When the beneficiaries of the law were no longer libraries reserved for an exclusive social class but became open and public, we see the language in the law also changed and (public) libraries do not enjoy rights but rather they are granted some exceptions. Is it possible to argue that the change in rhetoric also caused a change in the purpose of copyright? Is the underlying rationale no longer the advancement of learning, as enshrined in the 1710 Statute, but now concerns the advancement and maintenance of a market? It seems that between 1710 and the reforms envisioned by Milton and Locke and today's statutory copyright exceptions there has been a gradual erosion of the purpose of the law.

1.5 PUBLIC LIBRARIES AND 'THE NEEDS OF THE POOR'

Several historians studied the first piece of legislation for libraries passed by the English parliament: the Public Libraries Act of 1850. The act allowed for the establishment of public libraries to be accessible to all citizens. Most importantly, some have looked into the heated parliamentary debates during the legislative stages of the bill and how its final text was a product of much compromise and substantial concessions.[93] The studies, for example, of Stanley Max and of W.J. Murison demonstrate how the proponents of the Public Libraries Act were strongly motivated by efforts to counter the effects of poverty. The reformative role of libraries was perhaps more important to them than the educational role.[94] The tranquilizing effects of books could presumably bring order to the working class, particularly the restless tendencies of the working class in cities. The appointed parliamentary committee, the Select Committee on Public

[91] See *infra* at 1.3.
[92] See D'Angelo, E., 2006. *Barbarians at the gates of the public library: how postmodern consumer capitalism threatens democracy, civil education and the public good*. Library Juice Press, LLC.
[93] Max, S.M., 1984. Tory Reaction to the Public Libraries Bill, 1850. *The Journal of Library History (1974-1987)*, *19*(4), pp.504-524 and Murison, W.J., 1988. *The public library: its origins, purpose and significance*. London: C. Bingley.
[94] Ibid, Max, p.504.

Libraries, proposed the bill as one aimed at promoting "the temperate and moral habits of the working classes."[95]

Max emphasizes the effectiveness of the opposition to the bill, which was first proposed by William Ewart.[96] The opposing conservatives did not succeed in killing the bill but managed to narrow its scope significantly. As Max explains, the passing of the bill was ground-breaking not so much because of its scope – since it was very limited – but because it was the first time to recognize the necessity of public intervention for popularized access to knowledge. Prior to this, libraries were private and for the poor to gain access to books and educational material, they relied on private voluntary initiatives and charity. Opponents to the bill argued that establishing reading or lecture rooms for the poor "might give rise to an unhealthy agitation."[97] As Max describes the debates, William Ewart, who introduced the bill, together with his supporters had to argue against a general reluctance to spend taxes to benefit the poor. The debates over the spending of taxes to create public libraries also initiated more progressive arguments with respect to the effects of libraries. Bearing in mind that an accepted and very explicit class division existed in the 17th century English society, parliamentarians, like John Bright, argued that the "diffusion of intelligence" would not only bring order but also "open discussion, among all classes." [98] Bright, as we have seen, spoke of *taxes on knowledge*.

Overall, the bill that William Ewart was struggling to pass was already very limited in scope. It "would permit town councils to adopt the prospective act and levy a rate of no more than a half penny in the pound (that is, a ratio of 1:240) for library support".[99] It did not even provide for the procurement to purchase books; the libraries would have to hope for book donations. Therefore, while books were rightfully reserved by law for the advancement of learning for individuals with access to the royal library and certain university libraries, public libraries were not meant to enjoy a similar status. Books available in public libraries would hopefully and at best have a tranquilizing effect on the agitated poor.

The reason why we engaged in this historical discussion has been to highlight a paradox. The official rhetoric supported libraries as institutions dedicated to learning while at the same

[95] Report from the Select Committee on Public Libraries; together with the Proceedings of the Committee, Minutes of Evidence and Appendix, ordered by the House of Commons, to be printed 23, July 1849. Record available online at the HathiTrust catalogue: https://catalog.hathitrust.org/Record/011254579.
[96] Max, supra note 93, p. 506 et. seq.
[97] Ibid, p. 505.
[98] Ibid, p. 512.
[99] Ibid, p. 508.

time it discriminated against communities that belong to a certain 'public' and therefore marginalized their access to the institution. The historical debates, of course, reflect societies where social divisions were much more explicit. To this day, however, we have somehow retained the paradox in our copyright laws: while access to information is recognized as an important right to all citizens in democratic societies, access to material is not guaranteed.[100] Nominally, the copyright doctrine does not allow for information to be proprietized; it is only the expression that gets 'locked' into property rules for a certain period of time. It is, however, difficult to completely separate access to information and access to material (expressions that is copyrightable). There are many instances when the application of intellectual property rules has arisen exactly where these lines blur.[101] In fact copyright, especially if viewed from its protracted and at times unnecessary protection, contributes to keeping works from the public's reach. It is, therefore, a problem that ultimately affects freedom of information when public libraries are essentially not supported in providing access and copyrighted materials to those who cannot afford market access. Publically accessible libraries have the potential to fully contribute to more equal access to knowledge. Policies and laws that do not support public libraries in this role are practically endorsing, or at least tacitly accepting, unequal access to knowledge for many disadvantaged groups in the society.

[100] Besides copyright and its scope, there are different ways in which regulation tries to achieve freedom of information. Further forms of regulation that are relevant in this respect are subsidized prices for books and, also, public service broadcasting. The rationale behind these forms of regulation is complementary to that of sustaining publically accessible libraries. The analysis of these forms of regulation is beyond the scope of this thesis – it remains, however, relevant to keep the bigger picture in mind.

[101] See Brennan, T.J., 1992. Copyright, property, and the right to deny. *Chi. Kent L. Rev.*, *68*, p.675 and Samuelson, P., 1988. Information as property: Do Ruckelshaus and Carpenter signal a changing direction in intellectual property law, *Cath. U. L Rev.*, *38*, p.365.

CHAPTER TWO - JUSTIFICATIONS FOR COPYRIGHT PROTECTION ON BOOKS

This chapter reviews theoretical justifications for applying copyright protection on books. It examines both moral and economic justifications for copyright and focuses on the relevant discourse. The economics of books is a challenging topic that has been addressed by scholars in several different ways. Lawyers together with economists have been trying to justify or define the optimal copyright framework applicable to books, employing justifications grounded on Locke's moral theory of property, a romantic understanding of authorship, or utilitarianism. First, we will look at moral theories that justify the vesting of property rights to intellectual creations. The property rights vocabulary that remains tied to copyright, as part of intellectual property, finds its origins in Locke's moral theory of property. This does not necessarily mean that Locke had much to say about intellectual property but it has become commonplace to interpret him in this way. Secondly, we will look at political economy approaches to copyright on books as they evolved into modern utilitarian views. Utilitarianism seems to have prevailed on both sides of the Atlantic, but it is also prevalent on the international level judging from the current international copyright laws. Both property theory and the utilitarian rationales, however, fit the contemporary understanding of copyright. It is commonly said that the American model is more utilitarian while the continental European is predominantly based on moral considerations. This view is rooted in a notion of the European protection of moral rights and an older deference to the so-called 'romantic author.' Thirdly, we will examine the debate over the optimal copyright length. The debate is usually framed in utilitarian terms since it rests on the effort to define the optimal balance between copyright protection to sustain incentives and access. This balance is of utmost relevance to the thesis if we are to look at libraries as institutions that provide access outside the market that copyright protects. This is, however, the link that the following chapter three will make.

2.1 THE PROPERTY DISCOURSE AROUND COPYRIGHT AND THE ROMANTIC IMAGE OF THE AUTHOR

The vocabulary of property rights has and continues to be employed to refer to the rights over the content of books. With copyright amounting to a form of intellectual property, this almost appears to be straightforward. In economic and legal terms, however, the nature of property rights, applied to the content of intellectual works, is far from straightforward.

The property vocabulary attaches to moral justifications for property that goes back to John Locke. There is some scholarship examining Locke's views on intellectual property and whether it falls under his famous property theory.[102] Locke's property theory awards men for the results of their labor, which they own like they own their bodies: "Though the earth, and all inferior creatures, be common to all men, yet every man has a property in his own person […] The labour of his body, and the work of his hands, we may say, are properly his." [103] Not only textually but also historically we can be sure that his theory was meant to apply to tangible goods, such as apples picked from a tree or wood collected from the forest.[104] It is far from clear whether Locke also understood intellectual works to be the subject of property in the same way. Seana Valentine Shiffrin, for example, rejects that the Lockean theory would be able to support "strong, natural rights over most intellectual products."[105] She stresses, first, the importance of the status of common ownership as a starting point before the appropriation of goods by man, and second, how Lockean property theory disapproves of wasteful appropriation contrary to the benefit of humankind.

If we were to apply Locke's property theory on books, as the products of the intellectual labor of man, then authors should have a natural property right on their books, which are products of their intellectual labor. The theory, or at least the way that it was first expressed by Locke, does not translate entirely well, especially if we consider follow-on creation. Do authors remove thoughts out of a 'state of nature' which is essentially the public domain? Or do they actually use the works of others, get inspired by previous authors or as the expression goes 'stand on the shoulders of giants'? Before we accept a Lockean justification of property rights on

[102] See in particular Sherman and Bently, supra note 70, p 21. *et seq.*
[103] Locke, J. [1689] *The Two Treatises of Civil Government* (Hollis ed.), paragraph 27.
[104] Ibid, para. 28.
[105] Shiffrin, S.V., 2001. Lockean arguments for private intellectual property. *New essays in the legal and political theory of property*, *1*, pp.138-167, at p. 141.

intellectual works, a debate on whether there is any such thing as parthenogenesis is almost unavoidable.[106]

Scholars have long debated whether intellectual property is property.[107] A legitimate question is why does it matter.[108] After all, property is arguably only a legal construct or a "conversational habit."[109] Perhaps it carries some weight if copyright "goes too far in turning information into property."[110] Then, the question that actually matters is not whether property is the appropriate legal framework or term for intellectual property, but whether its limitations are sufficient not to restrict information flow. Timothy Brennan, for example, in his article *Copyright Property and the Right to Deny*, argues that copyright limitations, including fair use, and the notion of copyrightability are consistent with the copyrighted work being thought of as property and serve as guarantees so that information is not 'locked'.[111] This conclusion is contestable, especially when we look at practice.[112] A major critique to the effects that the proprietization of intellectual property has had is made by James Boyle in his work, *The Second Enclosure Movement and the Construction of the Public Domain*, where he describes a phenomenon of fencing off what was initially free information and intellectual commons.[113]

The property metaphor has actually extended not only to copyright but also to the public domain. Next to the understanding of property rights on books, it is necessary to grasp how and for what reasons the public domain has evoked real property metaphors. According to Pamela Samuelson, the first use of the term public domain referred to unsettled lands in the West of the

---

[106] Let us briefly note that Locke's property theory was developed as a response to Robert Filmer's defense of the divine right of the king. It is generally important to remember how the historical context in which Locke wrote is not easily transferable to different contexts. About Locke's 1694 Memorandum opposing to the renewal of the licensing act see Hughes, supra note 73, p.555.

[107] Child, J.W., 1990. The moral foundations of intangible property. *The Monist*, *73*(4), pp.578-600; Easterbrook, F.H., 1990. Intellectual property is still property. *Harv. J. L. & Pub. Pol'y*, *13*, p.108; Becker, L.C., 1992. Deserving to own intellectual property. *Chi.-Kent L. Rev.*, *68*, p.609. A number of intellectual property scholars have started to invest further in the study and critisism of the property metaphor (and analogies) and its implications. Among many see for example: Peter S. Menell, Intellectual Property and the Property Rights Movement, 30 Regulation 36 (2007); Cohen, J.E., 2011. Copyright as Property in the Post-Industrial Economy: A Research Agenda. *Wis. L. Rev.*, p.141; Van Houweling, M.S., 2013. Land Recording and Copyright Reform. *Berkeley Technology Law Journal*, *28*(3), pp.1497-1510.

[108] Carter, S.L., 1992. Does it Matter Whether Intellectual Property is Property. *Chi.-Kent L. Rev.*, *68*, p.715.

[109] Jeremy Waldron, *The right to private property*, 27-30 (1988).

[110] Brennan, supra note 101.

[111] Id.

[112] See Samuelson. Information as Property: Do Ruckelshaus and Carpenter signal a changing direction in intellectual property law, supra note 101.

[113] Boyle, J., 2003. The second enclosure movement and the construction of the public domain. *Law and contemporary problems*, *66*(1/2), pp.33-74.

United States, which before the 20th century public domain was thought of as common property or public property information resources.[114] Samuelson provides a very interesting account of the advantages and disadvantages of the terms: she explains how the terms 'common property' and 'public property' suggest that the public has ownership rights in the information resources, which arguably preclude privatization. Whereas the public domain is "carrying over the not-yet-privatized metaphor" arguably making the public domain "presumptively privatizable by those who appropriate and invest in commercialization of public domain resources".[115] The question whether to place the public domain under the terminology and ultimately under the umbrella of property, or not, is a critical but also very difficult one. Samuelson also considers those who reject the very idea of 'mapping' the public domain because the term can possibly reinforce the real property metaphor that already exists for information goods. In other definitions, however, property is used to negatively define the public domain. According to Paul Heald, for instance, the public domain can be characterized as the repository of works and creative goods that are not subject to the costs and constraints of ownership, thus the costs of property law.[116]

Eli Salzberger argues that labeling the debate between copyrighted works and those in the public domain as for or against proprietization dispute is inaccurate.[117] This, he explains, is primarily the case because it is not even clear "whether creating intellectual property rights by law is a manifestation of the free market or a case of government intervention."[118] He adds that for intellectual property, traditional law and economics approaches essentially advocate central intervention by granting intellectual property rights. A good example of this approach and stance is the classic analysis by Landes and Posner, which we will look at below.[119] It is more useful, however, to realize how the public domain can be seen as a market. In this situation, moral theory must extend to also cover the reasons why property assigned to individuals for their labor transforms into common property when the copyright term expires.

---

[114] Samuelson P., Challenges in Mapping the Public domain, in Hugenholtz, P.B. and Guibault L. (eds), 2006. *The future of the public domain: identifying the commons in information law*. Kluwer Law International, pp.7-21. In the same volume see also Cohen J., Copyright, Commodification and Culture: Locating the Public Domain (pp. 121-166).

[115] Id. (Samuelson), p. 20.

[116] Heald, P. 2013. The Public Domain. Illinois Public Law Research Paper No. 14-21. available at SSRN: https://ssrn.com/abstract=2362983.

[117] Salzberger, E. 2006. Economic Analysis of the Public Domain, in Hugenholtz, P.B. and Guibault L. (eds), *The future of the public domain: identifying the commons in information law*. Kluwer Law International, p. 27-57.

[118] Ibid, p. 28.

[119] Ibid, p. 30. See also Posner and Landes, supra note 68.

Most relevant to us is how moral theories of copyright, based on property terms, are translated with a particular romanticism when it comes to books. The image portrayed is not so much that of a man harvesting the earth - the Lockean illustration of property in the making - but a vintage portrait of a man, or even a woman, at their desk battling with thoughts on paper. This vision of authorship invites a strong protection of property rights on books, including long copyright terms, an underestimation of the input authors obtain from external sources, and a corresponding deference towards the author's right to exclude others from accessing the work. Again, how much this romantic view of the author actually remains influential on the contemporary copyright doctrine is contested in scholarship.[120]

2.2 POLITICAL ECONOMY JUSTIFICATIONS AND UTILITARIANISM

Today the prevailing theories of copyright are those grounded on an economic rationale. We will see how these theories evolved from a simple understanding of monopoly rights as a necessary evil to more elaborate utilitarian theories and economic analyses. Utilitarianism has its own moral dimension as it ultimately aims at the maximization of happiness or preference satisfaction. We will see how this plays out in the context of copyright. The common denominator of all these theories is that they understand and justify copyright as an artificial monopoly that the government created. The government decides to carve out exclusive privileges for authors to incentivize the creation of a presumably needed market. Thus, as opposed to insisting on some kind of moral rationale, they are basically political economy approaches. I argue, however, that political economy and property theory approaches to copyright end up adhering to the same consequentialist rationale so long as they rationalize an access/incentives tradeoff. Let us first start with the political economy justifications for copyright.

We saw how Adam Smith mentioned the exclusive privileges on books as a necessary compromise in order to reward authors for their merit. Smith was the first to express in simple terms what later thinkers developed as a political economy approach to copyright and transformed into more advanced utilitarian thinking. We also mentioned Thomas Babington

[120] See the debate between Boyle and Lemley: Boyle, J., 1996. *Shamans, software, and spleens*. Harvard University Press and Lemley, M.A., Romantic Authorship and the Rhetoric of Property'(1996). *Tex. L. Rev.*, *75*, pp.873-887. On the topic of romantic authorship see mostly: Rose, M., 1993. *Authors and owners: The invention of copyright*. Harvard University Press.

Macaulay and his contribution to the early parliamentary debates in England about the extension of the copyright term. He justified copyright as a necessary tax on readers for writers. At the time of Adam Smith, copyright on books lasted for fourteen years. In 1841, however, Macaulay was already arguing against an extension of a lengthier copyright term, which at the time the law was copyright for life or for twenty-eight years, whichever is longer. He opposed the then-proposed extension of copyright term to life plus sixty years. In his speech, Macaulay posed the critical question before making any final decision on the extension: on what principles is the question of the term extension to be argued? In his own words, the question was framed as follows: "Are we free to legislate for the public good, or are we not? Is this a question of expediency, or is it a question of right?"[121] Although at the time many were arguing for extension on a rights-based discourse, in particular the sacred and natural right of authors to the property of their own ideas or fruits of their imagination, Macaulay understood the issue differently. In principle he agreed that if this was the case, if it was a question of rights, then they should be protected irrespective of the cost: "Now, Sir, if this be so, let justice be done, cost what it may". [122] He differentiated the case because at that time the protection of authors' rights during their lifetime was not contested. What was at stake was the right of the legislator to determine what happens with the exclusive right it granted the author during his lifetime after his death. And in this determination, it was clear to him that the legislator should account for the public good. Any other rationale would cause a cost to the public that is not justifiable.

I find this old debate worth citing because Macaulay made what today seems like an economic analysis of copyright, mentioning the costs and benefits of providing copyright privileges and suggesting caution to extend such privileges. For him, the benefits of affording copyright to authors was that of having professional authors, who do not need to rely on other professional activities to survive, and also of not excluding those who are not independently rich. In sum Macaulay, like Adam Smith, insisted that the evil effects of the necessary monopoly that copyright creates should be carefully restricted to what is strictly proportionate and necessary. For both thinkers, the public benefit was still the primary rationale for making this proportionality balance. It is noteworthy that these early views considered the public benefit to be *prima facie* harmed by the monopoly rights on books created by copyright laws. Copyright is

[121] Young, G.M. ed., 1967. *Macaulay: Prose and poetry*. Harvard University Press.
[122] Ibid.

accepted as the necessary evil or what seems to be the second best solution to avoid dependency of authors on patrons or the struggle to find alternative means to survive and produce less or bad books. In the balance between what is beneficial for the authors and what benefits the public, the latter carries much more weight.[123]

Moving to the 20th century, Arnorld Plant was perhaps the first to provide a modern economic analysis of copyright on books in his article, *The Economic Aspects of Copyright in Books*, published in 1934.[124] Plant wrote this article with the explicit intent to provide a relevant economic analysis to a social policy issue.[125] He can be added to the political economy approaches to copyright in view of a number of arguments made in his article. He adheres to a general understanding of copyright and therefore supports professional authors (as oppose to authors that write irrespective of monetary returns).[126] However, he does not agree with another general understanding, that patronage is a necessary evil which copyright must under all

[123] Ibid. The relevant part of Macaulay's speech is worth quoting:

> The principle of copyright is this. It is a tax on readers for the purpose of giving a bounty to writers. The tax is an exceedingly bad one; it is a tax on one of the most innocent and most salutary of human pleasures; and never let us forget, that a tax on innocent pleasures is a premium on vicious pleasures. I admit, however, the necessity of giving a bounty to genius and learning. In order to give such a bounty, I willingly submit even to this severe and burdensome tax. Nay, I am ready to increase the tax, if it can be shown that by so doing I should proportionally increase the bounty. My complaint is, that my honorable and learned friend doubles, triples, quadruples, the tax, and makes scarcely and perceptible addition to the bounty.

There is obviously also a certain kind of romanticism associated to the passage's reading and writing, just like the one attributed to moral theories of copyright. It nevertheless fits the historical context which started, as noted elsewhere, with battles against a protectionist system that only favored some printers and the censoring powers of the State and the church.

[124] Plant, A., 1934. The economic aspects of copyright in books. *Economica*, *1*(2), pp.167-195.

[125] Ibid, p.167 citing Frank Knight (Risk, Uncertainty and Profit):

> Having started out by insisting on the necessity, for economics, of some kind of relevance to social policy- unless economists are to make their living by providing pure entertainment or teaching individuals to take advantage of each other," he discusses the conditions of relevance of economics to social policy, and his " first and main suggestion" is that an " inquiry into motives might well, like charity, begin at home, with a glance at the reasons why economists write books and articles. "Direct monetary profit from the sale of what they write does not figure in Professor Knight's suggestive discussion of the motivation of economist-authors; although for three, if not four, centuries the advocates of property in the right to copy have argued as though book production were the conditioned response of authors, publishers and printers to the impulse of copyright legislation. An inquiry into the rationale of copyright seems therefore both worthwhile in itself and likely to prove of general interest among students of economics.

[126] Ibid, pp. 169-170.

circumstances replace.[127] Citing Macaulay's speech at the House of Commons, he sees how patronage was understood as a system that is "fatal to the integrity and independence of literary man."[128] However, Plant's analysis immediately questions this common understanding on the mid-19th century figures coming from the profitable book trade between England and the United States absent any copyright law.[129] After reviewing the history of copyright law in England, he arrives at the high prices that the Copyright Act of 1842 ultimately resulted in and discusses the arguments of publishers: "the higher the profits from the copyright monopoly, the greater the willingness to publish the doubtful successes."[130] Plant further discusses the economic consequences of copyright, such as anticompetitive practices of publishers against authors and discriminating price policies for foreign markets. He ultimately agrees with retaining copyright as a 'necessary evil' that has already created many vested interests but insists on keeping the length of protection as short as possible.[131] Evidence showed then, as remains the case today, that the commercial life-span of books is very short. Thus, Plant concludes that while he is in favor of retaining copyright so that the book market, as has evolved, is not fundamentally disrupted, there is a need to minimize the length of protection in support of a more important public interest:

> There cannot be any question that it would be in the public interest to ensure low prices for books as early as possible after the fate of the first edition has revealed that a demand exists for them; and it is an important feature of the proposed amendment that it involves no new administrative principle, that it enables the continuance of, if not indeed an increase in, the remuneration of authors whose books the public want, and lastly that it simply confirms and extends throughout the book-publishing business the price policy which successful publishers already pursue in their own interests.[132]

[127] Ibid. p.170
[128] Id.
[129] Ibid, pp.172-175.
[130] Ibid, p.183
[131] Ibid, p.194.
[132] Ibid, p.195.

A clearer view of copyright as a dynamic private finance system arose in the work of economists that emerged much later than Plant. Ruth Towse, Christian Handke and Paul Stepan offer a comprehensive review of the economics literature on copyright.[133] In their intervention, the authors review most of the economics literature from the mid-1920's until today.[134] They explain that the literature on copyright places an emphasis on the balance between social costs and benefits on a Pareto scale (welfare economics). Thus, copyright is generally considered the second alternative following an imaginary state of Pareto efficiency, which is basically the same conclusion that we witnessed in earlier analyses before the 20th century. This imaginary state is a static state of perfect competition for books, marginal cost-pricing, no public goods externalities, no transaction costs and, as Towse and her co-authors note, a state of no technological progress.[135] In such an imaginary state, copyright would be unnecessary. However, the high sunk costs involved in the production of intellectual works, which make marginal cost pricing impossible, together with the public good nature of information (non-excludability and non-rival consumption) naturally lead to a monopoly. This is a monopoly where the supplier, in our case the authors and publishers, can supply the market more efficiently than competition.[136]

Finally, Towse, Handke and Stepan observe how most economists have been slipping too easily from static to dynamic reasoning. Then, the imaginary Pareto efficient state described above would also be static, which is contrary to the dynamic feature of copyright. The nature of the copyright incentive is dynamic, however. As the authors explain, the resulting social benefit of copyright "comes at a private cost borne by earlier generations of consumers, the time lag depending upon the duration of the copyright."[137] This time lag is usually portrayed as a deadweight loss. Towse, Handke and Stepan emphasize that the term deadweight loss is misleading because it confuses the static and dynamic effects of copyright.

---

[133] Towse et al., supra note 68.
[134] Their analysis includes among others:
Liebowitz, S.J., 1985. Copying and indirect appropriability: Photocopying of journals. *Journal of political economy*, *93*(5), pp.945-957; Caves, R.E., 2000. *Creative industries: Contracts between art and commerce*. Harvard University Press; Towse, R. 2001. *Creativity, Incentive and Reward: An Economic Analysis of Copyright and Culture in the Information Age*, Edward Elgar Publishing; Lipsey, R.G., 2007. Technological transformation, intellectual property rights and second best theory. *Review of Economic Research on Copyright Issues*, 4(2), pp. 5-28
[135] Towse et al, supra note 68, p.4
[136] Ibid.
[137] Ibid, p.3

Of the most cited works providing economic analysis of copyright is found in the *Economic Structure of Intellectual Property Law*, a 2003 book by William Landes and Richard Posner, who represent the Chicago school of law and economics.[138] Their theory of intellectual property assumes "continuity between rights in physical and in intellectual property."[139] Looking into the rationale behind intellectual property protection of intangible goods, they start with the broadly accepted incentive/access tradeoff but also add their own argument for avoiding congestion externalities. In other words, they think that without intellectual property we would face a tragedy of the commons. They are also concerned about free-riding in light of the public goods characteristic of intellectual property goods.[140] They continue to see value in maintaining a public domain but support limited intellectual property protection in order to avoid the tragedy of overconsumption of intellectual property goods.[141] In contrast to the benefits of maintaining intellectual property rights, there is the issue of costs, which according to the authors are basically the same costs attached to every property right: transaction costs, rent-seeking and costs of protection.[142] They seem to place particular emphasis on the costs of protection and argue against long protection of intellectual property rights.

Landes and Posner offer their own economic model of optimal copyright protection. Important for us is how their model applies to books. As they explain, the cost of producing a book, "cost of creation", consists of the cost of expression, which is fixed, and the cost of printing, which is variable.[143] In their model, the decision whether to create a work, or not, rests on the difference between the expected revenue and the costs of creating and making copies of the work. If the revenues and the cost of making copies equals or exceeds the cost of expression, then the book will be made. If there was no copyright protection, the expected revenue reduces because the market price will "bid down to the marginal cost of copying."[144] In that case authors and publishers will lack the incentive to create the work in the first place. Notably the authors seem to ignore differences in the costs and incentives between authors and publishers and in general portray a market of risk-averse authors and risk-neutral publishers.[145] The actual

[138] Posner and Landes, supra note 68.
[139] Ibid, p.9. Yet they do somehow differentiate property from intellectual property (see p. 36).
[140] Ibid, p.23.
[141] Ibid, pp. 14-15
[142] Ibid, p.16 et. seq.
[143] Ibid, p.37 et seq.
[144] Ibid, p.40.
[145] Ibid, p.38. See also at p.48 the proclaimed differences between authors and publishers.

economic model they describe fits how the publishing market generally operates: the risk of a lack of demand for the published book is calculated into the price. Thus the price starts higher than the marginal cost of creation.

Despite their being highly influential, or at least much cited, Landes and Posner's theory has not diverged significantly from the previous economic analyses, which were sketched already. While Landes and Posner fit squarely under the utilitarian model, at least in their study of copyright on books, they have not broken new ground. Their bigger contribution possibly lies with their congestion argument, which will be discussed when we study objections to the tragedy of the commons arguments and commons approaches to intellectual production. In short, Landes and Posner argue that negative consequences arise as a result of the overuse of the public domain, which can be compared to the consequences of unlimited drilling from a common oil pool. They also argue that works will be under-exploited once copyright expires and they enter into the public domain.

Utilitarian views based on the work of Landes and Posner have dominated copyright scholarship. There is, however, an already clear tendency that questions utilitarianism and its effects. The normative opposition is also linked to the transformations in intellectual production and consumption in the digital era.[146] There is a case to argue that we are nearing the end of the utilitarian era of copyright. A relevant critique to the evolution of copyright rules according to an economic rationale is offered by Paul David in 2004 where he claims that copyright protection has been shaped "more by the economics of "publishing" than the economics of "authorship.""[147] David claims that the development of advanced printing methods since the 19th century led to strategic advocacy for longer copyright terms by publishers who invested in mass production and mass distribution and sought to ensure high profits and fixed costs.[148] David further explains how the digital revolution has brought about the most serious disruption to the publishing industry so far. The next part is devoted to understanding this transformation and connects it to the revolution that advanced methods of copying have instigated.

---

[146] Among many, see the critique of Julie Cohen in Cohen, J.E., 2012. *Configuring the networked self: Law, code, and the play of everyday practice*. Yale University Press.

[147] David P. 2004. The End of Copyright History? *Review of Economic Research on Copyright Issues*, 1(2), pp. 5-10.

[148] Ibid, at p.6: "The greater fixed costs of the high-speed presses, the introduction of which technically complemented the replacement of moveable type by stereotype, gave rise to the exploitation of the economies of mass-production and mass-distribution, transforming business strategies and market structures in both the book trade and the newspaper industries."

Emerging out of our analysis is the question of whether the copyright framework conceptualized to grant property rights to authors, and other right-holders, for a limited period of time is the second best option, which most economic theories seem to assume. Against this position, which solution is the best? And is the reason behind its second best status why determining an optimal copyright length has been so difficult? Below we will take a closer look at the issue of the optimal copyright length. Before that let us add an important challenge in this discussion: since copyright rules apply not only to books but also to very different intellectual products, the problem might also lie in the one-size fits-all copyright framework that currently exists. Even within the distinct category of books, it is possible to argue for different types of protection in accordance to the need; for example, to sustain incentives for authors or an industry. The backlash that copyright laws has created in academic disciplines is apparent and clearly expressed by the Open Access movement or the design of alternative licenses that try to work around copyright, through the Creative Commons licenses. These fundamental questions on the future of copyright are beyond the scope of this thesis but should be taken into consideration. For now we will continue with the existing debate over the optimal copyright length, which is still framed in utilitarian terms.

2.3 THE DEBATE OVER THE OPTIMAL COPYRIGHT LENGTH

The optimal length of the copyright term has been long debated, particularly when legislative extensions were proposed. There have been scholarly efforts to present a model of the optimal scope of copyright for protected works.[149] While some formal models tend to focus on the optimal length for the recovery of sunk costs during a period of supra competitive pricing, without always considering the question of access and distribution of existing works or the costs imposed on follow-on creation, the biggest problem is not on debating the effectiveness of the proposed model. The ultimate issue comes down to whether the legislators actually take any of these models into account and if yes, then to what degree.[150] Let us look at the past two

[149] Posner and Landes, supra note 68, p. 89 et. seq.

[150] See, for example, the famous brief of the seventeen economists submitted as amicus to *Eldred v. Ashcroft*, 537 US 186 (2003): Brief of George A. Akerlof, Kenneth J. Allow, Timothy F. Bresnahn, James M. Buchanan, Ronald H. Coase, Linda R. Cohen, Milton Friedman, Jerry R. Green, Robert W. Hahn, Thomas W. Hazlett, C. Scott Hemphill, Robert E. Litan, Roher G. Noll, Richard Schmalensee, Steven Shavell, Hal R. Varian, and Richard J.

extensions of the copyright term in the US and the scholarship related to these legislative developments.

Stephen Breyer, a current Justice at the US Supreme Court, published an article in 1970 on the occasion of a serious extension of the copyright length in the US from 56 years of protection to the life of the author plus 50 years.[151] The article provides an extensive analysis of the moral and economic justifications of copyright and makes the case against an extension. Most importantly, Breyer focuses on the book market and evaluates several possible moral and economic justifications to copyright against actual factual evidence on the American book market. One of his most compelling arguments centers on the question of what would happen if copyright protection were abolished.[152] On the basis of his analysis, abolition of the copyright system would not seriously harm book production in the two groups of works under study: textbooks and tradebooks. He also discusses alternative compensation schemes for publishers; for instance, the organization of groups of buyers who would pay in advance to see works that have particularly high fixed costs published or the option of government subsidies. Replacing copyright with government subsidies is an idea that raises the danger of censorship. In contemporary scholarship, lawyers and economists have tinkered with the idea of replacing copyright with some kind of subsidies or rewards system. Some have also written about abolishing copyright entirely.[153] Yet the term of protection continues to extend. International norms raised the protection to the life of the author plus 50 years as a minimum standard.[154]

Thirty-three years after Breyer's article, extension was debated again in the American legal context. This time it was a judicial challenge that reached the US Supreme Court. In 2003, the court infamously upheld the last extension of the 'Sony-Bono Act' in *Eldred v. Ashcroft*.[155]

Zeckhauser, *Amicus Curiae* in support of Petitioners (May 20, 2002). For a critical view on the brief see Liebowitz, S.J. and Margolis, S.E., 2005. Seventeen famous economists weigh in on copyright: The role of theory, empirics, and network effects. *Harv. J. of L. & Tech.*18(2).

[151] Breyer, S., 1970. The uneasy case for copyright: A study of copyright in books, photocopies, and computer programs. *Harvard Law Review*, pp.281-351. At the time that this article was written US Congress was considering the first major revision of the Copyright Act of 1909.

[152] Ibid, p.293 et. seq.

[153] See Shavell, S. and Van Ypersele, T., 2001. Rewards versus intellectual property rights. *The Journal of Law and Economics*, *44*(2), pp.525-547; Fisher, W., 2001. Theories of intellectual property. *New essays in the legal and political theory of property*, *168*(1) and on the abolitionists side: Boldrin, M. and Levine, D., 2002. The Case against Intellectual Property. *The American Economic Review*, *92*(2), pp.209-212.

[154] Berne convention for the protection of literary and artistic works, of September 9, 1886 (as revised, Rome 1928, Brussels 1948, Stockholm, 1967).

[155] 537 U.S. 186 (2003). Breyer repeated his arguments against lengthy copyright protection in dissent together with 2 Stevens.

By that time, most European jurisdictions were also offering the same lengthy protection. It is remarkable that during the *Eldred v. Ashcroft* litigation seventeen economists, among which five Nobel-prize winners, wrote an amicus curiae brief of to the Supreme Court arguing against the extension, which the court upheld.[156] The brief cited a Congressional Research Service study that found only a small percentage of works, which were under copyright during the 1920's and 1930's and renewed in the 1950's and 1960's, had any commercial value in 1998 (the numbers were below 12% for some categories of works, such as books and musical works).[157] Economists generally agree on the point that the majority of books do not have commercial value for long, even if they argue with respect to the lifespan of the commercially valuable ones.[158]

In any event determining the appropriate length of copyright protection over books on the basis of figures calculating their commercial lifespan is relevant not only as part of utilitarian theories that try to pinpoint the economically optimal span of protection but also for a number of additional reasons. The legal treatment of out of print books is an issue closely related to this discussion and is especially relevant to this study. We shall return to the issue later when we look at authorizing digitization of out of print works.[159] We will first turn to the dichotomy between books under copyright and books in the public domain. We will then continue with a categorization of books, in accordance with a more flexible division starting with copyrighted works and arriving at the public domain. The categorization is based on the effects that copyright has on their availability and uncovers two intermediate states that are more 'obscure': out of print works and works orphaned by active right-holders. This discussion takes place in chapter four of the thesis, which explains the range of possibilities but also difficulties that libraries face as a result of the copyright status of the books in their collections.

In the following chapter, libraries will once again be center-staged, zooming in on their functions and roles with regard to providing access to books outside of markets.

---

[156] Supra note 147. For the text of the *Amicus Curiae* brief see, online available at: https://cyber.harvard.edu/openlaw/eldredvashcroft/supct/amici/economists.pdf.

[157] Ibid.

[158] Liebowitz, S.J. and Margolis, S.E., 2005. Seventeen famous economists weigh in on copyright: The role of theory, empirics, and network effects, supra note 147. Liebowitz and Margolis cite another study, that of Liebowitz himself (unpublished) from a small data set consisting of a sample titles reviewed in Book Review Digest in the 20s along with bestsellers. He found that more than half of the best-sellers in the sample remained in print for more than fifty-eight years, with differences according to the category of books (academic, philosophical, history, biography, religion etc.). They demonstrate that "for the small number of titles generating the lion's share of economic value, life expectancy is rather long. Extending copyright might have only a small change in expected revenues for these books, but not because they have gone out of print or have lost commercial potential."

[159] *Infra*, in chapter four, (4.3).

CHAPTER THREE - BOOK DE-COMMODIFICATION: LIBRARIES AS RESERVED PLACES FOR LEGAL DISTRIBUTION OF BOOKS OUTSIDE OF MARKETS

Libraries are praised for their neuralgic role within a democratic society: they serve as an independent and public space for lifelong learning, irrespective of income and access to formal education.[160] Libraries are welcoming, study-friendly environments that provide library users learning spaces and resources. They also provide professional help to navigate the sources and a space which is not limited to individual work but encourages also collaborative endeavors if so desired. It is little wonder that libraries have been traditionally recognized spaces for scholarship, creation and innovation.[161] Libraries hold collections that are curated, systematized and have a historical continuum. Usefulness and collaboration of the collections are more important than their popularity and marketability. Finally, libraries, especially national ones, hold historical and rare collections which they preserve for generations, past and future, thus they carry out an archiving and gate-keeping function on behalf of knowledge and history.

This chapter focuses on the importance of libraries as reserved places that are lawfully permitted to distribute books outside of markets. The chapter will present some of the most important functions of libraries and explain the significance of library space and accessibility to a public space and public services. It will then proceed with an analysis of the institution's positive externalities and unpack what it means to maintain such public service within a society. Ultimately, the emphasis is placed on the role that libraries play in book de-commodification.

The founding scholars of law and economics explained how the production and distribution of goods can be implemented through either the market or command structures. While copyright helps the production and distribution of goods via markets, there are exceptions to copyright that allow institutions, like libraries, to operate in a space that is arguably outside the market and the command spectrum. Both for the book market and libraries, the relevant good is the book, which is apparently a good that is not successfully optimized either by pure market or

[160] Palfrey, supra note 40, p.58

[161] Pettegree, A. The Renaissance Library and the Challenge of Print, in Alice Crawford (ed.) 2015. *The meaning of the library: a cultural history*, Princeton University Press.

by pure command structures. Quoting Calabresi, "it is hard to handle a good that cannot be successfully optimized by pure market or pure command structures."[162] With this premise in mind, the present chapter will try to explain the space where libraries fit, which is situated between a pure market and a pure command structure.

3.1 LIBRARIES FROM THE ANCIENT TIMES TO THE THIRD MILLENNIUM: FUNCTIONS AND SYMBOLS

The history of libraries can be told as a story about culture, ideas and fantasy. According to historians, we inherited the idea of a library from the ancient Mediterranean and near eastern worlds.[163] Since ancient times, libraries have been built and maintained as institutions combining functions and symbols.[164] The functions of a library include the collection, organization, storage and preservation of documented knowledge in the form of books, journals, and other similar mediums as well as provide access, restricted or more open, to these works. Traditionally the works that libraries have held were literary works in the broadest sense of the term, which in a historical context included scrolled papyrus, and then shelved books and, later, media content.

The core functions of libraries have evolved over time. In the words of John Wilkin, librarian and previous executive directive of the HathiTrust consortium, libraries have always preserved the cultural record as well as provide access and ensure use of that record.[165] What this 'preserving a cultural record' means precisely is a concept that has evolved as the collections and spaces that libraries occupy evolved. Wilkin distinguishes four pillars of activities that are clearly library functions according to our understanding of what constitutes a modern 'brick' library. These pillars are: (i) curation: the selection, preservation, maintenance, collection and archiving of, and provision of access to, materials pertaining to the cultural record which is mostly books and manuscripts, but also images and audiovisual items; (ii) engagement with research and learning; (iii) publishing; and (iv) creation and management of spaces for users and

[162] Calabresi G. 2016. *The Future of Law and Economics. Essays in Reform and Recollection*. Yale University Press, at p.137.
[163] Alice Crawford. 2015. *The meaning of the library: a cultural history*, Princeton University Press.
[164] Wilkin J., Meanings of the library today, in Alice Crawford (ed.) 2015. *The meaning of the library: a cultural history*, Princeton University Press.

[165] Ibid, at p. 236. Notably, Wilkin adds that libraries are increasingly involved in the creation of the cultural record.

collections.[166] Not all libraries are engaging with all four pillars of activity with the same intensity; the categories, however, cover more or less the principle functions of most libraries.

Because the library collects and preserves our cultural record, itself a sum of valuable resources, by extension the library is in its own right valuable as a resource. It is important to note that libraries are a resource greater than the sum of its parts. The library's wealth is its collections. They constitute the main occupation of librarians and attract patrons, thus make the libraries useful. Depending on the library, the collection can be broad or specialized in a specific field or science. If compared to the past, library collections have today become more complex, going beyond print works and manuscripts to images and audiovisual materials, e-resources, such as electronic journals and periodicals. At the same time the data and metadata that libraries have to deal with are simply big. Thus, the nature of the curating function has changed substantially.[167]

While performing their functions, libraries have maintained a significant symbolic role within societies, which continues to be celebrated in multiple and usually ambitious words: we attribute to libraries mandates amounting to the *preservation of the memory of mankind* or the storing of the *sum of human knowledge*. Thus, they transform to almost mythical places where one can engage with the pursuit of knowledge and finding the *truth*.[168] Libraries have also been associated with legends, which tell stories about their destruction. All myths and stories associated with the world's greatest libraries suggest that the creation of libraries has been understood as a self-evident good and their ruin as a real disaster. The library of Alexandria, well-known for being burnt down by a fire started by Julius Caesar in 48 BC, is the most cited example.

The historical symbols also contribute to the value attached to this institution. Libraries built in the kingdoms that Alexander the Great established, for example, served to institute new cultural hubs in continuance of the empire. To take the example of the library of Alexandria again, it was an institution ambitiously aimed at acquiring every work ever written and representing the cultural strength of the empire. In Athens, one can still find the plaque

[166] Ibid, p. 237.
[167] See also the discussions about libraries and big data and concerns about privacy. Since November 2016 the New York Public Library, for example, published its new privacy policy regarding the data it is gathering, available at https://www.nypl.org/help/about-nypl/legal-notices/privacy-policy/intro.
[168] If we link this to the Socratic idea that knowledge has a moral value, we can see how the pursuit of knowledge and truth has been established as a noble cause already since ancient times.

indicating the opening hours of the library of Hadrian, built around the first century A.D., indicating an honor code system for maintaining the books where they are and available for all free men of the city. The plaque reads: "No book will exit [the library] because we so vowed; [the library] will be open from the first to the sixth hour."[169] Before that and tied to the city's strong literary past, during the so-called Golden Ages, there are historical references to private collections of books belonging to intellectuals, mainly politicians or philosophers, of the time. Indeed, the city's attention to knowledge production and dissemination differentiated to rival cities. In ancient Rome we find libraries in public baths, just like we find them in records offices. We find them providing rooms where important men of the time gathered, important discussions on public affairs were held, and important decisions were taken. After the invention of the press, with the vast proliferation of books and book trade libraries also proliferated. Some of the most important libraries to this date started with small collections within the first academic institutions in Europe and later spread across the rest of the world.

It is safe to say in general that historically the existence and functioning of libraries has been widely appreciated as something positive for society. As institutions, libraries enjoy until this day respect and appreciation for their symbolic and actual functions within the communities they serve. Of course, library skeptics have made their voice heard.[170] Umberto Eco had famously expressed his critical views on libraries on several occasions, including in a speech at the Bibliotheca Comunale di Milano in 1981.[171] Eco's particular relationship with libraries has been seen as critical, especially in the *Name of the Rose*.[172] We will discuss Eco again later, when we analyze what is meant by his criticism of 'bad libraries' (la biblioteca cattiva) for, *inter alia,* 'hiding material'.[173] Suffice to say that he was not against the institution, quite the contrary, but he definitely has a clear vision of how the institution should work and what it should be.

---

[169] *ΒΥΒΛΙΟΝ ΟΥΚ ΕΞΕΝΕΧΘΗΣΕΤΑΙ ΕΠΕΙ ΩΜΟΣΑΜΕΝ ΑΝΥΓΗΣΕΤΑΙ ΑΠΟ ΩΡΑΣ ΠΡΩΤΗΣ ΜΕΧΡΙ ΕΚΤΗΣ* - Supplementium Epigraphicum Graecum, XXI 500.

[170] As for example Umberto Eco (see in particular *The Name of the Rose*) or Michel Foucault (*The Archeology of Knowledge*). See also Hall E., Adventures in ancient Greek and Roman Libraries. in Alice Crawford (ed.) 2015. *The meaning of the library: a cultural history*, Princeton University Press, at p. 23: "Libraries might have killed innovation and experiment in Greek poetry but offered a fertile ground for other genres such as biographies and sciences such as geography and astronomy to advance."

[171] Winter, M.F., 1994. Umberto Eco on Libraries: A Discussion of" De Bibliotheca". *The Library Quarterly*, *64*(2), pp.117-129.

[172] Garrett, J., 1991. Missing Eco: On Reading "The Name of the Rose" as Library Criticism. *The Library Quarterly*, *61*(4), pp.373-388.

[173] Winter, supra note 171. See also *infra* in the concluding chapter.

Starting as far back as Socrates, many philosophers are known for having particular views on documenting knowledge, on mainstream ambitions to 'gather the sum of human knowledge', and perhaps also on the hypocrisy of intellectual elitism when it comes to fancy collections and fancy buildings.

3.2 LIBRARY SPACES, ACCESSIBILITY AND PUBLIC SERVICE

The library's usefulness – or broadly speaking its value - is attributed not only to its functions but also to its spaces: today these are spaces are mostly publically accessible and operate within larger or smaller communities. Publically accessible libraries can include national, communal, general and specialized ones as well as academic libraries which, depending on their regulations, can be more or less accessible to a general public beyond the campus community. Literally speaking libraries are public or semi-public spaces (for the purposes of this thesis we exclude strictly private collections). They may set certain requirements for patrons they allow onto their premises, they may require registration, and they have certain hours of operation and so forth. These requirements and restrictions, however, do not disprove their public character. On the contrary, as long as their admission to the public is free or at a minimal cost, and regulated by objective rules that do not discriminate or substantially restrict anyone, they are an accessible place for people irrespective of income and other differences. They are also accessible and welcoming to younger and older generations. Given their functional public spaces, libraries have been brought closer to social work. This is, for example, a predominant phenomenon in public libraries in the United States. In many cities library space has become a shelter for the homeless and the jobless.[174] Similar challenges are faced today by European libraries in view of the ongoing social crises. This results in complex situation for the librarians who frequently engage in social work, collaborating with other public services and social workers. In spite of the challenge that they face as institutions, the libraries can present underprivileged groups with opportunities. One can note a proliferation of programs that assist job-seekers in many ways.

[174] See IFLA's current project on developing guidelines for library services to people experiencing homelessness. Draft guidelines: https://www.ifla.org/files/assets/lsn/projects/homeless/draft-guidelines-homeless-call-for-review.pdf.

Free access to computers and to the Internet are fundamental and assist in the direction of employment activities.

There is important data indicating the reach of the library's public service. As already noted, libraries are used by citizens of all different age groups. In 2013, for example, 54% of all Americans aged 16 and older said they had visited a library within the previous twelve months. In view of such data, it is hard to explain why public libraries are constantly underfunded.[175] John Palfrey devoted his latest book to the vital role played by libraries in today's digital age, underscoring how libraries have become more important than ever. Many of his sources demonstrate the usefulness of libraries in the lives of millions of patrons around the United States. Palfrey goes on to say:

> Adults tell researchers that they use libraries to borrow books, DVDs, CDs, materials for job-seeking, and materials related to becoming citizens, as well as to spend time with others in a public place. Circulation of materials is up across the board, from 43 million materials in 2002 to 69 million in 2011 in the libraries of the city of New York alone. Attendance at public events in libraries is also on the rise. In New York City, 17,000 programs in 2011 attracted more than 2.3 million attendees, a 40 percent increase since 2002.[176]

Library accessibility is linked to an anti-elitist culture. Of course there are many elite institutions around the globe that are proud of their rich library collections and they are accessible primarily to their patrons, usually graduate students that pay large sums in tuition. By contrast, national and municipal libraries, however, are mostly associated with an accessibility and common goods discourse. The Boston Public Library, known to be the first large and free public library in the United States, famously displays the 'Free to All' inscription above the main entrance of its main buildings, the McKim Building. This sign simultaneously does and does not mean free for all. In spite of the beautiful motto, there are of course rules in place regulating access to buildings and access to materials: "A library card from the Boston Public Library provides access to library materials and services for users in all Boston Public Library locations

[175] Palfrey, supra note 40, p.57.
[176] Ibid.

and via the library's website."[177] And those who can obtain a 'full service' Boston public library card are Massachusetts residents, Massachusetts property owners, people commuting to Massachusetts for work, students residing in Massachusetts while attending school and residents of temporary housing in Massachusetts.[178] This example shows the correlation between access and the community, which is mostly defined by two criteria: (i) a geographic association and (ii) payment of tax. Still, accessibility is remarkably wide. First, there is no requirement to pay fees. Second, practically no person within the whole State is excluded; this library is also accessible to children, whose card is issued with the consent of their parents if under twelve years old. And third, the richness of the resources available to the patrons is immense; the Boston public library is one of the only two public libraries that belongs to the Association of Research libraries and holds a particularly rich collection of millions of books and audiovisual items.[179]

Funding for public libraries in the United States usually comes from mixed public and private sources. Most of the well-known public libraries started with generous donations of private donors, as is the case of New York Public Library. The public service character of libraries is also clear in Europe, where the funding for public, city or municipal libraries, usually comes directly from the State, thus the pockets of taxpayers. Despite the important role and functions of libraries as spaces and service providers, the status of library funding today is widely considered poor.[180] One can cite numerous studies, news reports, librarians' and citizens' ongoing complaints on this issue. For the purpose of dialogue that this thesis wishes to engage in in the following sections but mostly in the subsequent chapters, I will at this point cite a concise but powerful comment found in the Library Quarterly in an article written by librarian at UC Davis, Michael Winter, in 1994:

> Librarians are of course well aware of the importance of cost, but what is interesting here is the notion of a more general theoretical, historical, and even

[177] Boston Public Library access policy online available at: http://www.bpl.org/general/policies/accesspolicy.htm.
[178] Boston Public Library service rules online available at: http://www.bpl.org/general/circulation/borrowers.htm.
[179] Quoting from the library's website (http://www.bpl.org/general/history.htm): "Today, the Boston Public Library system includes a Central Library, twenty-four branches, a map center, a business library, and a website filled with digital content. Last year, 3.7 million people visited the Boston Public Library system, many in pursuit of research material, others looking for an afternoon's reading or the use of the computer or to attend a class, still others for the magnificent and unique art and architecture of many library locations. There were 7.9 million visits to the library's website and 3.7 million books and audiovisual items borrowed or downloaded." (accessed on May 6, 2017).
[180] See also the contemporary discussion on the possible cut of federal funds: Templeton, supra note 16.

> ideological dimension of the subject. Cost, that is to say, is much more than a practical problem; it is a historical and socioeconomic phenomenon that reflects broader social processes and values. And while we tend to view material and intellectual value as incommensurable, they are in this respect closely intertwined.[181]

This passage ties well with the general discussion on copyright, the commodification of books and the role of libraries as spaces reserved for the legal distribution of books outside of markets. Libraries have distributive effects that regulation has, to a certain extent, aimed at supporting. These distributive effects are considered valuable externalities also in economic terms, as will be outlined in the next section of this part.

3.3 POSITIVE EXTERNALITIES

In light of their functional and symbolic roles, libraries are important to a broader part of the society than those actually who use them. There is a great deal of positive externalities associated with the library's general institutional role which surpasses the value of the sum of its parts as well as the benefits it generates to its immediate users. We understand this value as part of the positive attributes of literacy, scientific progress, and the *advancement of knowledge* in general. Without preservation of a cultural record and without access to the material preserved, new production of knowledge could not easily rely on and be built upon knowledge already produced.

To better understand the value that societies assign to libraries we need to think further about why the maintenance of libraries as institutions is useful. Is it useful just to a few people, perhaps only library users that benefit each time from working in library spaces or from checking out a book, or does it serve a general use? The social value that library users derive from libraries can vary. Likewise, the patron's uses can also vary. Having addressed some of the uses and the social role of libraries in the lives of patrons let us now discuss the issue of positive externalities.

[181] Winter, supra note 173, at p. 121.

Libraries have an institutional role that differentiates them from other public spaces, such as a square or a park. In economic terms the institutional dimension of a library is perhaps closer to a *public utility*. Economists define public utilities as organizations that maintain the *infrastructure*[182] for a *public service*.[183] First, as we saw already, libraries offer a service to the public: the preservation of knowledge materials and access to intellectual resources. Second, they offer an infrastructure for the service that consists of publically accessible spaces, resources and catalogues to navigate the resources and professionals who provide the service working individually with the materials - acquiring, organizing and curating collections - and collaborating with patrons. Public utilities are generally *lifeline services* offered at low rates or, most often, *zero rates* and have *wide availability* and a *redistributing character*. This is traditionally reflected in their pricing policies which ensure that the services are available and affordable to different members of society.[184] Their broad availability is based on the principle of universal service.

Libraries' infrastructural characteristics are particularly important to our discussion on positive externalities. Thus, we will first look at infrastructure as a concept, and then the positive externalities resulting from the infrastructural characteristics. Functionally, infrastructure is a large-scale organization of resources. Familiar examples of infrastructure are: transportation systems (highways, railways, airline systems, ports), communication systems (postal services, telephone networks, and telecommunications, including broadband Internet, wired or wireless), governance systems (courts), basic public services and facilities (schools, sewers, water systems).[185] It has been rightly pointed out that as there is no distinct field of infrastructure studies the meaning of the term varies.[186] More precisely, scholars focus on various

[182] The term infrastructure generally refers to structures or facilities, usually technical, essential for the production and communication of goods, services, or whatever society considers as valuable. For our purposes here I use the term infrastructure, quite broadly defined, as a set of structures that facilitate communication of value - be it physical and non-physical resources, information resources, and even human resources. More specifically, I follow Brett Frischmann's definition of infrastructure as a particular set of resources defined functionally in terms of the manner in which they create value: Frischmann B. 2012. *Infrastructure – The social value of shared resources*. Oxford University Press.

[183] Black J., Hashimzade N., and Myles G. *A Dictionary of Economics*. Oxford University Press (ed. 4); Geddes R., Public Utilities, in *Encyclopedia of Law and Economics* 1161, 1165-67 (Boudewijn Bouckaert & Gerrit De Gees eds., 2000).

[184] Brown S. and Sibley D. 1986, *The theory of public utility pricing*. Cambridge University Press.

[185] Frischmann, supra note 182.

[186] Ibid, p. 3.

characteristics of infrastructure depending on the discipline they come from.[187] Relevant for us, infrastructure has certain characteristics that arguably justify a particular economic and regulatory approach. It is therefore necessary to identify and evaluate these infrastructural characteristics found in libraries and to take them into consideration when assessing their current and future regulatory status.

Leigh Star and Karen Ruhleder offer a comprehensive classification of a number of important characteristics in infrastructure.[188] First they noted its embeddedness: infrastructure is 'sunk' inside other structures, social arrangements and technologies and invisibly supports other tasks. Since infrastructure is embedded, we also take its existence for granted but it becomes visible when it breaks down. In practical terms, this means that while the infrastructure is physically there, actual users might not properly articulate their demand for its continued availability. Thus, there is a demand-side problem. The second feature concerns its reach or scope: when we think of infrastructure we think of structures with broad reach and the bigger the scope the better. Third, the fact that infrastructure is learned which stems from community membership: as member of a community, you take the existence of infrastructure within the community for granted and you understand the use of this infrastructure as an entitlement resulting from your community membership. Fourth, it shapes and is shaped by conventions of practice. Here, Leigh Star and Karen Ruhleder refer to the QWERTY keyboard as an example. It makes sense for an infrastructure to adopt rules or conventions of practice, be it tangible like a keyboard structure, or intangible. The merits of mass adoption of the same rules of practice are due to obvious benefits of standardization and network effects.

Now we will move on to the application of these infrastructural characteristics on libraries: First, libraries are embedded within communities and are convenient in a subtle, less visible manner. Thus, their value is best understood in the case of their loss. Library users, who are occasional, active or potential, may not adequately signal their social demand for libraries. As noted, this is a demand-side problem for libraries that the book market does not face in the same way. Bookstores can perhaps face demand-side issues in a different context but avoid or

[187] See, for example: Borgman, C.L., 2000. *From Gutenberg to the global information infrastructure: access to information in the networked world*. MIT Press; Star, S.L. and Ruhleder, K., 1996. Steps toward an ecology of infrastructure: Design and access for large information spaces. *Information systems research*, *7*(1), pp.111-134. Both works define infrastructure from the information studies perspective. See also Shane Greenstein, Internet Infrastructure, in *The Oxford Handbook of Digital Economy* (Martin Peitz and Joel Waldfogel eds.).
[188] Ibid (Star and Ruhleder), at p. 113.

solve these issues by simply directing more resources and advertising to popular works through creating or following trends. Second, their usefulness depends on their scope. They are useful to those who have access to them as members; however, membership is defined. The degree of accessibility is also important for libraries to remain relevant and useful.[189] Fourth, there is a certain protocol that one needs to follow when using them. It involves using search terms, understanding the cataloging system, following certain paths to find a resource and obeying the rules to access them. Standardization in the cataloguing system has obvious benefits because it allows patrons to use multiple libraries easily and also enables libraries with different collections to communicate (network effects).

There are a number of consequences that we must consider when thinking about the regulation of infrastructure: (i) that free markets often fail to meet society's demand for infrastructure, (ii) that the degree of accessibility defines the infrastructure and (iii) that infrastructure generates significant spillovers resulting in social gains.[190] In his book on infrastructure, Brett Frischmann focuses on the significant *value* of infrastructure in society as well as its embeddedness. Frischmann explains how there are two primary microeconomic perspectives on infrastructure: one is grounded in public welfare economics and emphasizes the public goods attributes of infrastructure, and another is grounded in regulatory economics and emphasizes the potential natural monopoly attributes of infrastructure industries.[191] He builds his theory primarily on a third point: infrastructure generates significant spillovers resulting in social gains. The social importance of spillovers, resulting from the creation of knowledge and also from innovation, is well recognized in theory. Scholars have expressed many theories to explain how to deal with these spillovers. Theories include government intervention, tax, subsidies or

[189] For brick libraries, accessibility is defined in the rules that the library operates (hours of operation, registration requirements). As per digital libraries what shapes accessibility is copyright, to which a substantial part of this thesis is devoted.

[190] Frischmann, supra note 182.

[191] Ibid, at p.9. A commons management strategy maintains openness, does not discriminate among users or uses of the resource, and eliminates the need to obtain approval or a license to use the resource. Frischmann contends that "managing infrastructure resources as commons eliminates the need to rely on either market actors or the government to "pick winners" among users or uses. It facilitates the generation of positive externalities by permitting downstream production of public and social goods that might be stifled under a more restrictive access regime."
Similarly Yochai Benkler had argued, already in 1998, the commons have been, mistakenly, a neglected factor of information policy: Benkler, Y., 1998, September. The commons as a neglected factor of information policy. In *26th Annual Telecommunications Research Conference* (Oct. 3-5, 1998), online available at http://www.benkler.org/commons.pdf.

well-defined property rights.[192] Under section 3.4 we will discuss in more detail the positive externalities that libraries generate and the importance that societies attribute to their role in book de-commodification. Copyright law is still relevant as a legal framework, given that libraries operate under its umbrella and also have to adhere to the copyright restrictions that apply to the books in their collections. The copyright framework is in theory one of the possible frameworks to deal with intellectual creation that is expected to generate spillovers and as we saw in the previous chapter: it is meant to combine the rationale of well-defined property rights together with government intervention (here we refer to the political economy justifications for copyright). The case for libraries, however, is more complex. The regulator here needs to encompass not only spillovers from widespread access to books but also from the existence of the library as a public service beyond the mere collection of books.

The key point with infrastructure, according to Frischmann, is that spillovers are not necessarily a problem to be solved.[193] In certain cases, society may be better off if it lets some externalities go without aiming to internalize them, and further it encourages participation in activities that generate externalities, again, without aiming to internalize them completely. In a sense a public service embraces and also promotes these externalities so that disadvantages, like free-riding, are institutionalized and ultimately their impact is eliminated. In other words, the benefits of having the public utility, or the public service, outweigh the costs. This point is, I argue, particularly relevant for understanding the value of libraries as institutions. In its role as

[192] For example Pigou, A.C., 2013. *The economics of welfare*. Palgrave Macmillan; or the older work of Coase: Coase, R.H., 1937. The nature of the firm. *economica*, *4*(16), pp.386-405.
The work of Coase has, indeed, been foundational to later law and economics scholarship:
In 1937 Coase made a subset distinction of the firm and the economic system. He observed the paradox of having two fundamentally different coordinating instruments within our economic system, the one within firms, which is coordination by command of the 'entrepreneur-coordinator', and that of the economic system itself which is a market or price-mechanism instrument of coordination. In other words the economic planning within our economic system and that within firms, which constitute integral parts in the economic system, is in a market-economy very different (p.388). Coase wanted to understand the reasons for this difference and also the basis for choosing to allocate resources in one way in some cases and in another way in others. While exploring these reasons he also noted that "exchange transactions on the market and the same transactions organized within a firm are often treated differently by Governments or other bodies with regulatory power" (p.393). According to Coase the reason why some resources are better allocated with a price mechanism and some not lies in the costs and benefits of using each method. There are both benefits and costs in using the price mechanism and the same applies for alternative solutions, as for example government regulation. In his 1966 article about the Federal Communications Commission, Coase explained how when certain actions produce too many externalities finding solutions via general regulation becomes easier compared to solutions through the market. [Coase, R.H., 1959. The Federal Communications Commission. *Journal of Law and Economics*, pp.1-40].
[193] Frischmann, supra note 182. See particularly pp. 37-42.

an infrastructure that provides access to knowledge to the public, the existence of libraries is something valuable not only to the those that actually use them but to society in a broader sense.

The discussed characteristics and their consequences will be important for understanding how libraries are regulated and, furthermore, in critically assessing the existing regulatory framework as long as it needs to adapt to a new era. As will be argued further, digitization and the building of digital collections have the potential to transform the existing institution of brick libraries into an infrastructure that becomes substantially more complex.

3.4 LIBRARIES' ROLE IN BOOK DE-COMMODIFICATION

Connected to the above analysis of the positive externalities that libraries generate, is their role in the de-commodification of books and, thus, the generation of equalizing effects; the rich can still access books at their convenience but publicly accessible libraries also allow the poor to have access. As institutions that provide public access to knowledge on the basis of objective and egalitarian criteria, libraries ultimately enable social mobility by encouraging lifelong learning. They belong to a broader 'education and culture' circle that intersects but for good reason is distinct to their related market circles: the book market, the publishing market and the market for education. To better understand their importance as reserved spaces operating outside the market, it is perhaps helpful to consider what would be the societal reaction if they were eliminated. What would the elimination of libraries or the full commodification of their available services imply? The loss of the redistributive effect that libraries have is the biggest moral cost to be calculated in case of such elimination. Given that access to knowledge and education has historically been highly unequal, the existence of libraries, especially public libraries, has been compensating at least on a moral level for such inequality. This moral compensation is important in democratic societies that claim to value equal opportunities and social mobility. A second and distinct loss would be that of cultural preservation which the book market does not offer. All the functions that libraries cover and are unique to them (meaning that they cannot be easily replaced by other institutions in the book market or in culture and education) would count as additional costs.

The moral cost is perhaps due to the particular position that books have as goods, which brings them close to a category that Guido Calabresi refers to as "merit goods". In his latest

work, the *Future of Law and Economics*, Calabresi defines merit goods as goods whose commodification is costly or whose commodification is possible but highly undesirable to a large section of society. In both cases the moral costs of placing such goods on markets are high.[194] In other words there are goods, as for example life, whose "too obvious pricing is painful to many in our society"[195] The important claim made by Calabresi is that not only commodification but also commandification of these goods is problematic. Putting a price on life is difficult but equally difficult is the acceptance that a command/structure authority, like the State, will make decisions balancing the value of peoples' lives. Besides life we feel similarly about goods, such as body organs, military service and to a certain degree education, healthcare and even environmental protection.[196] In all such cases, as Calabresi clarifies, pure market or pure demand solutions are costly. In order to get around the obvious costs, mixed market/command methods are employed. [197] For example if we look at tort law and the rules on accidents, we see that the general rules avoid putting a price on lives and safety but by choosing certain compensation schemes, they ultimately engage in indirect pricing or valuing.[198]

The pricing of merit goods is then connected to inequality. The primary reason why pricing can result in moral costs is based on an underlying understanding that wealth distribution is highly unequal, and therefore some can afford the prices of some goods while many others cannot. Calabresi presumes that the general aversion to the commodification of certain goods would be far less present in a world of equal wealth distribution.[199] One way in which society has dealt with this problem is, he explains, by modifying command structures, market structures or both.[200] Calabresi makes his point with a number of examples ranging from the financing of the

[194] Calabresi, supra note 162. Calabresi thinks that these moral costs, which are external and often indirect, are not properly captured by Coasean transactions as traditional externalities (p.28).
[195] Ibid. p.30.
[196] Ibid. p.43.
[197] Ibid, pp.35-40.
[198] Ibid, p.37. See also Calabresi, G. and Babbitt, P., 1978. *Tragic Choices: The Conflicts Society Confronts in the Allocation of Tragically Scare Resources*. W.W.Norton & Company.
[199] Ibid. p.45. He also gives what he calls a more "utilitarian" explanation: "such sales may perhaps be prohibited because the very presence of such sales tells the rest of us something about how unequal our wealth distribution is, something that we are, literally, pained to hear." The examples in this part of the text are the fact that people are not permitted to sell themselves into slavery, or forego a minimum level of education and health.
[200] Ibid. Calabresi gives the example of central decisions on who is to be drafted for war and the possibility of exceptions that allow for some to buy their way out of the draft (see p.50). He explains how variables matter: the level of local, regional or national command decision reflects not only the wealth distribution of a community but also the distribution of power. The variables will determine which commands will have the lowest external moral costs and make the decisions acceptable. The three categories also analyzed in his work with Babbitt, Tragic Choices, are: (a) system akin to rationing as used in World War II (the system of food allocation in the US during

military and the allocation of goods during wartime to the markets for body parts or the limits in political contributions during the elections. He employs these examples in order to demonstrate how certain goods are 'special' or societies treat them as such, and he invites more refined thinking on how to deal with these goods, contending:

> […] what we can do- and what we, in fact already do – to remove some goods from the prevailing wealth distribution or to attenuate the influence of this distribution on the allocation of such goods. The analysis of how best to do that, as to what goods in which societies, is exactly the work that I hope sophisticated lawyer-economists will undertake.[201]

This is where education comes into play. Basic education is generally thought of as a merit good that governments regulate by command: the first years of schooling are compulsory in most societies.[202] There is something that societies value with regard to the production and distribution of knowledge and the process of education so that it is treated differently to the production and distribution of, borrowing from Calabresi, shoes and bananas because "there are sets of goods and bads whose optimization through pure market and pure command is counterproductive, impossible even." [203] These "goods and bads" are usually both a means and an end. It is an end of itself for societies to have a standard of education. At the same time, it is a means for educated citizens to be more productive and less prone to criminal behavior, and so forth. This dual attribute is key to understanding the value that societies place on merit goods, such as education.

It is important to note that the line between market and command structures is not clear-cut. Calabresi in particular stresses "the existence and utility of the interplay of nontraditional or semi-markets and of nontraditional or decentralized command structures to the proper allocation of goods and bads."[204] Overall, he believes that we need to have a more nuanced understanding

---

the war, color tagging certain goods in the market); (b) tax or subsidy structured, (relatively) wealth distribution-neutral market; and (c) a market in which the medium of exchange is not money but some other widely held good, such as time. (see p.53)

[201] Ibid, p.70.

[202] Ibid, p. 93. In this case, according to Calabresi, the command is much more like persuasion, meaning that certain behavior is not strictly commanded but promoted.

[203] Ibid, p.102 and 115.

[204] Ibid, p. 22.

of the relationship between command and market structures within the law. This is arguably the strongest claim in his book. Calabresi encourages legal scholars to engage in law and economics analysis in the way that institutional economists have been doing for a long time.[205] Subsequently, he urges legal scholars to study the command/market interplay and take it into account when assessing the desirability of (existing) legal structures.[206] The preceding analysis aimed at setting the frame for assessing the laws applicable to books when they are in the market and when they are in libraries. As an example of goods that are being distributed both within and outside of market structures, the case of books is difficult, which is similar to several other goods that preoccupied foundational law and economics scholarship, from the allocation of goods during wartime to the allocation of radio spectrum. Allowing the co-existence of book markets and libraries, the legal system encourages the treatment of these goods as market goods and merit goods at the same time. The library's role in this equation is to guarantee a certain degree of de-commodification that is considered essential. Again, flipping the question around, the moral cost associated with the elimination of libraries and their privilege of operating outside of markets is big exactly because access to knowledge and also to education and essential educational material is still generally unequal. Thus, libraries compensate for this inequality providing access to knowledge to a broader public that might not have access to the services that come with the library infrastructure as a package and particularly to knowledge materials in the broader market – because of age, income, and social background or for any other reason. This explanation also fits well with a general appreciation of the importance of a knowledgeable society and goes back to our normative benchmark for evaluating the rationale behind the copyright norms that are applicable also to libraries: the encouragement of learning.

In addition to the above, the access the libraries guarantee is not just to knowledge, but specifically to organized knowledge – thus access to an infrastructure of systematic collection of books. One needs to distinguish between the books as merit goods in general (they can enjoy special treatment also while they are also distributed in markets, for example with subsidized pricing) and *books in libraries*, as part of library collections. The latter are special merit goods in their own right: the merit goods that this part focuses are the books inside the library, thus added

[205] Ibid, p.17 and pp. 18-21. Calabresi identified various areas to which the application of law and economics scholarship "seems to be most promising and needed." (p.21). He discusses the advantages of having lawyers making the analysis rather than economists: their results are less prone to be labeled as ideological because of the empirical nature of the analysis.
[206] Ibid, p.23.

the value of a systematic library collection. In other words, we emphasize the fact that the book is available outside the market, not standalone but as a package together with the rest of the library functions offered around it.

After the more abstract analysis of the institutional roles and functions of libraries, the next chapter concentrates on the more practical question of how can libraries work with their materials, what they can and what they cannot do with the books in their collections due to copyright restrictions. We will explore the various copyright rules applicable to libraries, and more specifically to library collections that include copyrighted works and works in the public domain. We will also explore in more detail the legal puzzles around certain categories of copyrighted works and the still problematic issue of e-lending.

CHAPTER FOUR – WHAT ARE LIBRARIES ALLOWED TO DO WITH THEIR BOOKS?

What can libraries do with the books that they hold in their collections? What are the restrictions and limitations that copyright imposes? The main point of this chapter is to give a clear picture of a framework that was conceptualized at a time when digitization was neither conceived nor advanced. Indeed, this last chapter of the first part of the thesis, which serves as a bridge to the second part, introduces the focus on digitization. As we have seen thus far, libraries are allowed to operate outside the book market; they have a public service mandate. Yet, they operate under the umbrella of copyright law both directly, benefiting from exceptions and limitations to copyright rules, and indirectly adhering to the copyright status of the books in their collections. What libraries can do with the collections they own differs in accordance to the copyright status of the books.

The first section of this chapter starts with an overview of what libraries wish to do with their collections and a flexible categorization of books based on the effects that copyright has on their availability. This categorization seeks to provide a helpful map to explain the spectrum of possibilities that libraries have with regard to their collections. We will refer to both preservation and access services that libraries wish to offer with respect to their book collections. The chapter continues with an overview of the copyright rules, limitations and exceptions, which apply to libraries. We will review international rules, and the US and EU legal framework, looking at current laws and proposed legislation. The final part in this section concentrates on the particular issue of e-lending and summarizes the main legal challenges, mostly linked to the puzzle of 'digital exhaustion.'

The subsequent sections are devoted to the two 'special' categories: (i) orphan works and (ii) out of print works, also known as out of commerce works, and (iii) to the notion of the public domain, as the ultimate state in which books are free for libraries to use – copy and distribute – in any way they choose. The latter section provides some context in order to understand the notion of public domain. It, therefore, stands in contrast to the copyrighted categories. Since copyright is meant to be limited in time, ultimately all books end up in the public domain status and become part of a larger space of intellectual commons. Libraries are still dedicated to

provide the same services for public domain works as they do for copyrighted works, unlike other market players, whose interest mostly rests in the commercial value of books. Making the switch to digitization, this part concludes touching upon the new value that is found both in public domain works and in the 'in-between' categories of orphans and out of print.

4.1 COPYRIGHT RULES APPLICABLE TO LIBRARIES

4.1.1 What do libraries want to do with their collections?

As regards their collections, the core functions of a library are, first, to build and preserve them and, second, to provide access and ensure their use. Here we basically repeat Wilkin's definition of the core functions of libraries as discussed in the previous chapter.[207] Wilkin summarized certain pillars of activities which included on the one hand curation (this for him entails *the selection, preservation, maintenance, collection and archiving of, and provision of access to materials*), and on the other the providing of services (research and learning, sometimes publishing, and the management of their spaces).[208] The question this chapter poses is: what are libraries allowed to do with their books? We will therefore focus on their curation activities and not so much on their other services. To simplify the discussion we will narrow our focus on the binary: *preservation* and *access*. Preservation will include, or presuppose the selection, maintenance, collection and archiving of the materials.

With respect to the above functions, there are both practical and legal considerations to be made. The obvious practical ones include financing and labor. Obviously a library will be restricted by its budget when deciding which books/collections to acquire and how many specialists can it engage in order to manage the rest of the preservation activities. These practical restrictions are vital but beyond the scope of this dissertation. Our concern regards the legal restrictions that libraries need to face on top of the practical ones. As the legal restrictions are imposed by copyright rules applicable to the books, what is ultimately relevant is the making and then distribution of copies: whether libraries need to have permission to, first, copy the works they own (the utmost relevant question also for digitization) and, second, to allow their patrons to browse and consult these copies – in the library premises or by lending (thus, in the digital era

[207] Wilkin, *supra* note 164.
[208] Ibid.

the issue of e-lending comes to play). Below we will look at the different copyright status that books within library collections can have, which then dictates libraries' latitude in dealing with these books. At the end of this section we will briefly touch upon the issue of e-lending and defer to a further discussion that continues under subsection 6.2.4 which analyses the landmark European e-lending case, *Vereniging Openbare Bibliotheken v. Stichting Leenrecht*.[209]

4.1.2 A flexible categorization of books based on the effects that copyright has on their availability

The two most obvious categories of books based on their copyright status are: books under copyright and books whose protection has expired and are part of the public domain. Besides the clear dichotomy between books under copyright and books in the public domain, there are arguably two additional categories that require special attention: out of print books, also referred to as out of commerce books, and the so-called orphan works. We can call them 'gray categories'. As far as the clear boundary between books under copyright and those in the public domain is concerned, the deciding factor depends on when copyright protection expires. Copyright protection has a particular duration defined by law, which as discussed elsewhere, has expanded over time. Today most jurisdictions follow at a minimum the rule set down in the Berne Convention for the Protection of Literary and Artistic Works, which is the life of the author plus fifty years after his/her death. Many jurisdictions, including the United States and countries in the EU have exceeded this minimum and provide protection for seventy years after the author's death.

All the above four categories can be viewed flexibly and represent different stages of copyright status during the lifespan of a book. These stages demonstrate the effects that copyright has on their availability. I employ the term availability in order to capture both visibility and accessibility to the general public. The less visible and accessible books are, the less available they become for multiple uses. Both copyright books and books in the public domain can suffer from invisibility and difficult, or impossible, access. This is why the categories need to remain flexible. Out of print and orphan works can essentially overlap.

[209] Case C-174/15, supra note 9.

Libraries can do whatever they want with the public domain books in their collections. The books that are in the public domain remain free for libraries to use without copyright imposing restrictions and obviously without the need for any copyright exceptions and limitations. Libraries can make as many copies of public domain works as they wish both for preservation and for access purposes. When the question of digitization comes to play, again this is the 'easy' category to handle. Libraries can without legal restrictions digitize and further provide access to the public domain corpus of their collections in whatever manner they wish to define their public. The New York Public Library, for example, has been digitizing and providing full access under their website of a series of public domain pictures and maps that it owns.[210] Greg Cram, the associate director of copyright and information policy at the NYPL has explained the copyright challenges that the project entailed.[211] Indeed, determining the public domain status of a work is not always straightforward. Such massive projects require both time and financial resources. Still, as Cram adds with a disclaimer, despite librarians' best efforts mistakes can be made: "We also want to be sure our users understand that we are not a court of law, nor are we able to guarantee that we have not made a mistake in either the facts or the law."[212] Last but not least, among the various guidelines and best practices for libraries to determine the public domain status of works in their collections we must mention a recent very comprehensive publication prepared by the Copyright Review Management System (CRMS) team at the University of Michigan library: Finding the Public Domain: *Copyright Review Management System Toolkit*.[213] Indeed the CRMS at the University of Michigan is another great

[210] In January 2016 the NYPL released more than 180,000 digitized items encouraging their sharing and reuse. The release encouraged many innovative projects. See Kimball S., Free for All: NYPL Enhances Public Domain Collections For Sharing and Reuse, *NYPL Labs*, *Spotlight to the Public Domain* (NYPL blog-posts), 5 January 2016 [online available], at https://www.nypl.org/blog/2016/01/05/share-public-domain-collections.

[211] Cram G., How we expand access to our public domain, *Spotlight to the Public Domain* (NYPL blog-posts), 17 February 2016 [online], available at https://www.nypl.org/blog/2016/02/17/rights-public-domain.
Cram explains:

> Unfortunately, investigating the copyright status of collection items is not always easy. We often are not able to make conclusive or definitive determinations. This is due, in part, to the complexity of U.S. copyright law and changes to the law over the last one hundred years. For example, the publication date of an item can be a critical fact to determine whether the item is in the public domain. But the definition of publication is not always intuitive. For example, the broadcast of a speech on national television was not considered a publication of that speech.

[212] Id.

[213] Finding the Public Domain: Copyright Review Management System Toolkit, supra note 53.

example of librarians' invaluable contribution to an accessible public domain. This is despite the many copyright challenges, which we will incrementally examine in the sections below.

In sum, determining the public domain or copyright status of works is complex. Once the status is (at least preliminarily) determined, however, there are obviously more complex questions worth exploring with regard to copyrighted works, encompassing orphan and out of print categories. The next sections (4.2 and 4.3) will examine these two categories further. Prior to this we will examine the overall copyright framework applicable to libraries, focusing on international rules, and the rules in the US and the EU.

4.1.3 International framework

The Berne Convention for the Protection of Literary and Artistic Works does not explicitly mention libraries among the limited exceptions under Articles 10 and 10bis.[214] It is rather remarkable, however, that in the most important international copyright treaty the term library is not even mentioned. One needs to refer to Article 9, and its general provision, which introduces a 'three-step test' to be applied to copyright exceptions: "It shall be a matter for legislation in the countries of the Union to permit the reproduction of such works *in certain special cases*, provided that such reproduction *does not conflict with a normal exploitation of the work* and *does not unreasonably prejudice the legitimate interests of the author*." This test is meant to provide guidance to national legislators on the exceptions and limitations provided in the convention, which are also enforceable via the treaty's incorporation into the Agreement on Trade Related Aspects of Intellectual Property (TRIPs).[215] Article 13 of TRIPs provides: "Members shall confine limitations or exceptions to exclusive rights to certain special cases which do not conflict with a normal exploitation of the work and do not unreasonably prejudice the legitimate interests of the right holder." TRIPs basically expands the same test to all exclusive rights (not only the right of reproduction) of right-holders (a term broader than the term 'authors' in the Berne Convention). Both articles are relevant to the interpretation of the national copyright exceptions that apply to libraries. We will take a look at some of these exceptions more concretely below, when we discuss the relevant provisions of the US Copyright

[214] Berne Convention for the Protection of Literary and Artistic Works, supra note 154.
[215] TRIPS: Agreement on Trade-Related Aspects of Intellectual Property Rights, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1C, 1869 U.N.T.S. 299, 33 I.L.M. 1197 (1994) [hereinafter TRIPS Agreement].

Act and the EU Copyright Directive. Last but not least, international norms impose a duty on libraries to maintain anti-circumvention protection.[216]

There have been discussions at the international level about an agreement addressing exceptions applicable to libraries. A number of stakeholders, led by the International Federation of Library Associations and Institutions (IFLA), have even proposed a treaty to WIPO for exceptions applicable to libraries and archives.[217] The stakeholders flag the need for a clearer international framework that guarantees the following:

i. Preservation of materials for posterity, with the flexibility to access cultural works in copy-protected formats;
ii. Support of education, research, and private study;
iii. Making or receiving copies of works lawfully acquired by a library or archive for personal and private purposes;
iv. Supply of copies of works in response to requests from individual users;
v. Provision or lending of lawfully acquired content on a not-for-profit basis;
vi. Support of people with disabilities to exercise their right to access content;
vii. Circumvention of TPM for the purpose of permitting a non-infringing use of a work;
viii. Limit on the risk of liability to libraries and archives with respect to orphan works.[218]

Until now, there has been no legislative follow-up on the treaty proposals. Further developments, however, are reasonably expected especially after the adoption, under WIPO, of the Marrakesh Treaty which aims to facilitate access to published works for persons that are blind, visually impaired, or otherwise disabled.[219] With this treaty a new international exception to copyright rules has been construed for the benefit of a specific group and its access to

[216] WIPO Copyright Treaty, December 20, 1996, article 11.
[217] IFLA continues pushing the agenda for an international treaty under WIPO. See https://www.ifla.org/copyright-tlib. The treaty proposal came from IFLA in collaboration with the International Council on Archives (ICA), Electronic Information for Libraries (EIFL) and Corporación Innovarte.
[218] Quoted from: The Case for a Treaty on Exceptions and Limitations for Libraries and Archives: Background Paper by IFLA, ICA, EIFL and INNOVARTE, WIPO Doc. SCCR/23/3 (Nov. 15, 2011).
[219] Marrakesh Treaty to Facilitate Access to Published Works for Persons Who are Blind, Visually Impaired, or Otherwise Disabled, supra note 11.

materials. This new exception can also benefit libraries that fall under the definition of an "authorized entity", which is "an entity that is authorized or recognized by the government to provide education, instructional training, adaptive reading or information access to beneficiary persons on a non-profit basis." In this case libraries that are part of institutions dedicated to education, training and access for the blind, the visually impaired and the otherwise print disabled are able to benefit from the new exception.

4.1.4 US framework

In the United States, the most relevant provisions applicable to libraries are Sections 107 and 108 of the Copyright Act (17 U.S.C.). Section 107 provides for the famous fair use exception on the basis of four factors, which are:

> […] the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include— (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for or value of the copyrighted work.

Libraries can benefit from this fair use exception even if there is another provision designated to provide a special exception for them in Section 108. This has been clarified by Judge Bear in his District Court decision for the *Authors Guild Inc. v. HathiTrust* case; libraries are basically able to invoke the fair use defense in Section 107 in addition to the special library exceptions in section 108.[220] As chapter six of the thesis includes a wider analysis of this case,

[220] *Authors Guild, Inc. v. HathiTrust– 2012 District Court Opinion*, supra note 5.

we will defer our discussion on how the court applied the fair use doctrine to the benefit of HathiTrust Digital Library. Notably, Section 107 is also important because it can provide an exception from copyright not only to libraries but also to patrons who use library works.

Section 108 of the Copyright Act provides for special exceptions for libraries and archives. The exceptions are quite tightly defined. From a contemporary perspective, the section reads like a rather old text concerned about older methods of copying. It allows employees of libraries and archives "acting within the scope of their employment, to reproduce no more than one copy or phonorecord of a work (…) without any purpose of direct or indirect commercial advantage (…)." There are a number of additional conditions set down to allow this one copy, including the collections of the library or archive should be open to the public, available to researchers in general and also that the reproduction or distribution includes a copyright notice. There is also a provision allowing for three copies to be reproduced or distributed but this applies to unpublished works and "solely for purposes of preservation and security or for deposit of research use in other library or archives." Moreover, digital reproductions are not to be distributed in digital format or made in any way available outside the premises of the library or archive. The three copies exception further applies to works that are damaged, deteriorating, lost or stolen, or existing in a format that has become obsolete. Again, the requirements around the exception are quite strict, including a "reasonable effort, determined that an unused replacement cannot be obtained at a fair price." If the user makes a request and under certain requirements (mostly, for purposes of private study, scholarship or research), Section 108 allows "no more than one article or other contribution to a copyrighted collection or periodical issue, or to a copy or phonorecord of a small part of any other copyrighted work." The law, however, does not impose liability on librarians who have no reason to believe that the use of the copied work will be other than private study, scholarship or research.

The US Copyright Office has acknowledged how some of these rules are outdated and need review, especially after the mass digitization projects and the corresponding case-law that they triggered. Since June 2016 there has actually been an official inquiry calling interested parties to discuss how to revise Section 108 "with a goal to updating the provisions to better reflect the facts, practices, and principles of the digital age and to provide greater clarity for

libraries, archives, and museums."[221] The Copyright Office recognizes that Section 108, a legislative product of 1976, was intended to address the use of print-based analog technology at the time and that the current technological developments were not anticipated. In the words of the notice of inquiry "the exceptions in section 108 therefore are stuck in time." [222] The office even appointed an independent study group, the so-called "Section 108 Study Group" to work on the issue. In its report issued in 2008, the group pointed to the changes in needs and expectations of users regarding the reproduction and distribution of copies of library material.[223] In its notice of inquiry, the Copyright Office stated that it agrees in principle to the need to change in the following ways:

i. Adding museums as eligible institutions;
ii. Expanding the preservation, security, and deposit for research exceptions to include published/publicly disseminated works;
iii. Creating a new exception to permit the reproduction and distribution of publicly available internet content for preservation and research purposes, with an opt-out provision;
iv. Allowing the outsourcing of certain section 108 activities to third-party contractors;
v. Removing or revising the three-copy limitation for preservation and security, deposit for research, and replacement copies.[224]

4.1.5 EU framework

At the EU level, there is a list of exceptions in the EU Copyright Directive (2001/29/EC).[225] Point 34 of the preamble puts libraries in first place: "Member States should be given the option

[221] Library of Congress & Copyright Office Notice of Inquiry, Section 108: Draft Revision of the Library and Archives Exceptions in U.S. Copyright Law, Docket No. 2016 – 4, available at https://www.gpo.gov/fdsys/pkg/FR-2016-06-07/pdf/2016-13426.pdf.
[222] Ibid.
[223] The Section 108 Study Group report is online available at http://www.section108.gov/docs/Sec108StudyGroupReport.pdf.
[224] Ibid. The report is also explicitly recognizing a "widely known" lack of clarity and problematic organization in the section 108 provisions and prefers redrafting to re-organizing the section.
[225] Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the Harmonisation of Certain Aspects of Copyright and Related Rights in the Information Society. OJ L 167, 22/06/2001 P. 0010 – 0019.

of providing for certain exceptions or limitations for cases such as educational and scientific purposes, for the benefit of public institutions such as libraries and archives, for purposes of news reporting, for quotations, for use by people with disabilities, for public security uses and for uses in administrative and judicial proceedings." Libraries, therefore, fall under the radar of the Directive.

However, point 40 further specifies exceptions or limitations for the benefit of non-profit establishments, such as publically accessible libraries "and equivalent institutions" and archives, which should be limited to "certain special cases covered by the reproduction right" and also "should not cover uses made in the context of on-line delivery of protected works or other subject-matter." We can see firsthand how the Directive does not cover online delivery, which is significant in the digital era. The same point also mentions the exclusive public lending right adopted in Article 5 of the Rental and Lending Rights Directive which is relevant once we see the library-related provisions in the body of the Copyright Directive.[226] Article 5 of the EU Copyright Directive provides for a list of acceptable and optional exceptions and limitations to the right of reproduction. The relevant exception is found in paragraph 2(c): "Member States may provide for exceptions or limitations to the reproduction right […] (c) in respect of specific acts of reproduction made by publicly accessible libraries, educational establishments or museums, or by archives, which are not for direct or indirect economic or commercial advantage." The Directive does not impose uniform rules to the Member States. Until now, we face a situation where the copyright exceptions and limitations for libraries vary across the EU.

In the context of litigation before the Court of Justice of the European Union, the above-mentioned articles in the EU Directive have been judicially interpreted in cases involving libraries. The second part of the dissertation will discuss in detail the relevant cases and the most recent interpretation of these rules. The cases are: (i) *Technische Universität Darmstadt v. Eugen Ulmer KG*,[227] a case involving a university library in Germany that allowed the consultation of electronic versions of its books at specified terminals on the premises of the library; (ii) *Vereniging Openbare Bibliotheken v. Stichting Leenrecht*,[228] the case concerns the association of public libraries in the Netherlands seeking recognition of the libraries' right to e-lend; and (iii)

[226] Directive 2006/115/EC of the European Parliament and of the Council of 12 December 2006 on rental right and lending right and on certain rights related to copyright in the field of intellectual property (codified version)
[227] *Technische Universität Darmstadt v. Eugen Ulmer KG*, supra note 8.
[228] *Vereniging Openbare Bibliotheken v. Stichting Leenrecht*, supra note 9.

Marc Soulier & Sara Doke v. Premier ministre & Ministre de la Culture et de la Communication,[229] which discussed a new French law in France allowing cultural institutions to proceed in certain uses of out of commerce works. As we will shortly see in our analysis of the case-law, libraries encounter difficulties when applying the current copyright rules to their digital needs. Thus, the current framework does not offer straightforward rules for e-lending or digitization on mass or small scale. In the first case, the letter of the rental Directive does not cover electronic lending. In the second scenario, while most jurisdictions have exceptions that allow for copying for preservation purposes (especially of rare or degrading documents), they would not have much to say about small or mass digitization for any other purposes, such as indexing, creating searchable databases and so forth.

Despite these general observations, there have been some concrete attempts at a top-down level to create new rules that would facilitate needs, such as digitization. At the supranational level this is done with EU legislation, usually Directives, wishing to guide Member States in their legislative reforms. As already mentioned, a specific exception applicable to orphan works is now in place, following the introduction of the Orphan Works Directive.[230] Let us note for now that the Directive basically allows Member States to make uses of orphan works further to what is allowed for any other copyrighted work. This means that Member States can enact legislation allowing institutions like libraries to make such uses under certain conditions. According to Article 6 of the Directive these uses include their availability to the public and also acts of reproduction "for the purposes of digitisation, making available, indexing, cataloguing, preservation or restoration." Among the cultural heritage institutions, the scope of the Directive covers libraries, together with archives, educational establishments and museums.

The Orphan Works Directive has been the first complete legislative act within the EU to modernize rules applicable to libraries, which has taken into account the inefficiencies arising in certain types of copyrighted works. We will discuss the inefficiency around orphan works at length in the following section of the chapter. More importantly, there is an ongoing effort to review copyright rules even further.[231] Relevant to libraries, the latest legislative proposals touch upon the issue of works out of commerce, text and data mining and access to the blind and the

[229] *Marc Soulier & Sara Doke v. Premier ministre & Ministre de la Culture et de la Communication*, supra note 10.
[230] Directive 2012/28/EU of the European Parliament and of the Council of 25 October 2012 on Certain Permitted Uses of Orphan Works, OJ L 299, 27.10.2012.
[231] The new legislative proposals are available at the Commission's Digital Single Market webpage: https://ec.europa.eu/digital-single-market/en/modernisation-eu-copyright-rules.

otherwise print disabled. Of further relevance to libraries is the new copyright Directive proposed and the legislative proposals to pass the Marrakesh treaty into EU law: (i) Proposal for a Directive of the European Parliament and of the Council on Copyright in the Digital Single Market[232] and (ii) Proposal for a Directive of the European Parliament and of the Council on Certain Permitted Uses of Works and Oher Subject-matter Protected by Copyright and Related Rights for the Benefit of Persons who are Blind, Visually Impaired or Otherwise Print Disabled and amendment to Directive 2001/29/EC for the Harmonisation of Certain Aspects of Copyright and Related Rights in the Information Society.[233]

Reading the preamble and explanations accompanying the proposals, it is clear that the Commission's objective is to assist cultural institutions, including libraries, in their contemporary needs and functions. Because the old system of exceptions and limitations is restricted and does not properly guide Member States in implementing rules on how libraries and other institutions are to function in the digital era, it explicitly acknowledges:

> As a result of a modernised framework of exceptions and limitations, researchers will benefit from a clearer legal space to use innovative text and data mining research tools, teachers and students will be able to take full advantage of digital technologies at all levels of education and cultural heritage institutions (i.e. publicly accessible libraries or museums, archives or film or audio heritage institutions) will be supported in their efforts to preserve the cultural heritage, to the ultimate advantage of EU citizens.[234]

We will briefly refer to the text of the new proposed rules, which offers us important insight. If the proposals enter into the EU copyright law, it is true that the framework applicable

---

[232] COM/2016/0593 final - 2016/0280 (COD) (14.9.2016), text available at http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016PC0593.

[233] COM/2016/0596 final - 2016/0278 (COD) (14.9.2016), text available at http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016PC0596 which is the proposed implementation of the Marrakesh treaty combined with one final proposal for a regulation:
Proposal for a Regulation of the European Parliament and of the Council on the cross-border exchange between the Union and third countries of accessible format copies of certain works and other subject-matter protected by copyright and related rights for the benefit of persons who are blind, visually impaired or otherwise print disabled, COM/2016/0595 final - 2016/0279 (COD) (14.9.2016), text available at http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016PC0595.

[234] Explanatory memorandum accompanying the Proposal for a Directive of the European Parliament and of the Council on Copyright in the Digital Single Market (supra note 12).

to libraries becomes much more elaborate and in many ways supports digitization and e-lending. The proposed Directive is applicable to libraries as part of a broader category of cultural institutions. Article 2(3) defines cultural heritage institutions as "a publicly accessible library or museum, an archive or a film or audio heritage institution." The relevant provisions for libraries are the proposed Article 5 which authorizes reproduction for the preservation of cultural heritage "permitting cultural heritage institutions, to make copies of any works or other subject-matter that are permanently in their collections, in any format or medium, for the sole purpose of the preservation of such works or other subject-matter and to the extent necessary for such preservation." This is an obviously broader exception than the one currently in place under the existing Directive. Still the provision is generous only with respect to preservation, but not access. Libraries would immediately benefit from this exception because they could digitize and make back-ups of their entire permanent collections. Patrons would also indirectly benefit even if their access is limited due to copyright restrictions; first, so long as they consult the works in the premises of the libraries, and, second, they would have potentially digital access when the works do enter the public domain. Second, Article 7 of the proposed Directive deals with the use of out of commerce works by cultural institutions. This article calls on Member States to allow collective management organizations to conclude non-exclusive licenses with cultural heritage institutions for non-commercial purposes "for the digitization, distribution, communication to the public or making available of out-of-commerce works or other subject-matter permanently in the collection of the institution." It provides for a number of requirements around such licensing schemes including, for example, that the collective management organization is representative of right-holders in the relevant category of works, or that right-holders may at any point object to their works being under the licensing scheme.

The entire reform package of new legislation is extensive and more complex, covering more issues that are less directly relevant to libraries and, thus, will not be discussed in the dissertation. It is still relevant to follow the progress of the proposals and at a later stage the implementation at the Member State level. The difficulty that the Directive-based, half-baked harmonization process poses to copyright rules across the EU is a larger phenomenon that

necessarily affects libraries, which still operate under specific national rules while 'guided' by EU directions.[235]

Summing up the current applicable rules, one can see that the flexibility given to libraries in managing works in their collections remains fairly limited since the exceptions that allow libraries to do what they wish with their copyrighted works are tightly construed. Pending the above-mentioned reform plans at the international level and also the two broader jurisdictions of the US and the EU that we studied, today's libraries have to adjust to the changing realities and needs of their patrons, while operating under the existing legal framework. For our purposes, we see how this legal framework, at least on its face, does not allow libraries to engage in mass or small scale digitization projects unless it solely targets public domain works, or on the EU level orphans works are covered to a certain extent. Policy-makers have explicitly acknowledged the need for reform but reform measures are slow. One of the main reasons why the discussions around reforms of rules applicable to libraries are slow and not straightforward is, I think, due to the differences among different types of books and their corresponding legal statuses. Permitting copying and some form of distribution, which is an exception to the default copyright protection, has potentially diverse implications and depends on whether the book in question is scarce, or commercially relevant to its right-holders, or an academic work or novel and so on.

The most common form of distribution that libraries engage in is lending. Before moving to the next sections, let us briefly mention here the challenges of e-lending as the legal framework currently stands.

4.1.6 Library e-lending

The issue of electronic lending of books (e-lending) remains a big puzzle for copyright mostly in view of the general lack of clarity over digital exhaustion.[236] The *in rem* right over a physical book is reflected in the copyright rule of exhaustion, or the first sale doctrine. This rule basically means that, once a copy is sold, the right of the copyright-holder over that particular copy is

[235] Some Member States have already expressed several criticisms of the new legislative package and, as the Commission proposals are still under review in the European Parliament, numerous amendment proposals are being circulated. For an idea on how the debate formed during the past year in the Parliament see the related opinions written in The Parliament Magazine including: Stihler C., EU copyright rules holding back cultural institutions, *The Parliament Magazine*, 12 July 2016 [online available], at https://www.theparliamentmagazine.eu/articles/opinion/eu-copyright-rules-holding-back-cultural-institutions.

[236] See among many Rosati, E. 2015. Online copyright exhaustion in a post-All posters world. *Journal of Intellectual Property Law & Practice*, p.122.

exhausted. On this basis, books can be redistributed on second-hand markets. The doctrine is also the legal basis for the lending of books from libraries in the American jurisdiction (Section 109(a) of the US Copyright Act). In Europe the principle of exhaustion is a long-standing rule found in copyright statutes in several jurisdictions and has been established by the Court of Justice of the European Union and codified in Article 4(2) of the Information Society Directive of 2001. There is also a territorial effect in the entire EU jurisdiction as the first sale of a work in one EU Member State will exhaust the distribution right of the author in all other Member States.

One can also look at Article 6 of the WIPO Copyright Treaty which provides for the right of distribution and states that the parties of the treaty shall determine the conditions under which the exhaustion of the distribution right applies "after the first sale or other transfer of ownership of the original or a copy of the work with the authorization of the author."[237] In other words, the distribution right is limited by exhaustion. Once copies of works have been placed in the market with the right-holder's consent, further re-distribution (as is for example lending) does not require authorization. Traditionally, once a library purchases a book from a publisher, the exhaustion or the first sale doctrine releases the copy from further copyright control. At the moment, however, copyright law worldwide explicitly only grants exhaustion to tangible objects, such as printed books.[238] According to the Agreed Statements concerning the WIPO Copyright Treaty, in particular the statements regarding Articles 6 and 7 and subject to the right of distribution and the right of rental under these articles, the expressions "copies" and "original and copies" refer exclusively to *fixed copies* that can be put into circulation as *tangible objects*.[239]

In sum, both at national and international level the application of the exhaustion principle to digital works, including e-books, is still unclear.[240] Recent case-law has brought the issue of

[237] WIPO Copyright Treaty, December 20, 1996, article 6.
[238] See Müller, supra note 32, p. 152.
[239] Agreed Statements concerning the WIPO Copyright Treaty, adopted by the Diplomatic Conference on December 20, 1996, online available at http://www.wipo.int/treaties/en/text.jsp?file_id=295456.
[240] For the debate in the US see primarily Jenkins, J., 2014. Last sale? Libraries' rights in the digital age. *College & Research Libraries News*, *75*(2), pp.69-75 and Chiarizio, An American tragedy: e-books, licenses, and the end of public lending libraries, supra note 32. See also Reese, R.A., 2002. First Sale Doctrine in the Era of Digital Networks, The. *BCL Rev.*, *44*, p.577. Reese emphasizes the importance of affordability and availability of works in libraries and second-hand markets that exhaustion guarantees (pp. 644- 652). From the European side see mostly: Müller, Legal aspects of e-books and interlibrary loan, supra note 32; Afori, The Battle Over Public E-Libraries–Taking Stock and Moving Ahead, supra note 32; Dusollier, A manifesto for an e-lending limitation in copyright, supra note 32; Matulionyte, E-Lending and a Public Lending Right: Is it Really a Time for an Update?, supra note 32; van der Noll at all, *Online uitlenen van e-books door bibliotheken*, supra note 48; Last, from the perspective of

digital exhaustion to the forefront once again as illustrated in the case against ReDigi, the online market-place for used music files. The Capitol Records filed the case in 2013 and the District Court ruled against ReDigi and in favor of Capitol's claim that ReDigi infringed its distribution and reproduction rights afforded by copyright by permitting the reselling of used music files even though the seller's copy was deleted.[241] The main reasoning behind the decision of the Court of First Instance was that while ReDigi requested a reading of the first sale doctrine in the Copyright Act (Section 109a) that adjusts to the technological changes since the enactment of the Act, it is for the legislator to decide whether an amendment is needed and not for the court. Pending the appellate decision before the Second Circuit, we have to wait and see if the judgment will be upheld and how digital exhaustion will ultimately play out in the US.[242] The decision will be highly relevant to libraries. The *amicus curiae* submitted to the court of appeals by the American Library Association, the Association of College and Research Libraries, the Association of Research Libraries, and the Internet Archive supports the reversal of the previous decision. They argue that upholding ReDigi's fair use claim would encourage innovative digital lending services in libraries.[243] Interestingly, these library associations together with the Internet Archive cited the HathiTrust and Google Books decisions[244] as examples of fair use adjudications that provide "necessary breathing space in copyright law, making sure that control of the right to copy and distribute does not also become control of the right to create and innovate".[245]

In Europe the equivalent case is *UsedSoft GmbH v. Oracle International Corp*.[246] This was a similar case that involved an online market-place for used licenses for computer programs. Interestingly, the Court of Justice of the European Union came to the opposite conclusion of the ReDigi decision and ruled in favor of digital exhaustion in the case of software.[247] There is a new

---

IFLA see "IFLA Principles for Library eLending" (revised August 16, 2013), online available at: http://www.ifla.org/files/assets/hq/topics/e-lending/principles-for-library-elending-rev-aug-2013.pdf

[241] Capitol Records LLC v. ReDigi Inc., 934 F.Supp.2d 640 (2013).

[242] Capitol Records LLC v. ReDigi Inc., case number 1:12-cv-00095, in the U.S. District Court for the Southern District of New York

[243] Brief of Amici Curiae American Library Association, Association of College and Research Libraries, Association of Research Libraries, and Internet Archive, *Capitol Records LLC v. ReDigi Inc.*

[244] *Authors Guild v. Google* and *Authors Guild v. HathiTrust*, supra notes 4 and 5.

[245] Ibid, p. 16.

[246] Case C-128/11, ECLI:EU:C:2012:407.

[247] Furthermore, the European court has confirmed its view upholding digital exhaustion in the most recent Ranks/Microsoft case [Case C-166/15, ECLI:EU:C:2016:762]. In this case the resale of back-up copies of software was found to circumvent the exhaustion principle.

preliminary reference in the pipeline to be sent to the CJEU now and, this time it concerns the resale of e-books. This case comes from the Amsterdam Court of Appeals in the Netherlands which had decided that the reuse of used e-books was permitted: *NUV v. Tom Kabinet*.[248] To reiterate how confusing the situation currently stands, there are conflicting decisions at the national courts in Europe. While the Amsterdam court accepted the reuse of used e-books, German courts have denied the reuse/resell of audiobooks.[249]

Moving on with the relevant to libraries copyright rules, in the following sections of this chapter we will examine in more detail the particular concerns that arise with respect to the so-called orphan works and with the works out of print (sections 4.2 and 4.3). At the beginning of this section we tried to categorize books on the basis of the effects that copyright has on their availability, and attempted to transcend the standard binary categorization of books under copyright and books in the public domain. Having in mind the potentials that digitization opens up (this will be the immediate focus of the next part of the thesis) we will investigate these categories further. Even though orphan works and works out of print were always part of libraries' collections, the digitization potentials made the categories most relevant in practice and recently much more discussed. This, I will argue, is due to the additional value that digitization creates for these works – a value that is most apparent also in the case of the digitization of the public domain (section 4.4).

4.2 THE PUZZLE OF ORPHAN WORKS

Fay Kanin, the Chair of the Library of Congress National Film Preservation Board (NFPB), coined the term orphan works to comprehensively describe works protected under copyright when the copyright-holder cannot be identified or located. According to librarians, there is a substantial amount of orphan works for which it is difficult to find right-holders' information even after extensive research. However, knowing the exact size of the problem is important in order to be able to calculate the social and economic costs and benefits to finding

[248] *NUV v. Tom Kabinet*, Court of Appeal of Amsterdam, ECLI:NL:GHAMS:2015:66, (20 January 2015).
[249] On the German case (first instance and appellate level) see Rosati E., Waiting for a lower court to rein in resale? You'd sooner herd cats, *IP Kat Blog*, May 1, 2013, available at http://ipkitten.blogspot.gr/2013/05/waiting-for-lower-court-to-reign-in.html and Bellan A., Copyright exhaustion does not apply to digital goods other than software, Hamm Court of Appeal says, *IP Kat Blog*, June 13, 2014, available at http://ipkitten.blogspot.gr/2014/06/breaking-copyright-exhaustion-does-not.html.

possible solutions to the problem.[250] The very nature of orphan works renders the task of finding concrete quantitative and qualitative data difficult. This also explains why the size of the problem has not been calculated in a consistent manner. Besides the various numerical estimations,[251] it is necessary to look at the severity of the problem more generally. This is an important question since the value of these works matter for understanding the value of access to them. Among several studies, Hansen cites a Cornell University Library study indicating that library staff searching for right-holders of 343 monographs identified as still under copyright, spent (in time) over $50,000 and were still unable to identify 58% of the owners of these

[250] See also the JISC 2009 report, analyzing data from an online survey of over 500 organizations suggesting that many public sector organizations in the UK are themselves unsure of the extent of the problem. The report is available at: http://www.jisc.ac.uk/media/documents/publications/infromthecoldv1.pdf.

[251] According to librarians, there is a great amount of orphan works for which it is estimated that even after extensive research, no further information can be found. However, knowing the exact size of the problem is important in order to be able to calculate the social and economic costs and benefits of possible solutions to the problem. See for example the JISC 2009 report, analyzing data from an online survey of over 500 organizations suggesting that many public sector organizations in the UK are themselves unsure of the extent of the problem. The report is available at: http://www.jisc.ac.uk/media/documents/publications/infromthecoldv1.pdf.

- For example, the British Library has estimated that 40 % of its in-copyright collections are orphan. See: *The New Renaissance* - Report of the 'Comité des Sages' reflection group on bringing Europe's cultural heritage online (January 10, 2011), supra note 18, p. 16.

- David Drummond, Google's general counsel estimated that relatively few, under 20%, of the books in the Google Books corpus will ultimately turn out to be orphan (also relying on his positive predictions for Google's project incentivizing copyright owners to come forward). See: Samuelson, P., 2009. Google Book Search and the future of Books in Cyberspace. *Minn. L. Rev.*, *94*, p.1308, at p. 1323 citing to the Competition and Commerce in Digital Books hearing before the House of Representatives, available at: http://judiciary.house.gov/hearings/hear_090910.html.

- At the same time for the same project Jonathan Band estimated that around 75% of out-of-print books will remain unclaimed. See: Band, J., 2009. The long and winding road to the Google Books Settlement. *J. Marshall Rev. Intell. Prop. L.*, *9*, p.i.

David Hansen, who wrote the first white paper for the Berkeley Digital Library Copyright Project, poses the same question concerning the size of the orphan works problem (Hansen, Orphan Works: Definitional Issues, supra note 49).

He looks to measure the size of the problem, in terms of (a) the number of works considered as orphans and (b) the value of these works. According to his calculations, the problem is known to be significant but depending on the type of collection considered, estimates range in number from hundreds of thousands to millions orphans. Citing several studies, Hansen sees that for books, best estimates indicate up to 50% of twentieth century publications. (p.9)

This is affirmed also by John Wilkin's estimation after reviewing HathiTrust project holdings in 2011. John Wilkin reviewing HathiTrust project holdings estimates that roughly 50% of the volumes in the HathiTrust collection of 5 million are likely to be orphans, explaining that 12,6% will come from years 1923-1963, 13,6% from 1964-1977, and 23,8% from 1978 on. This amounts to 800.000 US orphans and nearly 2 million non-US orphans in the HathiTrust collection. See: http://www.clir.org/pubs/ruminations/01wilkin.

In a 2010 report for the European Commission, Anna Vuopala estimates orphan works range from 13% to up to 70% for certain collections. Anna Vuopala, Assessment of the orphan works issue and costs for rights clearance (2010, report for the European Commission), available at http://ec.europa.eu/information_society/activities/digital_libraries/doc/reports_orphan/anna_report.pdf. Beyond books estimates seem even more difficult.

works.[252] Overall, Hansen concludes that existing research, although showing a large number of orphans does not truly capture the size of the problem: "[existing] efforts to quantify the number of orphan works are a useful starting point, but inadequate to inform more particularized legal solutions."[253] For some, the number is almost incalculable. Before legal formalities were abolished in copyright law, owners in most cases only bothered to comply with the law for works that were commercially viable. After the abolition of formalities, copyright protection triggered automatically upon fixation to a medium and the problem came to be seen as structural. The extension of the term of copyright protection and the rapid growth of unpublished and personal works created and stored in digital form results in an untraceable amount of works that are, or can be, protected by copyright and whose owners are difficult or impossible to locate. Thus, for most scholars the root of the problem of orphan works, rendering the quest for a definite solution difficult, is found in the expansions of copyright law in the past few decades; extensions of the duration of copyright coupled with the elimination of registration, renewal and notice requirement for copyright protection (these results are also due to the Berne Convention rules). [254] There are also alternative accounts on why solutions to the orphan works problem are difficult. A good example is the Lydia Pallas Loren's explanation based on the misleading use of the orphans metaphor, which overemphasizes a romantic view of the author.[255]

---

[252] Response by the Cornell University Library to the Notice of Inquiry Concerning Orphan Works, 70 Fed. Reg. 3739 January 26, 2005, Comment OW0569, available at http://www.copyright.gov/orphan/comments/OW0569-Thomas.pdf.

[253] Hansen, Orphan Works: Definitional Issues supra note 49, p.11.

[254] As simply explained by Olive Huang, longer copyright terms create longer periods over which copyright owners can change hands and become even more difficult to trace: Huang, O., 2006. US copyright office orphan works inquiry: Finding homes for the orphans. *Berkeley Technology Law Journal*, pp.265-288, at p. 268. See also: Hansen D. 2012. Orphan Works: Causes of the problem, supra note 49. According to Hansen the orphan works problem may have existed in theory since first copyright laws came into effect, if one defines the problem broadly as the situation where the owner of a copyrighted work cannot be located and asked for authorization by someone who wants to use it. Besides the aforementioned reasons, the elimination of copyright formalities and the progressive extension of copyright terms, Hansen in his third white paper for the Berkeley Digital Library Copyright Project adds: the technological advances that allow authors to create – and preserve – more copyrighted works (thus changes in both information creation and storage practices) and technological changes in the way users access and consume copyrighted works (thus changes in formats for information consumption), especially in the shift from print to digital. One could say that the added two reasons made the problem much more apparent in the digital age than ever before.

[255] Loren, L.P., 2012. Abandoning the Orphans: An Open Access Approach to Hostage Works. *Berkeley Technology Law Journal*, pp.1431-1470. Loren proposes a change in terminology, and suggests we should call "orphans" "hostage" works instead. She proposes that law should give them immunity from copyright claims so long as they provide the "hostage" works for free and also have a mechanism to return to respecting the author's work if they appear.

Legislators have realized this particular category of works is problematic because of the inefficiency in approaching right-holders to obtain the consent to use the book and in accessing copies as they become rare. Notably, for our purposes, the category is particularly problematic for libraries who estimate that large portions of the books in their collections are orphans. If librarians wish to make use of such works they face the same obstacle: they cannot trace right-holders to obtain their consent for actions that are not clearly covered by the exceptions in place under the various current copyright rules (as seen in the section above); when they wish, for example, to make more copies than the limited number allowed, or digitize the works and include their corpus in searchable catalogues, or, further, display them in their e-portals and e-lend them.

There have been attempts to deal with both the orphan works and the problem of out of print works with soft law approaches, usually best practices, and legislation. Legislative efforts have been more successful in Europe, especially at the EU level, compared to the US. Neither jurisdiction has managed to fully tackle the issue, however. Starting with the US there have been two unsuccessful attempts to legislate orphan works: first, with the Orphan Works Act of 2006.[256] Later two other bills were introduced: the Orphan Works Act of 2008,[257] and the Shawn Bentley Orphan Works Act of 2008.[258] These legislative proposals would have introduced a statute limiting the remedies available against defendants that in good faith and after diligent search did not locate the right-holder to request permission to use orphans. The Copyright Office has further indicated that there will be another attempt to legislate this issue.[259]

In the EU, as mentioned also above, there have been concrete policy developments following the adoption of the Orphan Works Directive 2012/28/EU in October 2012.[260] This Directive 'on certain permitted uses of orphan works' sets out common rules for the digitization

[256] H.R. 5439, 109th Congress, 2nd session, 22 May 2006.
[257] H.R. 5889, 110th Congress, 2nd session, 24 April 2008.
[258] S. 2913, 110th Congress, 2nd session, 24 April 2008.
[259] It issued a broad notice of inquiry in the Federal Register, seeking comments from the public regarding the current state of play for orphan works (See at: http://www.copyright.gov/orphan/). The notice of Inquiry was issued on 22 October 2012 and invited comments on the state of play for orphans. The initial comments have now been published (91 stakeholders have sent their comments). The reply-comment submission period ran until the 6th of March 2013, and is now also completed (89 reply-comments available). From the comments it becomes evident that there is a heated debate between the stakeholders. Collecting societies seem to be taking the lead against orphan works legislation, while the academic world together with libraries (comments from librarians, associations of libraries and University libraries) recognize a real problem that needs comprehensive solution.
For a more extensive analysis on the orphan works problem as faced in the United States, see the most recent report prepared by David Hansen: Hansen, supra note 51.
[260] Available at: http://ec.europa.eu/internal_market/copyright/orphan_works/index_en.htm.

and online display of orphan works.[261] The Directive only applies to works that are first published, whether broadcasted or made publicly available by the beneficiaries, in an EU Member State. It concerns only uses of orphan works by publicly accessible libraries, educational institutions and museums, archives, film or audio heritage institutions and public-service broadcasting organizations. Thus, the Directive is not applicable to non-profit organizations. In sum, the solutions that the Directive provides are relatively narrow in scope since they apply to a limited class of users, a small number of uses and particular types of works. In order to establish whether a work is orphan these institutions are to carry out a 'diligent search' according to the requirements of Article 3 of the Directive, keeping records of their searches on a publicly accessible online database. What constitutes a 'diligent search' is outlined in more detail in the Memorandum of Understanding on diligent search guidelines for orphan works.[262] Once designated as orphan, it may be used (digitized or made available) by the institutions "only in order to achieve aims related to their public-interest missions, in particular the preservation of, the restoration of, and the provision of cultural and educational access to works and phonograms contained in their collection".[263] The Directive provides for a system of compensation if the right-holder(s) is found at a later stage.[264] Due to the narrow scope of the Directive, which leaves ample room for different interpretations, the implementation of the Directive at the member state level has result in *de jure* or *de facto* conflicting approaches.[265] Also, some member states have been slower to participate in the debate at a national level than others.[266]

Starting with the premise that the owner, be it an author or a subsequent right-holder, is absent, an overall question lingers with respect to orphan works: Why has it proven so difficult to introduce reform in the field of property law where owners of works are absent (thus by

[261] Directive preamble, point 3. The Directive is complementing and without prejudice to the existing 20 September 2011 Memorandum of Understanding on key principles on the digitization and making available of out-of-commerce works: Memo available at http://europa.eu/rapid/press-release_MEMO-11-619_en.htm.
[262]Directive 2012/28/EU of the European Parliament and of the Council of 25 October 2012 on Certain Permitted Uses of Orphan Works, OJ L 299, 27.10.2012.
[263] Article 6(2) of the Directive.
[264] Article 5 of the Directive.
[265] We will have a closer look at the contradictions when looking at the European litigation under chapter six.
[266] Relevant ongoing work is done through the collaborative project *EnDOW* ("Enhancing access to 20th Century cultural heritage through Distributed Orphan Works clearance") which was launched by agencies of eighteen countries and the European Commission as part of the 'Joint Programming Initiative in Cultural Heritage and Global Change'. The project focuses on right clearance for cultural institutions engaging in digitization and specifically on determining what would constitute diligent search to satisfy the Directive's requirements. See the projects website with thus far available resources available at http://diligentsearch.eu/resources/.

definition, they are unable to lobby themselves), while users of works lobby for reform?[267] Libraries are perhaps the most relevant stakeholders representing users, themselves and their patrons, who wish to benefit from a better and clearer regulatory framework in this matter. As we will see in the analysis of the Google Books and predominantly the HathiTrust case, in chapter six, libraries risk liability by digitizing or allowing third parties to digitize their corpus of orphan works. Thus, given the absence of the right-holders, the risk is arguably unnecessary and overly burdensome on an institution dedicated to public service.[268] Let us now study similar issues arising with respect to out of print works, also known as out of commerce works.

4.3 WORKS OUT OF PRINT (OR OUT OF COMMERCE)

The second of the two gray categories - the out of print works – has its roots in the way the publishing market has worked for books especially in an era when books were only produced in print form. Out of print is a book no longer published and no more copies are made for commercial circulation. This is why the term out of commerce is also used for this category. As noted, the term mostly captures the pre-digital era. Nowadays, technological advancements allow for widespread use of the possibility of print on demand. Employing the 'out of commerce' term, the above-mentioned new EU Directive provides a definition under Article 7(2) that is helpful, according to which:

[267] See also Katz, A., 2012. The Orphans, the Market, and the Copyright Dogma: a modest solution for a grand problem. *Berkeley Technology Law Journal*, *27*(3), pp.1285-1346, at p. 1337, where he remarkably notes:

> A discussion of solutions to the orphan works problem will not be complete before addressing why has it been so challenging to find an acceptable and workable solution to this problem in the first place. The difficulty is puzzling because owners of orphan works are, by definition, absent from the debate about orphan works, and normally, when discussions about contemplated reforms do not involve those who might be directly affected by them, one could expect that reform would be easy. Specifically, *one could expect that in a setting where users lobby for reform that would allow them to use orphan works, and owners of those works are absent, passing a pro-user reform (even overly pro-user) would be a breeze.* Therefore, the fact that it has been difficult to find an acceptable solution, and that many of the proposed solutions involve serious impediments on using orphan works, suggest that the political economy of the orphan works problem is complicated, and that there is much at stake—not necessarily for the interests of orphan owners, but for the interests of those who speak on their behalf. [emphasis added]

[268] See Hansen, supra note 51.

> A work or other subject-matter shall be deemed to be out of commerce when the whole work or other subject-matter, in all its translations, versions and manifestations, is not available to the public through customary channels of commerce and cannot be reasonably expected to become so. [269]

Out of print or out of commerce works become scarce because official printing of the works stops. Copyright then causes an inefficiency when it inhibits further distribution of these scarce works; for example, by restricting reprints of existing copies. Copyright becomes an obstacle either because right-holders have no interest in further distribution, or simply because the right-holders cannot be easily found to give permission. This is why the category overlaps with orphan works. For libraries, this category is an obstacle for the same reasons that the orphan works are – especially so when they wish to make more (digital) copies for both preservation and access purposes. Again, libraries would optimally want to be able to make more copies of these works than the limited number allowed, or digitize them and include them in searchable catalogues, or further display them in their e-portals and e-lend them.

According to David Hansen, there are two definitional approaches to orphan works which I think are also relevant to out of print works.[270] The first approach focuses on the inability of a potential user to identify and locate the right-holder from whom permission is sought. The other approach places the inability of the user to easily obtain permission to use a particular work central to the problem. Out of print books fit squarely in this category that overlaps with orphan works. This broader issue of the inability or difficulty to trace and contact the copyright owners has led to the perception that out of print and orphan works are problems of market failure.[271] A potential user - institutional or individual user - faces disproportionate transaction costs to obtain authorization from right-holders, as well as the risk of infringement liability. Thus, the user will usually forgo the use even though had he been able to locate the copyright owner, a deal would have been struck for that use.[272] Although it is to be expected that rights-clearance involves certain transaction costs, with these works the costs become disproportionally high and results are not guaranteed.

---

[269] Proposal for a Directive of the European Parliament and of the Council on Copyright in the Digital Single Market, supra note 12.
[270] Hansen, Orphan Works: Definitional Issues, supra note 49.
[271] Ibid, p.1.
[272] Loren, supra note 255, at p.1433.

Yet again, as the Google Books case teaches us, there is value in digitizing these works, both orphans and out of print. Google, as a private user, together with the libraries that partnered with the company to digitize their entire collections took the risk and digitized copyrighted works that mostly included books of these two categories. We will refer to the facts of the case and the parties' arguments in detail in chapter six. Suffice to say here that libraries, together with more stakeholders from the technology and other sectors, have shown a clear interest in surpassing the copyright default permission-to-use rule with regard to both orphan and out of print works. The EU legislative efforts mentioned above and opt-out logic that was *de facto* employed by Google and HathiTrust in the respective case-law might be proof that the direction of the law is already shifting.[273] This clear interest is arguably due to the computational value of a digitized corpus. The additional value of the book in digital form will be the center of our discussion in the next section, which adds the public domain works to the mix.

4.4 TOGETHER WITH ORPHAN AND OUT OF PRINT WORKS, THE PUBLIC DOMAIN IS REGAINING 'VISIBLE' VALUE

When copyright expires, books enter a space of intellectual commons that we call the public domain. There are many different definitions of the public domain because of the differences among jurisdictions on what exactly is worthy of copyright protection. The definition is easier with respect to books. Jurisdictional differences only concern the copyright length. Normally all books after a certain period, which is calculated in accordance with the length of the copyright protection, enter into the public domain.

Paul Heald models the public domain as a function of, first, the effect of term length on price, second, the effect of term length on access/distribution (the availability of copies of the work itself), and third, the effect of term length and policing doctrines on follow-on creation (the number and quality of derivative works).[274] He explains that the regulation of the copyright term and doctrinal rules that prevent the proprietization of facts, ideas, concepts, methods and procedures are both means of policing the costs associated with copyright protection. He explores the size of the public domain space as a function of copyright duration rules and

[273] See more extensively *infra*, sections 6.1.5 and 6.1.8.
[274] Heald, supra note 116, p. 6. See also his cited research finding that ownership is a drag on digitization (p. 11).

policing doctrines, such as the idea/expression dichotomy. Indeed, the optimal size, or scope, of copyright protection is tightly connected to the definition of the public domain. Heald observes that the proper size and content of the public domain raises fundamental issues concerning the efficiency of property rights. He notes that both issues are under-theorized and the subject is of limited empirical research.[275] Also, he contends that the value of the public domain will be its 'net' value: that is the value generated by the work being in the public domain over and above what it would generate under copyright.[276] Going back to the definitions of the public domain, complementary to the definitions formulated in negative terms (works not restricted by copyright) are positive definitions, such as the public domain is a space of intellectual commons. [277] The management of public domain works is, however, much more complex than first appears. The public goods attributes of books are, indeed, encouraging an analysis of the possibility for their management as common goods. This is why, for instance, the theory of economic governance of common resources proposed by Elinor Ostrom offers useful insight. Ostrom develops a theory to explain phenomena that do not fit into a dichotomous world constituted by the market or the state. [278] She challenged the traditional worldview of simple systems of two

[275] Ibid, p. 2.

[276] Ibid, citing Pollock, Stepan, & Valimakki, Valuing the EU Public Domain, Cambridge Working Papers in Economics 1047 (2010), 1 – 54, p. 4.

[277] Another important and related question is whether the public domain equates the commons. In this case the boundaries of property again negatively define the public domain. James Boyle points out the inconsistency in which legal scholars use the two terms, meaning commons and the public domain. See Boyle, J., 2003. Foreword: The opposite of property?. *Law and contemporary problems*, pp.1-32, at p. 9.

Salzberger, on the other hand, thinks that equating the public domain with commons provides for a good definition from a law and economics perspective in the sense that proprietization with of intellectual works might be seen as aiming to avoid a certain type of tragedy of the commons.

Salzberger, supra note 117, p.28.

[278] Ostrom, E., 2010. Beyond markets and states: polycentric governance of complex economic systems. *Transnational Corporations Review*, 2(2), pp.1-12. Doing mostly empirical research she found diverse institutional arrangements for governing common-pool resources and proceeded in institutional analysis and development framework ('IAD'). The results from her and her team's experimental studies were fascinating. In sum they found that isolated, anonymous individuals overharvest from common-pool resources, that studies challenge the presumption that governments always do a better job than users in organizing and protecting important resources and also focused on the central role of trust in coping with social dilemmas.

While classic economics differentiated only between pure private goods, both excludable and rivalrous, and public goods, nonexcludable and nonrivalrous, Ostrom's theory differantiates at least four types of goods. First, private goods to which she adds the possibility of subtractability, instead of rivalry of consumption. She conceptualized a broad spectrum of possibilities where subtractability of use and excludability vary from low to high rather than characterizing them as either present or absent. Second, there are pure public goods which are nonexcludable and nonrivarlous. Next, there are club goods, or 'toll' goods, and finally a distinct category of common-pool resources. Common-pool resources share the attribute of subtractability with private goods and difficulty of exclusion with public goods. Forests, water systems, fisheries, and the global atmosphere are examples of common pool resources. She also insisted on a cautious use of terminology. *Common property resources*, for example, is a usual term that she thinks is confusing, the confusion coming from the combination of the term property and resource.

optimal organizational forms, the market for private goods and government for non-private goods. Following on from this observation, we need to think harder about where books fit as a resource in a spectrum that ranges from purely private to purely public goods, including common pool resources. The division between copyrighted works and works in the public domain is already a clear indication that books are not only private goods or only public or common goods. This is even more the case because, as the law currently stands, the legal status of each book changes over time: it starts as copyrighted but ultimately moves into the public domain. The need to follow Ostrom's advice and maintain her proposed flexibility or rather complexity in the economic analysis of books is something to keep in mind when looking into the different status of works as we have done so far.

For our focus that rests with the libraries, it suffices to insist on the legal consequence of the public domain status which is that all users may appropriate freely and without interference from competing claimants the content of intellectual works. As institutional users, libraries can collect, copy in any form and distribute without restriction their copies of public domain works. These works pose no legal concerns when it comes to digitization projects and, indeed, libraries have been able to experiment with digital access and lending programs. To continue the claim we made at the end of the previous section, several stakeholders, including libraries and technology companies, which found themselves partnering in this, understood the additional value of books in computable form. Thus a very rich public domain corpus, together with orphan and out of print works, became relevant to mass digitization projects planned. The purpose of this section is to examine in more detail what the value of library collections had been before mass digitization was possible, and what digitization adds to that. The argument remains that digitization introduces additional value to the preexisting value of these books.

Below we will, first, examine the attributes and value of the public domain. There is a body of important scholarship recognizing the great value of maintaining a growing public domain.[279] This scholarship can be indicative for the value of library collections as well, encompassing a public domain corpus but also copyrighted books including above mentioned 'in-between' categories which will ultimately enter the public domain. According to Pamela Samuelson, the public domain entails social utility that can be broken down into at least eight

[279] See Boyle, supra note 113, Samuelson, supra note 114, Cohen, supra note 114, Heald, supra note 116, Salzberger, supra note 117 and Finding the Public Domain: Copyright Review Management System Toolkit, supra note 53.

distinct values: (i) building blocks for the creation of new knowledge, (ii) enablers of competitive imitation, (iii) enablers of follow-on innovation, (iv) enablers of low cost access to information, (v) enablers of public access to cultural heritage, (vi) enablers of education, (vii) enablers of public health and safety, (viii) enablers of deliberative democracy. A ninth value that promotes personal autonomy and artistic self-expression is seen by the author as already covered by the values of creating new knowledge (i) or enabling deliberative democracy (viii).[280] For another interesting perspective on the issue, Eli Salzberger describes the public domain as a public sphere, a place in which individuals meet each other, interact, exchange views and information, attempt to influence each other's opinions and preferences and absorb inspiration and ideas for creation.[281] The picture that Salzberger draws with this definition resembles the marketplace of ideas that we discussed in the beginning of the dissertation with respect to freedom of expression and its complex relationship with copyright law.[282] All of these values are transferable to a library collection: available to the patrons, the books in the collection are potentially building blocks for the creation of new works, encourage competitive imitation, enable education and low cost access to information and so forth. The value within a rich library collection mirrors the value we have seen in the library itself as a public service institution.

Public domain and copyrighted books stand next to each other in a library collection. They are usually important as a unity. Think for example of the needs of a thematic research on a medical issue. One would visit a library and find the appropriate collection (a real or a virtual shelf) which will most probably include both older (in the public domain) and newer books (copyrighted or with an obscure 'in-between' status). A comprehensive study will need both. Indeed, libraries' added value compared to the book market is that it curates and makes available with the same care copyrighted and public domain works, irrespective of their commercial demand. In sum, the institutional role of libraries with regards to the public domain in particular becomes very important when we think of their traditional curational functions; for example, preservation, and cataloguing can transform an incoherent and perhaps less accessible body of intellectual commons to systematic collections. In this respect, libraries can perform better than market players the long-standing role of making the public domain relevant and accessible to their patrons. The examples of the NYPL public domain release and the Copyright Review

[280] Samuelson, Challenges in Mapping the Public domain, supra note 114.
[281] Salzberger, supra note 117.
[282] See first Justice Holmes dissent in Abrams v. United States, 250 U.S. 616 (1919).

Management System at the University of Michigan, which we mentioned in the beginning of this chapter, are exemplary.

On a side-note, the public domain status of intellectual works does not necessarily mean that these works are *de facto* out of the market. Freedom from copyright means that they can be copied, used, read and distributed for profit and not for profit. A number of scholars have documented examples of books that have had commercial success without any copyright protection. Breyer, for example, points to the market for imported English books in the United States that operates without copyright protection.[283] Informal rules between importers and publishers, including exclusivity agreements, were effectively doing the job of copyright and maintaining a short-term market for these works. There are also examples of circulated government works that became popular, such as the Warren Report or the 9/11 Report, and have been successfully sold despite belonging to the public domain.[284] The argument that a market with public domain works is possible is also supported by the existence of reprints of the classics. All these examples prove that demand can create a market which at least covers the cost of the copies without copyright rules.

Coming back to the value of library collections, in view of the possibilities that digitization opens one can observe a renewed interest in all the aforementioned categories of books, irrespective of their commercial viability as physical copies. Besides the obvious benefit of preservation and 'easy' storage, a digital version of an entire corpus of books offers more advantages: librarians can experiment and innovate in a number of ways; for example they can create searchable catalogues, systematize the collections with multiple key words, or find new ways to group books. As per access, they can experiment with various platforms, especially with their digitized public domain works, and also with e-lending. We shall return to the issue of the computational value of digitized books in the next part of the thesis, which adds on top of the libraries' prospects the value that technology companies might seek in the digitization of library collections.

[283] Breyer, supra note 151, p.300 et seq.
[284] Ibid, at p. 302. See also Levine, D. and Boldrin, M., 2008. Against intellectual monopoly. *Cambridge University Press* where more examples are cited. Both reports have been financed by government.

MIDWAY SUMMARY

Part One of this dissertation aimed at setting the ground for the investigation of our topic: the phenomenon of digitization and the challenges that libraries face in their transition to the digital era predominantly due to the copyright framework. It starts from a time when books had only physical form and the copying was difficult and costly. In the beginning of this part we saw how copyright laws have applied to books and to libraries since their inception (Chapter 1). Copyright evolved as a field of market regulation for intellectual goods, including books, on the basis of moral and economic theories (Chapter 2). This part then concentrated on the libraries, their public service roles as institutions and their substantial and symbolic role in book de-commodification, widespread access to knowledge and ultimately contributing to the encouragement of learning – the latter being also the express rationale of copyright laws (Chapter 3). The final chapter devoted more time in the details of the copyright framework as applied today to libraries, examining a variety of rules - international, supranational and national - and a number of distinct issues - exceptions allowing the making of a number of copies, e-lending, and the problem with orphan works and works out of print (Chapter 4). In this context, our purpose is to evaluate whether the existing copyright framework indeed facilitates or inhibits the work that libraries need to do to fulfill their mandates (libraries' mandates have been discussed as consisting in two pillars: *preservation* and *access*). Furthermore, entering the digital age, it is necessary to ask whether it facilitates or inhibits the potentials that libraries have to serve the needs of their communities. Our overall assessment, thus far, is fairly positive with regard to the evolution of the role of the 'brick' library given the existing copyright rules in the pre-digital era. Entering the digital era our hypothesis and early assessment is to the negative. Chapter four has already touched upon a number of legal obstacles. We could perhaps isolate the digitization of entire library collections together with e-lending as the biggest puzzles. In addition, we saw how centrally debated, by legislators and policy-makers, is the need for reform of the current rules, to adjust to the needs of the digital era.

The **Second Part** of the thesis will aim at providing further answers to evaluate the current framework with more certainty. Maintaining our focus on the library, we will now concentrate on the reality and promises of digitization. We expect to see (i.) a process of transformation and

(ii.) a struggle: (i.) libraries' transformation to 'digital libraries' and (ii.) libraries engaging in mass or smaller scale digitization projects and struggling to remain relevant in a game dominated by powerful technology companies. Amidst this transformation and struggle, which we will portray as a phenomenon of 'digitization rush', the seen regulatory framework does not seem to offer the necessary support to libraries. We will test this hypothesis examining the various developments, mapping relevant stakeholders and, last but not least, delving into the complex litigation that several digitization projects started.

PART II

II.

DIGITIZATION RUSH

> *Project Gutenberg began in 1971 when Michael Hart was given an operator's account with $100,000,000 of computer time in it by the operators of the Xerox Sigma V mainframe at the Materials Research Lab at the University of Illinois. […] At any rate, Michael decided there was nothing he could do, in the way of "normal computing," that would repay the huge value of the computer time he had been given ... so he had to create $100,000,000 worth of value in some other manner. An hour and 47 minutes later, he announced that the greatest value created by computers would not be computing, but would be the storage, retrieval, and searching of what was stored in our libraries.*[285]

CHAPTERS' OVERVIEW

Having a better understanding of the context provided in the first part, Part II focuses on the phenomenon of book digitization. Digitization is a method of copying that results in a digital, thus computable form of the copy. Copying has always played an important role for the preservation and dissemination of books. While methods of copying gradually advanced after the invention of the printing press, it seems that digitization as a technological method for copying has drastically changed the playing field. It has now become easier and more 'affordable' – comparatively speaking – to digitize books on small and mass scale. This means that old books that only had a 'physical' lifespan can be copied and can have a parallel digital existence.

[285] Hart, M., 1992. The History and Philosophy of Project Gutenberg by Michael Hart, online available at https://www.gutenberg.org/wiki/Gutenberg:The_History_and_Philosophy_of_Project_Gutenberg_by_Michael_Hart.

This part of the thesis contains four more chapters. The first – chapter five of the thesis - explains **"How the Internet did not kill libraries"**. Digitization, together with the possibilities of computation, did not signal the end of the institution dedicated to preserving and providing access to books but actually reveals its unprecedented potential. Indeed, the economics of copying are complex, but the potential that digitization has, particularly at mass scale, is substantial. Previous utopian visions of creating a world library, which encompasses the knowledge of all mankind, might in fact becoming closer to reality. This chapter also demonstrates how libraries are actually deeply connected to the history of the Internet and its development. Lastly, it underscores the relevance of the physical library, which is far from obsolete in the digital age.

Chapter six, **"A timeline of digitization initiatives and the relevant case-law affecting digital libraries"**, creates a timeline of digitization initiatives, which are primarily geared towards the creation of digital libraries, starting in the United States and then in Europe. Most importantly, a body of case-law has developed, some of which involve these projects. The analysis of the litigation is part of the timeline and includes the following cases:

i. *Authors Guild v. Google*
ii. *Authors Guild v. HathiTrust*
iii. *Technische Universität Darmstadt v. Eugen Ulmer*
iv. *Vereniging Openbare Bibliotheken v. Stichting Leenrecht*
v. *Marc Soulier & Sara Doke v. Premier ministre & Ministre de la Culture et de la Communication*.

The timeline covers five decades and presents an interesting range of stakeholders and their efforts. While the initiatives are diverse and different in nature and in scope, they share one common vision: the creation of a digital library. By drawing conclusions based on the timeline of various projects and lawsuits, the overall purpose is to understand the position that libraries find themselves in amidst what we will call a 'digitization rush.' The competitive environment created by this array of stakeholders, together with the limitations that libraries face in view of the legislation and litigation contribute to their challenging position. Finally, this chapter places this era of digitization in the context of the so-called 'copyright wars.'

The next chapter discusses **"The adverse consequences of digitization on books and on libraries."** First, it focuses on the effects on the book as a medium. Books are undergoing a transformation in the digital age, lacking the physicality usually associated with them and also perhaps losing their attributes as merit goods. The chapter describes how the application of the concept of property on books, like many other goods in the digital age, is challenged. Most importantly, this chapter observes a failure in the availability of books in the digital age, despite it being the 'age of the copies'. This failure contradicts the reader's high expectations of book availability and access. File sharing of literary content and robust pirate libraries available online represent two attempts to fill this gap. The so-called 'shadow libraries' provide alternative access to many books that are otherwise unavailable to potential readers. This alternative access mostly covers out of print books that were either legitimately or illegitimately digitized. One can see this as an additional challenge for the position of libraries in the digital age. Thus, chapter seven makes reference to the role of libraries in equalizing access to knowledge, as discussed in chapter one, and the need to maintain this role also in the online environment. Finally, it explores catastrophic scenarios of books or libraries becoming obsolete but quickly dismissing them, however, as less realistic.

Part II is followed by a short final chapter, chapter eight, which draws some conclusions from the entire thesis: **"Conclusion – What's the Rush"**. In a more general fashion digitization has caused a broader phenomenon of competition among stakeholders to capture the value of a new resource: the digitized book. Chapter eight concludes with the observation of a 'digitization rush' phenomenon and further explores profounder motives for the participation of parties such as Google in the rush. As the competition took the form of copyright controversies, also before courts (the litigation that we saw under chapter six), it also forms part of the broader 'copyright wars' of this century. The conclusion will argue for the strengthening of libraries' position in this context and explain why this particular institution should be trusted with the guardianship of our literary treasures in the age of digitization.

CHAPTER FIVE - HOW THE INTERNET DID NOT KILL THE LIBRARIES

The relationship between libraries and digitization has been constructive despite doubts or fears expressed about the future of the institution. The concerns mainly centered on how libraries might become irrelevant in the digital age and in the age of the Internet. In short, digitization promises new ways of collecting, managing, and accessing content. Ultimately big ideas, such as the creation of a world library, a common reference for humankind, became technologically possible because of digitization. Interestingly, the story of libraries is tied to the development of the Internet itself. The two 'institutions' share visions of organizing the world's knowledge, make knowledge accessible and connect individuals who access it. These considerations are important background information to remember when we proceed to the next parts of the thesis that focus on the legal saga of specific digital library projects.

The chapter begins with ways in which libraries and the development of the Internet intersect. It then moves to the promises and perils of digitization and concentrates on the promises relating to the creation of a 'world library.' Finally, it introduces economic and legal considerations: in particular, how the economics of copying become more relevant for copyright when we move from small to mass scale copying capabilities.

5.1 How can a story of libraries actually be connected to the development of the Internet?

Robert Kahn and Vinton Cert invented the Transmission Control Protocol (TCP) and the Internet Protocol (IP) in the early 1980's. 1983 is considered the birth year of the Internet, which is the year when TCP/IP protocols were adopted by the ARPANET. In a paper published in 1999 Robert Kahn and Vinton Cerf, inventors of the TCP/IP protocols, explain in lay-terms the Internet, specifically what it is and what makes it work.[286] They explicate that the Internet is an architecture which, firstly, enables communications connectivity, packet delivery and end-to-end

[286] Kahn R. and Cerf V., 1999. What is the Internet (and what makes it work), paper online available at http://www.policyscience.net/cerf.pdf.

communication, and, secondly, enables the creation, storage and access of information resources. Thus, these two Internet 'founding fathers' underline two elements of the Internet architecture: (i) as far as the technology or scientific character of the Internet is concerned, it is an architecture that enables communications connectivity and (ii) the Internet is also an *information system "independent of its underlying communications infrastructure."* [287] In defining the Internet, Kahn and Cerf insist that it is an information system, therefore an accurate definition must include:

> […] not only the underlying communications technology, but also higher-level protocols and end-user applications, the associated data structures and the means by which the information may be processed, manifested, or otherwise used.[288]

The analogy that Kahn and Cerf use to help us visualize the Internet architecture is that of a highway. To this end, they use the now renowned term *information superhighway*. Their analogy refers not only to the highway lanes but to the entire highway system: entries and exits, signs, maps, rules and services, like gas stations.[289] While not contesting this iconic highway metaphor, in this section I try to play with another metaphor: that of visualizing the Internet as a library, and the library of the future operating similarly to the way the Internet operates today. The two metaphors are not entirely unconnected. As discussed in chapter three, libraries have infrastructural characteristics that place them, like the highway system, in the category of public services. The connection between the Internet and libraries is robust especially with respect to the second characteristic that Kahn and Cerf emphasize: they are both information systems. Kahn and Cerf note three functions of the Internet architecture: (i) creation, (ii) storage and (iii) access

[287] Ibid at p.3. In their own words: "In essence, the Internet is an architecture, although many people confuse it with its implementation. When the Internet is looked at as an architecture, it manifests two different abstractions. One abstraction deals with communications connectivity, packet delivery and a variety of end-end communication services. The other abstraction deals with the Internet as an information system, independent of its underlying communications infrastructure, which allows creation, storage and access to a wide range of information resources, including digital objects and related services at various levels of abstraction."

[288] Ibid, p.11. They cite the definition of the Internet by the Federal Networking Council (1995) as the best currently existing. See also Benkler Y. and Clark D. 2016, The Internet: Introduction, *Daedalus* 145(1): 5-8: "The Internet was born in 1983(…) Thirty-two years later, the Internet has become the most fundamental global communications and knowledge infrastructure of our age, and is fast becoming the basic data-and-control network of the coming decade."

[289] Kahn and Cerf, supra note 286, at p.11.

to information resources. All three, and particularly the last two, are traditional functions of libraries, as demonstrated in the previous section. The form in which traditional libraries store information resources is mostly of complete literary works, while the Internet running on computers allows for the creation, storage and access to data – digital recourses that can be translated on screen in many different forms of content, including complete literary works.

The comparison between the Internet and libraries is not new. In a paper published by the International Journal of Human-Computer Studies in 1995, Luciano Floridi wrote about the Internet and the future of organized knowledge.[290] In this paper he explicitly uses the Internet-library metaphor several times. Furthermore, he adds flesh to the metaphor by comparing specific characteristics of both. For example, he notes that "the Internet is free, in the same way as you would expect the information in a public library to be free." [291] Service fees for Internet service providers (ISPs) or costs of purchasing software would be, for him, the equivalent to the costs one has to invest to physically reach a library. In fact, he mentions software as a type of vehicle that help us reach the actual network and expands the metaphor noting how software offer services for users to reach the Internet, just like transportation can help us reach a library. Floridi maintains:

> Nobody is running the system or will be able to control it in the future. When Bill Gates speaks of his plans to dominate the Internet, what he really means is that he wants to make sure that Microsoft Windows '95 and its network will result in the most popular interface and route for single users who wish to manage their own access to the network through their personal computer. *In*

[290] Floridi, L. (1995) The Internet: Which Future for Organized Knowledge—Frankenstein or Pygmalion? International Journal of Human-Computer Studies 43: 261-274. The article is a revised version of a paper the author gave to the first international conference, promoted by the UNESCO Philosophy Forum, to celebrate the fiftieth anniversary of the founding of the organization (Paris, 14-17 March, 1995).

[291] Ibid, p 263. One might argue that this view is perhaps outdated. On the one hand, we have a school of Internet scholars who praise the free and collaborative ethos of the Internet. Among these lawyers, Yochai Benkler is perhaps the leading intellectual in this school. See Benkler, Y., 2006. *The wealth of networks: How social production transforms markets and freedom*. Yale University Press and Benkler, Y., 2011. *The penguin and the leviathan: How cooperation triumphs over self-interest*. Crown Business.
In view of today's apparent commercialization of the Internet other scholars have started to seriously dispute the hyper-romanticism of everything 'cyber'. Among many, see Morozov, E., 2014. *To save everything, click here: The folly of technological solutionism*. Public Affairs.

other words, he would like to run the main bus line that takes you to the library, not the library in itself. It is already a very ambitious project.[292]

However romanticized, the analogy between the Internet as a 'free' medium and public libraries as institutions that provide free access to information has significant merit. Like a library, Internet is a 'space' full of intellectual resources. To surf online we search via search engines that allow us to navigate vast amounts of information, which is similar to searching a library catalogue or browsing library shelves. Once in a library, resources are mostly freely available, which is like most information online. Of course the creation of both library resources and information placed and freely accessible online are not costless. The costs, however, are usually covered indirectly, for example by taxpayers, donations, and in the case of the Internet payment fees to ISPs.

Floridi predicted the transformation of libraries in view of the technological advancements that he described. His predictions were close to the contemporary understanding of digital libraries. In the same article he explains how he understands the transition "from a library which stores knowledge physically recorded on paper to the consulting library which provides access to electronic information on the network" or "from the library as a building to the library as a gate-node in the virtual space of the digital encyclopedia."[293] What he saw as a big digital library, however, is the Internet itself which he understood as an *evolving system of encoded knowledge*, a *human encyclopedia.*[294] Thus, his work can be read as affirming an interesting and deeper link between these two worlds. What Floridi envisioned back in 1995 has come closer to becoming true; first of all with the creation of online encyclopedias, such as Wikipedia, and second with the initiation of digital library projects. Arguably the advancement of search engines, such as Google, which are fed with mass content (including millions of books) is also a move in the same direction.

More importantly, with this analysis Floridi places libraries and computing and the Internet in a context of what he calls history of the human encyclopedia, in which they attempt to effectively manage and organize knowledge. He is, basically, less concerned about providing a

[292] Ibid (Floridi), p.263 [emphasis in the quote is added].
[293] Ibid, p. 264.

[294] Ibid, p. 265.

definition for the Internet and is more preoccupied with understanding its future, and particularly how it can and will affect *organized knowledge*.[295] Floridi saw the Internet content evolving in mixed ways which provoked him to think of a future human encyclopedia. Thus, in his analysis he ultimately jumped from the library analogy to that of an encyclopedia. His overall description, however, was that of an evolutionary process:

> Organized knowledge moved from the unstable memory of oral transmission to the written text, from volumina, providing only a linear access to their contents, to codices, which made possible a random access, from the appearance of the tables of contents to the elaboration of indexes, from the establishment of public libraries to the elaboration of a science of bibliography, from the appearance of new reference books, such as encyclopedias or lexicons, to the publication of reviews and ever more specialized forms of literature.[296]

This evolutionary story of the phases is fascinating. Today the mass amounts of content that is found on the Internet has only passed the indexing phase. As we will discuss further, digitization made it possible to elaborate and process content, which used to be only found in print magazines and books but is now computable and accessible online once digitized. Sticking to the image of the evolutionary process, however, the transition between phases might result in short and long-term winners and losers. A logical question to ask is whether the Internet will signal the end of all previous forms or mediums associated with the organization of knowledge. It seems that Floridi was at the time concerned or at least preoccupied by the possibility that the revolutionary effects of the Internet *vis-à-vis* knowledge production access and management might render libraries obsolete.[297]

5.2 THE PROMISES AND PERILS OF DIGITIZATION

[295] Ibid, p. 264.
[296] Ibid, pp.265-6.
[297] Ibid, at p. 264.

Will the Internet ultimately kill the library? The question is, of course, a more provocative gesture than accurate. Today's times already witness massive changes in the way knowledge is produced, accessed and managed online: Wikipedia has established itself as a successful example of an online encyclopedia; online retailers, such as Amazon, directly compete with the traditional bookstores in the distribution of books; e-books have also established their position in the marketplace. Last but not least, digitization of literary works is now technologically possible in more cost-effective ways than was previously the case. The revolution that digitization brings about, mostly on a mass scale basis, is central to this thesis and is, I argue, critical for the future of access to knowledge.

While most information and also literary production is being digitized and in one way or another is made available online, the way we conceptualize access to knowledge is also transforming. Perhaps digitization of the 20th and 21st centuries is of the same magnitude of the revolution brought about in the 15th century with the invention of the printing press. Throughout history, disruptive technology has been interpreted as a blessing and a threat; a blessing to those who welcome change and innovation in the traditional ways of doing things and a threat to those who for different reasons are opposed to the change. Established players belonging to several intellectual property industries, such as the movie and music industry, had many reasons to object to the changes that the Internet and digital technologies caused. Technology has been disruptive to certain business models and forced industries to either adapt or lose their dominance. Within a few years, changes first happened in the news market and immediately after the book market transformed. Libraries are arguably a different case because of their public service role, which differentiates them from other market players.

During this transition to the digital era, new players, in particular technology companies, emerged and showed interest in books and copying them. In the aftermath of the explosion of small and mass scale digitization, concerns were voiced by institutions and industries worried about losing their primacy in the 'book business' and ultimately about their survival in the new age. Similar to the music and movies industry beforehand, the entire book industry, authors, readers and libraries became increasingly concerned. Scholars studying the phenomenon had expressed concerns about the effects of digitization. In one of their famous blog debates, Gary Becker and Richard Posner discussed the effects of digitization on bookstores and libraries. Together with the real changes unfolding around them, such the closing down of bookstores,

music stores and movie theatres, they were concerned about copyright and how it was being bypassed through piracy. It was Posner who first asked whether bookstores and libraries are doomed because of digitization.[298] He cited Joseph Schumpeter's notion of creative destruction which describes "the process by which innovation […] promotes economic growth and welfare but at the cost of wiping out existing economic practices or institutions."[299] He argued that creative destruction best describes the transformation that the information sector is undergoing. Digitization, Posner maintained, has destructive effects on bookstores, libraries but the overall gains in access might ultimately offset costs, such as the disruption of the book industry, which has been built on a particular economic model, that of copyright.

Becker replied in a less optimistic tone. He did not really discuss libraries as much as bookstores, newspapers and movie theatres, whose decline he described as sharp as the decline of the postal system. According to Becker:

> It has been recognized for at least a decade that traditional bookstores and newspapers are essentially doomed by the growth of the Internet and digitization. Doomed also are postal systems, record albums, movie theatres, and most other traditional ways of providing information, entertainment, and other content to consumers. To take the US postal system as an example, after growing for many decades, the number of pieces of first class mail declined by 25% in past few years alone.[300]

In his blogpost Becker, like Posner but to a lesser extent, appears concerned about a changing reality. Neither are certain that the change is entirely negative. Becker, for example, adds that "hardly anyone would lament the decline of the US postal system, which traditionally has provided surly service by overpaid employees, introduced few innovations in mail delivery, and frequently failed to deliver mail on time. Email is much more efficient than the appropriately called snail mail, which explains why email communication has grown so rapidly, even among

[298] Posner R., Are Bookstores and Libraries Doomed by Digitization?, 03/04/2012 blog-post available at: http://www.becker-posner-blog.com/2012/03/are-bookstores-and-libraries-doomed-by-digitization-posner.html.
[299] Ibid.
[300] Becker G., Traditional Bookstores and Libraries are Doomed by the Internet, and so too are Many Other Ways to Distribute Content to Consumers, 03/04/2012 blog-post available at: http://www.becker-posnerblog.com/2012/03/traditional-bookstores-and-libraries-are-doomed-by-the-internet-and-so-too-are-many-otherways-to-distribute-content-to-cons.html.

older persons."[301] Given that both mention libraries and bookstores as institutions in decline, then it is natural to ask if they are happy to celebrate the new *more efficient* ways to distribute books or do they also forecast the end of the physical book era? Let us note that there is a distinction between the transformation of the book as a medium and the ways in which it is being distributed. Amazon's model, for example, has challenged small and big bookstores, arguably big retail bookstores more than small ones. Online shopping and home delivery of physical books is perhaps a bigger threat to bookstores than digitization. Libraries, on the other hand, cannot easily be grouped together since their functions and services differ to those of the book market. Access to electronic books, however, blurs the lines between bookstores, online retailers and libraries when the latter offer e-books to their patrons. In the case of electronic books, it can be argued that these actors are directly competing. Unrestricted e-lending programs provided by libraries would be able to cannibalize on sales of e-books in the market. Nevertheless, bookstores and libraries also need to remain distinct in our analysis because of their different objectives. Unlike bookstores, libraries do not consider books and e-books to be a commercial product. Therefore, the effects of digitization on these two actors should be studied separately.

Despite the challenges, it has been observed so far that most of the fears expressed with regard to the future of libraries are exaggerated.[302] On the other end of the spectrum emerge optimistic views on the most ambitious promise of digitization when it comes to the future of libraries: the realization of a world library providing access to the world's content. The following section discusses in depth what was once only utopian idea. Compared to the fears expressed of digitization, perhaps such visions are equally exaggerated. Yet they are deeply fascinating, and, while they used to belong to the realm of science fiction, they are arguably at least more realistic today.

5.3 Realizing the dreams for a World Library (or getting closer)

Visions such as the *World City*, a city of knowledge, or the *World Brain*, a machine containing all of the world's knowledge, can be traced back to the beginning of the 20th century. In 1910, Paul Otlet, a Belgian pioneer in information science, had imagined a central repository

[301] Ibid.
[302] See also section 7.3 *infra*.

and reference point for the whole world's information.[303] The *World City* (Cité Mondiale or in Otlet's words, Mundaneum) was a utopian city of knowledge, which he envisioned leading institutions of the world would construct. The city would serve as a central repository of the world's information, and would hold and manage all knowledge worldwide. Otlet worked hard to begin what he thought would become the world's repository, which was interrupted by World War II and signaled the end of his efforts. His legacy is significant, however, not only as a visionary but also because in a world without computers he revolutionized indexing preceding today's OPAC, the online public access catalogue held by libraries. Herbert George Wells, from England, had also famously imagined a super computer or a world encyclopedia which he described in a series of essays published in 1938 under the title *World Brain*.[304] In today's terms the World Brain would be a machine with the most advanced artificial intelligence capacities. For Wells, the first stage of the World Brain would entail the construction of a World Library. Together with Otlet and Wells, one could add Jorge Luis Borges, the Argentinian author and librarian, and his similar concept in the fantasy story of the *Library of Babel* which was a fantasy structure of an 'infinite' or 'complete' library (*La biblioteca total*).[305] Several decades later, these visions are often associated with contemporary accomplishments, such as Wikipedia and Google Books.[306]

ARPANET, the first network to use the Internet protocol, was developed only thirty years after Wells' idea of the World Brain. In 1969, the first host-to-host message was sent from Leonard Kleinrock's lab at UCLA to Doug Engelbart's lab at Stanford.[307] In the following chapter, while examining various digitization projects, we see the connection that the Stanford labs have had to the creation of digital libraries. For now, suffice it to say that the birth of the Internet, which occurred before digitization, brought visions such as those just described closer to reality. The utopian *World City* and even more so the *World Brain* machine could now plausibly transit from a sphere of science fiction to that of science.

---

[303] See Wright A. 2014. *Cataloging the World*. Oxford University Press.
[304] Wells H.G. 2016 (reprint). *World Brain*. Read Books Ltd.
[305] Borges, Jorge Luis.1941. *The library of Babel*.
[306] On the how the work of these visionaries predate Wikipedia see Reagle, J.M., 2010. *Good faith collaboration: The culture of Wikipedia*. MIT Press. The connection between the World Brain and the Google Books project is made in the 2013 documentary *Google and The World Brain*, directed by Ben Lewis.
[307] Leiner, B.M., Cerf, V.G., Clark, D.D., Kahn, R.E., Kleinrock, L., Lynch, D.C., Postel, J., Roberts, L.G. and Wolff, S., 2000. A brief history of the Internet, online available at the Internet Society webpage: http://www.internetsociety.org/internet/what-internet/history-internet/brief-history-internet#Timeline.

Contemporary scholarship becomes clearer about the relationship between the web and organization of knowledge. In the context of discussions about the future of the Internet and the future of libraries, the two worlds intersect. John Palfrey has been one of the primary academics to recognize and highlight the importance of this intersection. We have already mentioned his recent book *BiblioTech: Why libraries matter more than ever in the age of Google,* where he makes a strong case for libraries in the digital age, highlighting their importance in the civic and educational life of Americans across many age groups.[308] For Palfrey the role of a traditional library has not altered significantly despite the digital revolution. It is only the need to adapt to different formats that makes the current situation challenging. As part of a volume of the American Academy of Arts and Sciences' magazine, Daedalus, dedicated to the future of the Internet, Palfrey discusses design choices that libraries need to make in what he calls the *digital-plus* era.[309] The volume, published in the winter of 2016, is dedicated to design choices of the Internet as a network. Palfrey's input relates to the tension between the original 'public' character of the Internet and the ongoing privatization of the networked environment. Libraries as public institutions struggle to fit the network's much privatized space but at the same time might prove to be the "necessary counterweight" to this space.[310] Palfrey explains that libraries now face the *digital-plus* era, when materials have not only transformed from paper to digital, but also appear in multiple forms in the digital space.[311] Some material, for example, has been born in paper and then digitized while others are born digital and then made available in multiple forms and mediums. Libraries, while continuing to play their institutional role of access and preservation, now have to deal with this heterogeneity of formats.[312]

Within the context of the digital plus era, Palfrey urges us to think of libraries as *nodes in a network* which ultimately works towards a larger goal: that of the development, curation and accessibility of knowledge.[313] Following on from this remark, let me return to the basic point of this chapter: the strong relationship between libraries and the Internet architecture. Palfrey takes the network vision further and gives concrete examples of libraries that have digitized collections, offer online access, and are interconnected to other online platforms (mostly libraries

[308] Palfrey, supra note 40.
[309] Palfrey, supra note 41.
[310] See also See also Benkler Y. and Clark D. 2016, The Internet: Introduction, *Daedalus* 145(1): 5-8, supra note 288.
[311] Palfrey J., supra note 41, p.80.
[312] Ibid, p.81
[313] Palfrey, supra note 41, p. 83.

but also different kinds of platforms) forming a communicating network of available resources.[314] He gives the example of the Emily Dickinson papers digitized by Harvard University, which were then made directly accessible through the university websites. The digitized papers were also accessible through the Digital Public Library of America, which incorporated the metadata of the collection into the database and would point the user that searched for the material to the university website.

The proliferation of a network on the basis of such example show how, in the age of digitization, libraries can build on top of the existing web infrastructure a vibrant content which expands in a similar way to the Internet itself. The key benefit of connecting to the Internet is basically connectivity and scalability, both of which are at the very heart of the Internet's success story. Thinking how the Internet was also inspired by an idea of a digital library, we can conceptualize a network of digital libraries built at the content layer of the network. Palfrey himself makes such direct analogy and points to the success of the Internet, the World Wide Web, and Wikipedia as networks.[315] He argues that libraries operating as a network in the digital era can experience growth in a similar manner to the World Wide Web and Wikipedia.[316] With the proliferation of digitization activities around the world, the idea of a network of digital libraries forming something of a 'world library', which grows similar to the World Wide Web or to content projects like Wikipedia, becomes at least technologically plausible.

Within this dissertation we will examine the most relevant mass digitization projects in the United States and Europe, which aimed at creating either one or a network of digital libraries. Before doing so, and to conclude this chapter let us review some of the more practical considerations concerning digitization, namely economic and legal implications that arise with the planning of mass scale copying or, ambitiously speaking, with the planning of digitizing "the world's books."[317]

---

[314] Ibid.
[315] Ibid, p.85
[316] Let us briefly note here how the comparison to Wikipedia and to the world wide web would also bear a similar meaning. With Wikipedia digital libraries, and most importantly a digital network of libraries in the way that Palfrey envisions it, share the knowledge commons element. With the World Wide Web digital libraries share the information access element. The World Wide Web makes information accessible via hyperlinks and operates on the Internet, which is the broader network. In the library context, the World Wide Web equivalent would probably be the library portals that make digital catalogues accessible.
[317] Quoting from the Google Books webpage explaining the Google Books Library Project: "an enhanced card catalog of the world's books" at https://www.google.com/googlebooks/library/.

5.4 THE ECONOMICS OF COPYING AND GOING FROM SMALL TO MASS SCALE

How many books exist in this world? This is a serious question that has been asked with equally serious ambitions for an answer. In August 2010 Google claimed that it calculated all books worldwide and estimated that there are 129,864,880.[318] Google explained its method of combining the databases of different resources, including but not limited to the International Standard Book Number (ISBN), library identifiers such as WorldCat accession numbers and the Library of Congress control numbers. Google's researchers were determined to find out the actual number of all books counted as volumes, taking into account that many of them are copies of the same content of the same and different editions. Google explains what they considered to count as a book:

> One definition of a book we find helpful inside Google when handling book metadata is a "tome," an idealized bound volume. A tome can have millions of copies (e.g. a particular edition of "Angels and Demons" by Dan Brown) or can exist in just one or two copies (such as an obscure master's thesis languishing in a university library). This is a convenient definition to work with, but it has drawbacks. For example, we count hardcover and paperback books produced from the same text twice, but treat several pamphlets bound together by a library as a single book.[319]

That same year, the 'Collections Trust', which is a UK based charity working with museums, libraries, galleries and archives and is dedicated to collection's management, produced a report for the European Commission. The report estimated approximately seventy-seven million individual book titles held by European libraries – as opposed to Google's one hundred and thirty million estimate for the continent.[320] Thus, Google's accuracy in the numbers has been somewhat contested. The cited report actually states that Google's calculation cannot be

---

[318] Taycher L., Books of the world, stand up and be counted! All 129,864,880 of you, August 05, 2010 at: http://booksearch.blogspot.com/2010/08/books-of-world-stand-up-and-be-counted.html.
[319] Ibid.
[320] *The Cost of Digitising Europe's Cultural Heritage - A Report for the Comité des Sages of the European Commission*, Prepared by Nick Poole, the Collections Trust, November 2010, available at http://nickpoole.org.uk/wp-content/uploads/2011/12/digiti_report.pdf.

completely accurate. It is nevertheless most likely to be the best estimate so far, taken into consideration that the company aimed at all book volumes globally and made multiple controls of the duplications of the resources used for metadata. Mistakes are ultimately bound to depend on the validity of the sources.[321] Indeed metadata coming from sources, such as ISBN, are also problematic. For example, books are assigned a unique International Standard Book Number since 2007. Therefore, it is necessary to take into account the preceding varying valuations in different countries, libraries and other organizations as well as the fact that the number is only assigned to books placed in commercial circulation.

The motivation behind these efforts to gauge how many books actually exist has been to calculate costs of copying on a mass scale, and more specifically costs of digitization. Google aimed at counting all books in this world, while the European Commission focused on books in Europe. Common in both efforts is their starting point: the physical book, or the book in paper version. Most precious books in library and other collections are paper books since the phenomenon of digital works is very recent. Copyright rules today apply mostly in the same way to paper and only-digital work, with the exception of lending. Still, our focus is mostly books that are 'born' in paper format and then digitized to become part of a digital corpus. The complications related to this transition appear much more challenging due to the substantial costs of mass copying as well as complicated copyright issues, such as out-of-print and orphan works that amount to special categories mostly for paper books.

This section is devoted to the economic and legal complications related to the copying of paper books, a process which began on a smaller scale but now mass digitization promises to fulfill the universal digital library dream. It has always been possible to copy books. Before the invention of the printing press, copying was done manually. Thus, it was costly and timely. Technological progress basically contributed and continues to contribute to making the copying process cheaper, faster and easier. The rules of copyright are tightly related to the economics of copying which we will look at immediately. The phenomenon of digitization, which arguably generates a revolution like Gutenberg's invention of the printing press, is just another method of copying, only vastly more advanced. The key question on copyright rules in the digital age is whether the explosion in the volume of copying makes any difference.

[321] Ibid, p. 28

The distinction between economics of copyright and the economics of copying was initially introduced by Landes and Posner.[322] The distinction is extremely helpful. It has also been adopted by other scholars.[323] The economics of copyright is the study of the consequences that a set of legal rules has on certain goods; for example, on books that carry content on which there is copyright protection. The rules of copyright of course have a certain impact on the economy of these goods. The economics of copying, however, is the study of the technology of reproduction and its impacts on the economy. In other words, it is the study of the effects of new technology that makes the process of copying easier and cheaper.[324] There is ample literature on new copyright technologies that arrived in our homes such as the VCR, then recording cassettes and CDs and finally file-sharing on personal computers.[325] In addition, there is scholarship on illegal downloading of files online and peer-to-peer file-sharing, which exploded around the time of the Napster case decision at the turn of the millennium.[326]

With the advancement of technology, the quality of copies also advanced to the point that they are perfect substitutes of the original, or rather official, products. This was definitely not a negligible development for industries that profit from selling official copies as products, particularly in the music and movie industry. The study of the economics of copying and that of piracy were soon developing hand-in-hand and a number of empirical studies complemented the economic analyses.[327] Economic theories of copying have focused on the effects of copying on production and the overall social welfare costs. Some assume that consumers do not care about

---

[322] Landes, W.M. and Posner, R.A., 1989. An economic analysis of copyright law. *The Journal of Legal Studies*, *18*(2), pp.325-363.

[323] Towse et al., supra note 68; Varian, H.R., 2005. Copying and copyright. *The Journal of Economic Perspectives*, 19(2), pp.121-138.

[324] Ibid. (Towse et al.).

[325] See, among others, Samuelson, P., 1993. Fair use for computer programs and other copyrightable works in digital form: The implications of Sony, Galoob and Sega. *J. Intell. Prop. L.*, *1*, p.49; Gordon, W.J., 1982. Fair use as market failure: a structural and economic analysis of the" Betamax" case and its predecessors. *Columbia Law Review*, *82*(8), pp.1600-1657.

[326] *A&M Records, Inc. v. Napster, Inc.*, 239 F.3d 1004 (2001).

[327] Peitz, M. and Waelbroeck, P., 2003. Piracy of digital products: A critical review of the economics literature, CESifo Working Paper Series No. 1071. available at SSRN: https://ssrn.com/abstract=466063; Liebowitz, S.J., 2004. Will MP3 downloads annihilate the record industry? The evidence so far. In *Intellectual property and Entrepreneurship* (pp. 229-260). Emerald Group Publishing Limited; Zentner, A., 2005. File-sharing and international sales of copyrighted music: An empirical analysis with a panel of countries. *Topics in Economic Analysis & Policy*, *5*(1), p.21; Rob, R. and Waldfogel, J., 2006. Piracy on the high C's: Music downloading, sales displacement, and social welfare in a sample of college students. *The Journal of Law and Economics*, *49*(1), pp.29-62; Oberholzer-Gee, F. and Strumpf, K., 2007. The effect of file sharing on record sales: An empirical analysis. *Journal of political economy*, *115*(1), pp.1-42.

the legitimacy or illegitimacy of the copy, or consider it irrelevant.[328] Therefore, the consumer's primary concern is whether the costs of copying are low. This did not provide much evidence for early theories to prove underproduction or overall social welfare loss.[329] Stan Liebowitz conducted an interesting study in 1985, where he looked at the impact of photocopying on the market for academic journals.[330] Liebowitz introduced the concept of indirect appropriability whereby right-holders (in this case publishers) would get compensated, which included unauthorized copying "by an increase in demand for copiable originals and also by the increase in the total value of the copyrighted material."[331]

In 2005, Varian offered a comprehensive analysis of the changes brought about by digitization and the rise of networked economy.[332] Starting with the observation that digital media introduced unique challenges to economics, he looked at the history and rationale for intellectual property protection which by that time had been the focus of economists who questioned its need in the digital age. Most importantly Varian uncovered legal inconsistencies that do not make any sense to economists:

> Though you can photocopy articles from a journal that you have purchased for your own use, an employee of a company may not be able to photocopy an article in a journal that the company has purchased for its library. (American Geophysical Union v. Texaco, No. 92-9341 (2d Cir. October 28, 1994). However, such restrictions do not appear to apply to educational institutions.
>
> As these examples illustrate, there are many points in intellectual property law that may strike economists as peculiar. Over the years copyright law has

[328] Novos, I.E. and Waldman, M., 1984. The effects of increased copyright protection: An analytic approach. *Journal of political economy*, *92*(2), pp.236-246.
[329] See Towse et al., supra note 68.
[330] Liebowitz, Copying and indirect appropriability: Photocopying of journals, supra note 134, pp.945-957.
[331] Towse et al., supra note 68. Also: Varian, H.R., 2000. Buying, sharing and renting information goods. *The Journal of Industrial Economics*, *48*(4), pp.473-488.
[332] Varian, supra note 323.

> Today most newly created textual, photographic, audio, and video content is available in digital form. Even older content that was not "born digital" can relatively easily converted to machine-readable formats. At same time, the world has become more networked, making it easy to transfer digital content from one person to another. The combination of technological progress in both digitization and computer networking has been a challenge for traditional ways of managing intellectual property.

> evolved in somewhat haphazard ways to meet the challenges of new technologies and business developments […][333]

Moving away from the legal inconsistencies, Varian went on to consider the economics of copying by focusing on the balance between the number of works produced (x) and the number of works consumed, which includes the sharing of copies (y). Without copying, x and y are equal. If copying is permitted y can exceed x. What happens in this case? Varian endorses the concept of appropriability introduced by Liebowitz.[334] In this case, Varian argues, the producer might sell fewer units of the work but he will likely raise the price: "allowing greater freedom to copy will increase seller profits if the value to the marginal consumer (and thus the increased price that the firm can charge) more than offsets the loss of sales."[335] He further explains the tradeoffs between strengthening antipiracy enforcement to support copyright or focusing on pricing - setting prices that reflect the demand. Interestingly, Varian then factors in what he calls *heterogeneous values*. We cannot assume that every individual values intellectual works in the same way – in other words they are not all willing to pay the same price. A public goods problem is created: "When individuals have heterogeneous valuations for the work, the decision about whether to acquire a work that can subsequently be shared becomes a non-trivial public good problem." [336] He examines the group demand function, which simply means the ability of groups to solve the public goods problem by various forms of formal and informal group contributions. Thus, he introduces his analysis on sharing institutions such as video rental stores and libraries:

> […] the group has some way to "solve" the public goods problem and elicit contributions from the members that cover the cost of the item being purchased whenever the sum of the valuations is greater than that cost. This specification makes sense for, say, household members jointly trying to decide whether to purchase a DVD, or a librarian that is trying to decide whether to buy a book for the patrons of the library. To the extent that the librarian is familiar with the tastes of the patrons, the relevant number is the sum of the valuations of the

[333] Ibid.
[334] Liebowitz, supra note 134.
[335] Varian, supra note 323.
[336] Ibid.

> borrowers. […] the item will only be purchased if the value to the member of the group that values the item least exceeds the cost that he or she has to pay. This is motivated by a sharing institution like a video rental store. The store has to set a uniform rental price, and that price must reflect the value of the marginal purchaser.[337]

It follows that in some cases depending on the transaction costs of sharing and the pricing policy of the seller, it makes economic sense for individuals to form groups in order to purchase a copy. We will continue zooming in on the issue of copying books and in particular the line drawn between legitimate and illegitimate copying: copying books outside of a professional printing market, which publishers normally rely on, the availability of photocopying machines and the advent of scanners which have changed the picture. Copying books on a small scale has become mainstream technology and is cheap.

Photocopying machines, or as they were called in the seventies 'xerox machines', had troubled copyright long before digitization, which is basically an advanced method of copying through scanning.[338] The technology is similar; it is simply a matter of advancement and scale. No one back then seems to have argued that the wide distribution of photocopying machines that reached schools, libraries, the workplace and even homes was as revolutionary as digitization is thought of today. Yet there were already discussions on the effects that photocopying was having on copyright. Further discussions engaged with the question of how copyright laws should respond to the technological developments; these were not far off of todays' discussions. For example, Breyer's 1970 article, *A study of copyright in books, photocopies, and computer programs*, was written at a time that the use of photocopying machines in libraries, among other places, was debated.[339] At that time *Williams & Wilkins Co. v. United States* was a case pending before the courts.[340] The subject matter of this litigation is relevant as it dealt with the photocopying of journals for research purposes.

---

[337] Ibid, referring to his previous work: Varian, supra note 323.
[338] Technically digitization also covers the typing of content in forms readable by computers – See also the interchangeable term *digitalization*.
[339] Breyer, supra note 151, pp.281-351. See also the earlier work of Nimmer, M.B., 1968. *Project New Technology and the Law of Copyright: Reprography and Computers; Foreword: Two Copyright Crises*. 5 UCLA L. Rev. 939.
[340] 487 F.2d 1345.

The case was concerned about whether and to what extent photocopying scientific articles could be considered fair use. Williams & Wilkins Company, the plaintiff, was a publisher of medical journals and claimed that the National Institutes of Health (NIH) and the National Library of Medicine (NLM) infringed the plaintiff's copyright by making unauthorized photocopies of articles from the medical journals. While the trial court found the defendants liable for infringement, the federal court disagreed and upheld the libraries' fair use right, holding that "in any event, the hitherto uncodified principles of "fair use" apply to printing, reprinting, and publishing, as well as to copying."[341] The facts of this case were such that copying was somehow limited to 'non-excessive' requests and a maximum of fifty pages. However, the court added that even photocopies of entire works without limitation can be considered fair use.[342] The court carried out a detailed balancing analysis that included looking at subscription fees, patrons' usage and so forth. The final outcome in favor of fair use was also critically influenced by two additional factors: first, the non-profit use of the copies and, second, the fact that photocopying was generally accepted as a practice and had been done for years. This outcome was affirmed by the Supreme Court and did not grant certiorari.[343]

Almost five decades later a similar question has returned to the courts but this time it concerns electronic course reserves at Georgia State University. The plaintiffs are three prominent publishing houses: Oxford University Press, Cambridge University Press, and Sage. The case started in 2008 and the District Court issued its first decision in 2012, which for the most part upheld the university's fair use claims. In 2014, on appeal the federal court disagreed with the lower court's reasoning and the case was remanded. In March 2016, the District Court issued an order, following the deferral court's reasoning but again found most of the university's

---

[341] Breyer, supra note 151.

[342] *Williams & Wilkins Co. v. United States*, See 420 U.S. 376, at Part III of the decision:

> It has sometimes been suggested that the copying of an entire copyrighted work, any such work, cannot ever be "fair use," but this is an overbroad generalization, unsupported by the decisions n12 and rejected by years of accepted practice. The handwritten or typed copy of an article, for personal use, is one illustration, let alone the thousands of copies of poems, songs, or such items which have long been made by individuals, and sometimes given to lovers and others. Trial Judge James F. Davis, who considered the use now in dispute not to be "fair," nevertheless agreed that *a library could supply single photocopies of entire copyrighted works to attorneys or courts for use in litigation.* […] There is, in short, no inflexible rule excluding an entire copyrighted work from the area of "fair use." Instead, the extent of the copying is one important factor, but only one, to be taken into account, along with several others.

[343] Id, 377.

actions to be fair use.[344] Meanwhile, similar issues have been raised in more courts around the globe. Striking are the examples of two recent and much discussed cases in India and in Canada: (i) *The Chancellor, Masters & Scholars of the University of Oxford & Ors. v Rameshwari Photocopy Services & Anr.* (CS(OS) 2439/2012) 16 Sep 2016, Delhi High Court and (ii) *The Canadian Copyright Licensing Agency ("Access Copyright") v. York University* (2017 FC 669). For the purpose of this analysis we will complete the discussion on the basis of the American cases as they demonstrate the evolution of such type of fair use cases within the same legal system. The issues around the Indian and the Canadian case remain very current and relevant.

Looking at the facts of each case and also the debates in the two American cases, one sees that the change of the technology did not radically change the arguments: from photocopying, which could result in relatively small distribution, we went to scanning and e-reserves where digital resources are shared in a controlled manner but neither of these cases involved wide online distribution of materials. Therefore, in such cases the debate came down to proportionality: how much copying is considered acceptable within the context of professional libraries, academic institutions and similar bodies. The days of so called 'xeroxing' do not seem to be radically different to those of scanning, especially when it is limited to small scale copying. The discussion however shifts radically in the case of copying on a larger scale.

As already discussed, mass copying – mass digitization – brings larger discussions to the table related to long-term and wide-spread potential for access to knowledge. At the same time the challenges that mass digitization poses to copyright are definitely greater than the one posed by photocopying machines or e-reserves. The following chapter constitutes the core of this Part II of the dissertation. It describes, in chronological order, the most important digitization initiatives aspiring to create digital libraries and the relevant case-law, mostly started from such initiatives, affecting the future of digital libraries.

[344] The first two decisions for the case are *Georgia State University v. Becker*, 863 F. Supp. 2d 1190, 1209 (N.D. Ga. 2012) and *Cambridge Univ. Press et al v. Patton*, 769 F.2d 1232 (11th Cir. 2014). On remand, the District Court makes an elaborate case-by-case assessment of the fair use factors for every book assigned, in a 220 pages order (US District Court for the Northern District of Georgia, Atlanta Division, Civil Action No 1: 08-CV-1425- ODE, March 31, 2016). The case is pending on appeal before the Eleventh Circuit (Case No. 16-15726: *Cambridge University Press, Oxford University Press, Inc., and Sage Publications, Inc. v. J. L. Albert, et al*.)

CHAPTER SIX - A TIMELINE OF DIGITIZATION INITIATIVES AND THE RELEVANT CASE-LAW AFFECTING DIGITAL LIBRARIES

A crusade seems to be taking place in the western world today where various stakeholders expressed interest in digitizing books on a mass scale and at creating digital libraries. Tracing the developments over the past three decades, one encounters familiar clichés. Americans have cited their founding fathers who spoke eloquently about freedom of information and spreading ideas; and Europeans have talked about a New Renaissance. Digitization of literary works, however, leading to smaller or bigger digital library projects began more than forty years ago, which is several decades before any policy discussion in government or in other legal and policy *fora*. Our story begins in 1971 at the Materials Research Lab at the University of Illinois with Michael Hart's Project Gutenberg. Michael Hart 'invented' e-books and started what is arguably the first digital library. The timing is quite important: packet-switching and the connection of remote computers proved successful only two years earlier, in 1969, which can be called the birthdate of the Internet (the original ARPANET).[345] The first public demonstration of the "new network technology" took place later, in October 1972, when Robert Kahn demonstrated the ARPANET at the International Computer Communication Conference.[346] It is not accidental that Project Gutenberg, the first provider of eBooks, and the Internet are parallel stories or, depending on how one interprets them, intersecting stories.

This chapter creates a timeline of digitization initiatives leading to digital library projects, including the relevant litigation around or related to some of these initiatives. Integrating the case-law, it explores the limits of libraries' engagement with digitization given both the legislation and litigation. For organization purposes, the timeline is split in two sections: it starts with the US followed by the European initiatives. In reality the two timelines run in parallel, or rather merge, as indicated by the dates. The chapter concludes with a preliminary assessment of libraries' position amidst the 'digitization rush'.

[345] Leiner et al., supra note 307.
[346] Ibid.

6.1 THE AMERICAN PROJECTS AND RELEVANT LITIGATION

6.1.1 Project Gutenberg (1971)

Project Gutenberg proudly claims that it is the first provider of free electronic books, or eBooks. Its history and philosophy is tied to the work and vision of its founder, the late Michael Hart who invented electronic books in 1971.[347] As the story goes, Hart was given access to a Xerox Sigma V mainframe computer at the Materials Research Lab at the University of Illinois and obtained an account of a virtually unlimited amount of computer time, valued by him as worth "$100,000,000 of computer time."[348] To use this time Hart's idea was the following: items that can be entered into a computer are then indefinitely replicated so that an infinite number of copies can be available and everyone in the world can access them. As he saw it: "Everyone in the world, or even not in this world (given satellite transmission) can have a copy of a book that has been entered into a computer."[349]

Hart started typing in works, which he then made available online. The first work that he typed in was the US Declaration of Independence. The goal was to make material available in forms that most computers and programs can read and also to make the text searchable. Therefore, texts needed to be presented in "Plain Vanilla ASCII" which is e-text, or else plain text file. The features of Project Gutenberg's digital library are fairly simple. The website, where the library is located, is accessible freely by all Internet users and no subscription is required. The works are available to view and read in digital form, and now in audio form. The input of works is based on an open distributed network of volunteers who can contribute as 'distributed proofreaders'.[350] In general, the project prompted digitization in text-format and was concentrated on books that were already in the public domain.[351] The term digitization is used here *lato sensu* to include the manual typing in of text to make it readable on a computer and by

[347] Hart is also credited as one of the first information providers of ARPANET. See Schofield, J. *More about Project Gutenberg's Michael Hart, RIP*, September 14, 2011 blogpost available at http://www.zdnet.com/article/more-about-project-gutenbergs-michael-hart-rip/.

[348] See the history of the project available at its webpage: https://www.gutenberg.org/wiki/Gutenberg:The_History_and_Philosophy_of_Project_Gutenberg_by_Michael_Hart.

[349] Id.

[350] "Distributed Proofreaders provides a web-based method to ease the conversion of Public Domain books into e-books. By dividing the workload into individual pages, many volunteers can work on one book at the same time, which significantly speeds up the creation process." [from their webpage http://www.pgdp.net/c/].

[351] History of the project, supra note 342. As Hart explained: "Our eventual goal is to provide Public Domain Etext editions a short time after they enter the Public Domain. Of course, the period before a copyrighted work entered the Public Domain was extended from 28 years (with a 28 year extension available) to 50 years more than the life of the author, so this put a kink, to put it mildly, into our plans."

a computer. While Hart digitized works by literally typing them into his computers, today the term usually refers to works that exist in paper form and are then digitized (or 'digitalized') with advanced scanning technology. It is worth noting that most works these days are actually born-digital and exist in various digital forms and in paper as they are printed.

According to Marie Lebert, an information specialist writing about the history of Project Gutenberg, Hart's project got attention at the international level when the Internet became popular, thus around the mid-90's.[352] It was during that period that, as she documented, the number of electronic books rose from 1,000 in August 1997 to 25,000 in April 2008. Today, over 50,000 e-books are available on the website.[353] Project Gutenberg reached the milestone of 50,000 e-books in 2015, when Google, as we will shortly see, had already entered the business of mass digitization and by then had digitized millions of books.[354] What makes the Project Gutenberg initiative so special, even if it lacked the scale of Google and its successors, is that it was the very first project to actually engage in digitization, also when data storage was not as technologically advanced as it is today. And, most importantly, it was the first to envision and then communicate to the world the value of digitization. Equally important and rather symbolic was the explicit mandate to serve the public by providing public domain texts available universally in computable form. Legally, it was thus quickly registered as a non-profit organization and has to this day maintained its idealistic character, while widely recognized as the first digital library.

The project had an interesting self-understanding of a digital library that has been documented in its 'history and philosophy' page available online. Most probably drafted in its entirety by Hart, it explains the logic behind the selection of the first texts included in the Gutenberg catalogue. It states:

> There are three portions of the Project Gutenberg Library, basically be described as: Light Literature; such as Alice in Wonderland, Through the Looking-Glass, Peter Pan, Aesop's Fables, etc. Heavy Literature; such as the Bible or other

[352] Lebert, M. The Project Gutenberg EBook of Project Gutenberg (1971-2009), etext available at http://www.gutenberg.org/files/31632/31632-0.txt.
[353] For the progress of digitization and the relevant statistics see the project's blog with data as of July 20, 2011: http://www.gutenbergnews.org/statistics/.
[354] Hoffelder, N., Milestone: Project Gutenberg Releases eBook #50,000, October 5, 2015 post available at http://the-digital-reader.com/2015/10/05/milestone-project-gutenberg-releases-ebook-50000/.

> religious documents, Shakespeare, Moby Dick, Paradise Lost, etc. References; such as Roget's Thesaurus, almanacs, and a set of encyclopedia, dictionaries, etc. The Light Literature Collection is designed to get persons to the computer in the first place, whether the person may be a pre-schooler or a great-grandparent. We love it when we hear about kids or grandparents taking each other to an etexts to Peter Pan when they come back from watching HOOK at the movies, or when they read Alice in Wonderland after seeing it on TV. We have also been told that nearly every Star Trek movie has quoted current Project Gutenberg etext releases (from Moby Dick in The Wrath of Khan; a Peter Pan quote finishing up the most recent, etc.) not to mention a reference to Through the Looking-Glass in JFK. This was a primary concern when we chose the books for our libraries. We want people to be able to look up quotations they heard in conversation, movies, music, other books, easily with a library containing all these quotations in an easy to find etext format. [355]

Based on this quote, the starting point for book selection is similar to that of a community library. The project's main focus was to serve a community that uses computers in similar and complementary ways to community libraries at the service of patrons, young and old. The possibility of data mining is also clearly a priority, which is not surprising since this is one of the primary advantages of a digital compared to the brick library.

Michael Hart died on 6 September 2011 but his legacy did not. In addition to the several beautiful statements and essays explaining his vision and the project,[356] there is available online a discussion that he had in May 2011 with Bewster Kahle of the Internet Archive and Michael Keller, Stanford University librarian, on the National Public Radio on the topic of digital libraries.[357] The radio interview of the three men is particularly revealing since each speaker shares his motivation and goals. Michael Hart appears dedicated to making as many works as possible available online to every person in the world. He also expresses his frustration with the copyright extensions that resulted in longer waiting periods for the works to enter the public

[355] History of the project , supra note 342.
[356] Obituaries were written by many of the international news magazines, including the New York Times, the Economist and the Guardian
[357] *Libraries enter the digital age*, National Public Radio, Science Friday podcast, May 11, 2011, transcript and sound available at http://www.npr.org/templates/transcript/transcript.php?storyId=10136882.

domain. In a similar fashion, Bewster Kahle appears dedicated to his own project, the Internet Archive, which we will look at directly, having the same vision of making works widely available across the world. The third interviewee, Michael Keller, who was involved with the Google Books project in which Stanford had played a key role from the outset, expressed a slightly different view. He advocated for 'deep search' in books as data rather than free access to content, which was the main driving force of Hart and Kahle. Hart will remain an iconic figure in the history of digital libraries. Time will show whether he will stand in stark contrast to future leading figures of similar initiatives or whether he indeed set the tone for access to knowledge and to a common literary wealth online.

As we continue, we will take a closer look at all the subsequent projects, including the Internet Archive and Google Books, and the relationship between them. The Internet Archive is the next historically important example, which one can see as complementary in its targets to Project Gutenberg, and is perhaps the conceptual continuation of what Michael Hart had begun.

6.1.2 Internet Archive (1996)

The Internet Archive operates as a non-profit organization which is dedicated to building an 'Internet library.' Its mandate is to create an archive, which is basically a backup, of all Internet content and thus preserve the born-digital material. It is essentially a massive programming project that aims at preserving and archiving all our digital history. The Archive was founded in 1996 by Brewster Kahle, a computer engineer and digital librarian. Its headquarters are in San Francisco, located in an old Christian Science church, a pristine white building whose classic architecture is portrayed on the organization's logo.

Besides archiving Internet content and similar to Project Gutenberg, it also engages in the creation of a digital library of literary works. Since 2005 this initiative has been collaborating with libraries, which are mostly located in North America but also elsewhere in the world, to build digital collections. Today, having established two decades of successful operations, it collaborates with institutions, such as the Library of Congress, the Smithsonian, universities and libraries in the US and in Canada, and offers access to a significant number of collections. At this point, the Internet Archive together with the Open Library, a library originally developed for the print disabled community but later merged with the Archive, offers over 8,000,000 e-books and texts.

The Internet Archive was conceived in a way that resembles a traditional library, arguably more than Project Gutenberg. What is distinctive about the founder of the Internet Archive is how he has always been concerned about the role and the responsibility of a digital librarian. In the late 1980's Brewster Kahle was involved in the development of the Wide Area Information Server (WAIS), a client-server text searching system using "Information Retrieval Service Definition and Protocol Specifications for Library Applications" (Z39.50:1988) to search index databases on remote computers.[358] In 1992 he published a piece on the ethics of digital librarianship, where he discussed the responsibilities that digital librarians have in securing the searching information in their possession about their patrons. [359] This article was one of the first of its kind. Since its founding in 1996, he saw and shaped Internet Archive as a digital library project and advocated for the benefits of digitization and public access.[360] He has also been quite critical of the Google Books project detailed below, and voices concern mostly about accessibility and how to reconcile private interest motives and a cause that he deems to be clearly public.[361]

One of the factors that makes the Internet Archive different is its insistence on keeping a physical version of each book that is digitized by them. It also aims to be an archive of books in the traditional sense. The project's success makes it of utmost relevance, with Brewster Kahle reporting ten million downloads of books per month in a 2012 interview.[362] Like Michael Hart, Kahle is also an iconic figure and a visionary that has thought long and hard about and advocated for online access as a global common good for humanity to enjoy. In all, the two first non-profit projects to engage in the creation of online accessible digital libraries share a crucial common

[358] Ferguson, S.J. and Hebels, R., 2003. *Computers for Librarians: An Introduction to the Electronic Library*. Elsevier.
[359] Kahle, B. *Ethics of Digital Librarianship*, Thinking Machines (now brewster@archive.org), online available at https://archive.org/about/ethics_BK.php.
[360] Benson, H., A man's vision: world library online - Brewster Kahle hopes to realize his 25-year dream of an international book archive, *San Francisco Gate*, November 22, 2005 online available at http://www.sfgate.com/news/article/A-MAN-S-VISION-WORLD-LIBRARY-ONLINE-Brewster-2593527.php. See also Gillmor, D., The Creator of the Internet Archive Should Be the Next Librarian of Congress, *Slate*, September 10, 2015, online available at http://www.slate.com/blogs/future_tense/2015/09/10/brewster_kahle_creator_of_the_internet_archive_should_be_the_next_librarian.html.
[361] Kahle, B., How Google Threatens Books, The Washington Post, May 19, 2009, online available at: http://www.washingtonpost.com/wp-dyn/content/article/2009/05/18/AR2009051802637.html.
[362] *The Internet Archive - a chat with Brewster Kahle and a little look inside its book archive*, interview online available at https://www.youtube.com/watch?v=M6rsPgnA87Q.

denominator: a dedication to librarianship and to offering wide access. The same mandates are not necessarily the target of private companies that would subsequently engage in digitization.

To better explain an underlying tension between these two non-profit initiatives and subsequent digitization efforts by private actors, we will first take a look at the details of private initiatives. The focus is primarily on Google whose predominance in the game is, I think, associated with the very history of the company and the work of its founders in the laboratories of the Stanford computer science department.

6.1.3 The Stanford Digital Library Project (1994-2004)

The *Stanford Digital Library Project*, or *Stanford Integrated Digital Library Projec*t, was a research project mainly located in the computer science department at Stanford University and supported by the National Science Foundation under Cooperative Agreement IRI-9411306. DARPA, NASA and the industrial partners of the project provided funding for this agreement. As can be seen from the National Science Foundation award abstract for the cooperative agreement, the project was funded to "develop the enabling technologies for a single, integrated and "universal" library, proving uniform access to the large number of emerging networked information sources and collections." [363] The grant started in 1994, which was the first phase of the research. The second phase, running from 1999 to 2004, was the Stanford Digital Library Technologies project, which aimed at demonstrating the technologies developed on the then emerging California Digital Library.[364]

---

[363] Principal investigator was Hector Garcia-Molina, Professor at the Departments of Computer Science and Electrical Engineering at Stanford University. The NSF award with details for the project is available at http://www.nsf.gov/awardsearch/showAward?AWD_ID=9411306 and older information available online are still captured at http://webcache.googleusercontent.com/search?q=cache:http://www-diglib.stanford.edu/&gws_rd=cr&ei=RoD1VruXC4H8POT8rcAC
Available is also the summary of the project:

> The goal of the Stanford Digital Library Technologies Project is to design and implement the infrastructure and services needed for collaboratively creating, disseminating, sharing and managing information in a digital library context. The Stanford Digital Library Technologies Project was funded from three coordinated proposals, from The University of California at Berkeley UCB, the University of California at Santa Barbara UCSB, and Stanford University. One of our major goals is to demonstrate our technologies on the emerging California Digital Library, CDL and to implement and evaluate these technologies on a testbed system to be built with the help of the San Diego Supercomputer Center, SDSC. All three projects together yield a synergistic and comprehensive digital libraries project.

[364] Ibid.

Very similar to the story of the development of the Internet, academic research together with government funding were key factors for the development of the algorithms that ultimately made content search on the vast web more robust and more meaningful. This project produced important research results going exactly in this direction. Again similar to the development of the Internet, it was a coordinated project involving more than one academic partner together with industrial partners.[365] The ultimate goal, however, was far 'bigger' than digital search; the project aimed at providing *an infrastructure* to interconnect *heterogeneous* and *autonomous digital libraries*. As seen in the award abstract cited above, the vision was the so-called *universal library*.[366]

It is worth looking at the scientific papers that were published while the project was underway to understand why it held the digital library title and what exactly the participants understood by it, which consequently gave rise to their aims. Among them a paper entitled *Stanford Digital Library Metadata Architecture* that was published in 1997 explains in more detail the project participants' vision of digital libraries:

> In the Stanford Digital Library project, we view long-term digital library systems as collections of widely distributed, autonomously maintained services. Of course, a digital library system must include services that allow users to search over collections of information objects. Examples of searchable collections include traditional library collections, digital images, e-mail archives, video, online books, and scientific article citation catalogues […] While searching services are valuable, they are not the only kind of service in the digital library of the future. Remotely usable information processing facilities are also important digital library services [providing]

[365] According to the old webpage of the project, these were the University of California, Berkeley (Re-inventing Scholarly Information Dissemination and Use project), the University of California, Santa Barbara (Alexandria Digital Earth Prototype Project – ADEPT), the California Digital Library (CDL) and the San Diego Supercomputer Centre (SDSC)
See the partners listed under: http://webcache.googleusercontent.com/search?q=cache:http://ilpubs.stanford.edu:8091/diglib/pub/SponsorsAndPartners.shtml&gws_rd=cr&ei=04D1Vp6RJsX0PfKSg8gO.

[366] "The Integrated Digital Library is broadly defined to include everything from personal information collections, to the collections that one finds today in conventional libraries, to the large data collections shared by scientists. The technology developed in this project will provide the "glue" that will make this worldwide collection usable as a unified entity, in a scalable and economically viable fashion": In the abstract available at the NSF webpage supra note 367.

> support for activities such as document summarization, indexing, collaborative annotation, format conversion, bibliography maintenance, and copyright clearance.[367]

As we see from this excerpt, the project contemplates many autonomous services to be carried out by many autonomous digital libraries that are connected. At the same time, it understands that the role of a digital library is to provide searchable content, broadly defined to include books but also images, e-mail archives and more resources. Besides search, the added value of a digital library is the remote processing and use of information. For that, it was well understood that the functions of traditional libraries, such as indexing and bibliography maintenance, should be somehow replicated in the digital domain. Last but not least, copyright was also already seen as a concern at that point. Their 2000 report, *Building the InfoBus*, which reviews the technical choices of the project shows that the principle investigators were also concerned about interoperability since their goal was to connect distributed independent platforms. Even today, in addition to the question of funding, copyright and interoperability are perhaps the two main issues of concern with respect to the creation of digital libraries.[368]

The whole project apparently terminated around 2004; this is also roughly when Google Books first appeared. It is important to note how Google, the most successful search engine to date, is directly linked to this project. Larry Page and Sergey Brin, Google's founders, presented their search engine in a paper entitled *The Anatomy of a Large-Scale Hypertextual Web Search Engine.*[369] As it becomes clear in the paper, they were involved with the Stanford Digital Library Project although it was mainly Larry Page who was officially part of the project's research team. As they explain, Google started as a result of the research they conducted as part of the project. They created a large-scale search engine making heavy use of the structure present in hypertext and designed to "crawl and index the Web efficiently." They explicate how the key is *crawling* and *indexing* the web and developing a search algorithm with ranking functions. As Google developed, it is safe to say that the way they revolutionized web search is firmly linked to the

[367] Baldonado, M., Chang, C.C.K., Gravano, L. and Paepcke, A., 1997. The Stanford digital library metadata architecture. *International Journal on Digital Libraries*, *1*(2), pp.108-121.
[368] Paepcke, A., Baldonado, M., Chang, C.C.K., Cousins, S. and Garcia-Molina, H., 2000. *Building the Infobus: A review of technical choices in the Stanford digital library project*. Stanford InfoLab.
[369] Brin, S. and Page, L., 1998. The anatomy of a large-scale hypertextual Web search engine. *Computer networks and ISDN systems*, *30*(1-7), pp.107-117.

work and purpose of the Stanford Digital Library Project where they started. It is then no real surprise that Google engaged with digitization and with a library project itself once it had established its position as the most successful search engine worldwide.

6.1.4 Google Books (2004)

Google announced its new *Google Print Service* at the Frankfurt Book fair in 2004.[370] The plan was to incorporate books into the Google search service. By 2005, Google had renamed the service *Google Books.* Google Books, as it is known until today, consists of two projects. One is the 'Library Project' under which Google partnered with university libraries initially Harvard University, Stanford University, University of Michigan and Oxford University, as well as the New York public library to digitize their collections. Since then, many more library partners have been joined. Google says that it now works with over forty libraries around the world to digitize their collections. On its website, it provides a sample of its partners.[371] Also, many publishers were interested in the program and therefore joined. [372]

On the basis of the agreements that Google signed with libraries, it digitized their collections and in return it gave a digital copy of each scanned book. Google was creating more than one digital copy of the books on its back-up servers. Some of the libraries that Google partnered with allowed the company to scan only their public domain books, while others also allowed Google to scan in-copyrighted books. Participating libraries were not allowed to have digital copies of works from the other libraries and had to agree to abide by copyright laws in their uses of their digital copies. Google owns the digital corpus it has digitized. Moreover, since it has provided its partner libraries with a digital copy of their collections, the libraries each owns part of the same corpus. To this day Google Books is a free service for Internet users. Lastly,

---

[370] See Google's December 2004 announcement, supra note 28. At the same note, also Wyatt E., New Google Service May Strain Old Ties in Bookselling.

[371] Here are some of the additional to the initial partners: Austrian National Library, Bavarian State Library, Columbia University, CIC library partners, Cornell University Library, Ghent University Library, Keio University Library, Lyon Municipal Library, University of California, The National Library of Catalonia, Princeton University, Stanford University, University Complutense of Madrid, University Library of Lausanne, University of Virginia, University of Texas at Austin, University of Wisconsin – Madison.
See the Google Books Library partners available at http://books.google.com/intl/en/googlebooks/library/partners.html.

[372] For the history of Google Books see https://www.google.com/googlebooks/about/history.html
According to this page the first publishers to join the program were: Blackwell, Cambridge University Press, the University of Chicago Press, Houghton Mifflin, Hyperion, McGraw-Hill, Oxford University Press, Pearson, Penguin, Perseus, Princeton University Press, Springer, Taylor & Francis, Thomson Delmar and Warner Books.

next to the 'Library Project', Google Books has a 'Partner Program' under which right-holders provide Google with a printed copy of their book for scanning and decide how much of their work Google is permitted to display. Subsequently, works are displayed with the permission of the copyright-holder.[373] Initially Google shared the revenues from advertisements with copyright-holders until 2011 when it stopped displaying advertisements in connection with the books.[374]

Today Google Books provides the user with an index, a digital catalogue of the scanned books, matching search terms with books. According to the copyright status of each book or the agreement with the copyright-holder, the service gives access to full, less or no content. Each book includes an 'About this book' page that provides basic bibliographic information and links which direct the user to online or offline bookstores and libraries where the book is available. Besides the cases where full view is allowed, limited view allows access to some pages. In other words, snippet view allows a number of snippets to be displayed after a certain term search. Finally, no preview only reveals card catalogue information. Google Books at times operates more like an online reference catalogue database than a library. It would be more accurate, however, to say that it is a mixed reference catalogue and library as it provides full access to a significant number of public domain books or books it is allowed to give access to, just like Project Gutenberg or the Internet Archive. Unlike these two projects however, it digitized much more material, including in-copyrighted, which are equally used to feed the search engine.

The value that Google found in the content of books is an important question that is also a preoccupation of this study. Technological advancements, besides better search results, also include improving services such as text translation that Google offers. In general, it might be fair to say that by getting hold of a very rich corpus of material which was digitized, and thus transformed into computable data, Google has benefited from the functioning of its algorithms in more than one way. We will delve into this discussion in the final chapters but let us continue this timeline with what is perhaps the most vocal copyright litigation that is relevant not only to the Google Books project but also to the future of digital libraries in general. Indeed, when the case against Google reached the courts, its importance rose far beyond the main parties involved.

---

[373] On the basis of Google's submissions to the district court for the *Authors Guild, Inc. v. Google Inc.* case we know that as of early 2012 the partner program included approximately 2,5 million books with the consent of 45,000 right-holders. See *Authors Guild, Inc. v. Google Inc.* - 2013 District Court Opinion, pp. 5-6. [For the full reference to the decisions *see* supra note 4]

[374] Id. at p. 5.

Besides the publicity and the length of the case, the sheer number of intervening parties submitting objections or *amicus curiae* briefs over the course of the litigation is indicative.[375]

6.1.5 The Google Books case before the courts (2005-2015)

Very quickly after the announcement, Google faced two lawsuits. First, a class action brought by an American association of book authors, the Authors Guild, and three individual authors.[376] Simultaneously another action was brought by publishers McGraw-Hill Companies, Pearson Education, Penguin, Simon & Schuster and John Wiley.[377] All suing parties claimed that the act of digitization of works without permission as well as the display of snippets on the Google Books search platform constituted copyright infringement. As we have seen, the company's plan was to bring books into its search service and then allow Internet users to search the texts of the digitized books and see displays of parts, known as 'snippets', of the corresponding texts. It proceeded with the two programs that we already discussed: the 'Partner Program' and the 'Library Project.' Meanwhile, Google was creating more than one digital copy of the books on its back-up servers.

While the publishers sued in their own right, the Authors Guild attempted to represent a significant number of authors, besides the three involved in the suit. As we will see, the lawsuit brought by the publishers was settled along the way. Thus only the Authors Guild continued to litigate their case, which would become known as the renowned Google Books case. It turned out to be a complex case, with two decisions at first instance delivered by Circuit Judge Denny Chin, two unsuccessful settlement attempts, a federal decision delivered by Federal Judge Pierre Leval, and an unfruitful appeal to the US Supreme Court.[378] The case was initially brought before the District Court for the Southern District of New York and was filed as a class action. It is mostly, but not exclusively, due to the class action procedure that this case lasted longer than every other digitization case that we examine. It started in 2005 and ended in 2015. During the ten-year long litigation, there were three court decisions, which will be studied. Two are at the

[375] According to judge Chin's 2011 decision, for example, the objections to the Google Books settlement agreement were *"approximately 500 submissions"* (*Authors Guild et al. v. Google Inc.* – 2011 Opinion Rejecting the Settlement, p.10).

[376] Authors' Guild initial class action complaint, 20 Sept. 2005, available at https://www.authorsguild.org/wp-content/uploads/2008/10/Authors-Guild-v-Google-09202005.pdf

[377] Publishers' initial complaint, 19 Oct. 2005, available at https://www.authorsguild.org/wp-content/uploads/2008/10/McGraw-Hill-v.-Google-10192005.pdf

[378] After the death of the initially appointed Hon. Judge John E. Sprizzo.

first instance and one at the appellate level. We will start with the two decisions by Judge Chin and conclude with the appellate, and now final, decision of Judge Leval.

First came a settlement proposal to the class action. This settlement agreement, known as the Google Books Settlement, was filed twice. The first time was in 2008 and the second amended settlement agreement was filed in November 2009. The court granted preliminary settlement approval on 17 November 2008.[379] The settlement agreement received objections from hundreds of class members and the US Department of Justice. The district court received *amici curiae* in favor and against the amended settlement. A hearing was held in February 2010. In March 2011, Judge Chin delivered his opinion and concluded that the settlement agreement was "not fair, adequate, and reasonable."[380] This was the final rejection of the Google Books Settlement.

With the proposed settlement agreement, the case had transformed from a copyright and fair use case, in relation to the copying of entire books and the making available of snippets of books on a search engine, to a case of an elaborate commercial plan for the exploitation of full-text digitized works in a form of a massive online bookstore.[381] Thus, when Judge Chin issued his decision on the Google Books Settlement, he faced a very broad question. He began by stating that "while the digitization of books and the creation of a universal digital library would benefit many, the ASA [Amended Settlement Agreement] would simply go too far."[382] It is evident that he had associated the project with the vision of a universal digital library from the beginning, which echoes many scholars and commentators.[383] Reading the decision one can understand that the main preoccupation of the judge was the significant competitive advantage that Google would enjoy if the proposed settlement was approved. This is understandable, given the many submitted objections, which included an objection from the US Department of

---

[379] *Authors Guild Inc. v. Google Inc.* Docket: Filing #59, Second Amended Class Action Complaint (October 31, 2008).
For the history of the settlement from the Authors Guild point see the dedicated Authors Guild webpage: http://www.authorsguild.org/advocacy/authors-guild-v-google-settlement-resources-page/ and http://www.authorsguild.org/advocacy/125-million-settlement-in-authors-guild-v-google/
[380] *Authors Guild et al. v. Google Inc.* – 2011 Opinion Rejecting the Settlement, p.45.
[381] *Authors Guild et al. v. Google Inc.* – 2011 Opinion Rejecting the Settlement pp. 5-10.
[382] Id. p.1.
[383] Kelly, supra note 29. Some commentators saw a big threat in this vision, primarily due to the monopoly and the private for-profit organization of the company. See for example, the in-depth commentary of Robert Darnton, "The Library in the New Age," *New York Review of Books*, June 12, 2008, as well as his later series of articles at the New York Review of Books as the Google project and the litigation advanced.

Justice.[384] The main problem with litigating this settlement under a class action procedure was indeed its scope and the inadequate representation of interested and immediately affected parties.[385] The judge, in his decision, cited the comments of the Department of Justice to the court stating that the agreement "is an attempt to use the class action mechanism to implement forward-looking business arrangements that go far beyond the dispute before the Court in this litigation".[386]

After rejecting the settlement, the same judge reviewed both parties' claims on whether the fair use defense applies in this case. He handed down a new decision in November 2013. Thus, the initial fair use defense of Google once again became the main subject in the case. Moving away from the settlement, the court case had returned to the issues concerning search function and snippet views of Google Books. To reach his decision, Judge Chin engaged in a legal assessment of the fair use requirements. Under US copyright law, the fair use exception releases what is a *prima facie* copyright infringement from liability. There are four requirements to be assessed by the courts, which are codified in the US Copyright Code: (i) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes; (ii) the nature of the copyrighted work; (iii) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (iv) the effect of the use upon the potential market for or value of the copyrighted work.[387] US copyright case-law has consistently placed the emphasis on the first and the fourth criteria since they, particularly with regard to the first criterion, emphasize the significance of transformative uses (uses where the original work is being transformed in a manner that adds new value to it and transforms into new expression).[388] Following this precedent, the decision places a strong emphasis on whether the use made of the copyrighted works was transformative.[389] Indeed, the judge found Google's use

[384] *Authors Guild et al. v. Google Inc.* – 2011 Opinion Rejecting the Settlement, p.5.
[385] Id., pp. 11-13. The concerns that ultimately led to the rejection of the settlement were many and related both to the ruling of a class action (adequacy of class notice, adequacy of class representation etc.) and to copyright and antitrust laws, not to neglect the expressed privacy and international law concerns.
[386] Id. p. 21.
[387] 17 U.S. Code § 107.
[388] Of seminal importance for the establishment of the criterion of 'transformativeness' has been the Supreme Court decision *Campbell v. Acuff-Rose Music*, 510 U.S. 569 (1994). See Samuelson, P., 2015. Possible futures of fair use. *Wash. L. Rev.*, *90*, p.815.
[389] He also made the analogy with the thumbnail images of copyrighted photographs in *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1168 (9th Cir. 2007) and in *Kelly v. Arriba Soft Corp.*, 336 F.3d 811 (9th Cir. 2003), and with small sized display of images of posters in *Bill Graham Archives*, 448 F.3d all cases that found

of the books *highly transformative* mainly in view of the abovementioned benefits of the service created.[390] On the basis of this reasoning, the fair use defense was finally upheld in favor of Google.

While reviewing how Google Books works, the new decision discussed the benefits of the related library project, and as a result, the digitization agreements with the libraries.[391] Judge Chin enumerated five benefits, which on this occasion was in a more elaborate manner than his earlier decision in 2011.[392] First, he noted how Google Books is a new and efficient search and reference tool that became essential to researchers and librarians; second, that it promotes data mining and text mining, methods that open up new research opportunities; third, that it expands access to books, especially (i) benefiting groups of print disabled[393] and (ii) enhancing the search capabilities of "remote and underfunded libraries that need to make efficient decisions as to which resources to procure for their own collections or through interlibrary loans."[394] In this part of the judgment it became clear, and clearer than his previous decision, that for the judge access to knowledge has a solid value and must be taken into account in copyright cases. Moreover, it is obvious that he referred to libraries as public service institutions in the servitude of this value. The fourth benefit that the judge stated was how Google Books as a service helps preservation of books and also "give[s] them new life".[395] He was referring to the new digital life that is given to out-of-print books that were until their digitization "buried in library stacks."[396] This phrase is repeated from his 2011 decision and reads as an affirmative endorsement of digitization; a worthwhile process for the sake of the searchability and readability of books. Preservation and accessibility are two notable benefits in their own right, especially when the print versions are close to becoming obsolete. The fifth benefit enumerated by Judge Chin was that because the

---

transformative use of displays similar to that of the Google Books snippets. *Authors Guild, Inc. v. Google Inc.* - 2013 Summary Judgment, pp. 19-20.

[390] *Authors Guild, Inc. v. Google Inc.* – 2013 Summary Judgment, p. 19.

[391] Id. at p. 9 et. seq.

[392] Id. pp. 9-13.

[393] Id. p. 11. On this issue the court cited a letter from Marc Maurer, President of the National Federation for the Blind, saying the following: "Google books provides print-disabled individuals with the potential to search for books and read them in a format that is compatible with text enlargement software, text-to-speech screen access software, and Braille devices. Digitization facilitates the conversion of books to audio and tactile formats, increasing access for individuals with disabilities."

[394] Ibid. In the judge's words: "Google Books facilitates the identification and access of materials for remote and underfunded libraries that need to make efficient decisions as to which resources to procure for their own collections or through interlibrary loans." Here the judge cites the *amicus curiae* brief of the American Library Association.

[395] Ibid.

[396] Id. p.12: "Older books, many of which are out-of-print books are falling apart buried in library stacks, are being scanned and saved."

service helps readers and researchers identify books, it ultimately benefits authors and publishers. The 'About the book' page that appears with the search results offers links to sellers and/or libraries where the book is available. Thus, the judge found that Google Books generates new audiences and ultimately creates new sources of income-making works more visible and to a broader public.[397]

The Authors Guild appealed to the US Court of Appeals for the Second Circuit.[398] Judges on the bench were Judges Pierre Leval, José Cabranes and Barrington D. Parker Jr. The final decision was delivered in October 2015 with Judge Leval writing for the court. It is noteworthy that this judge is also a distinguished expert in fair use cases, having written the highly influential 1990 Harvard Law Review article *Toward a fair use standard*.[399] Similar to the district court, this decision concluded that Google's unauthorized digitizing of copyright-protected works, creation of a search functionality, and display of snippets from those works are non-infringing fair uses. In addition, Google's provision of digitized copies to the libraries that supplied the books is also not an infringement of copyright.[400] Judge Leval set the tone at the outset. He started by stating that "[t]his copyright dispute tests the boundaries of fair use."[401] Reviewing the factual and procedural history of the case, Judge Leval, like Judge Chin, reiterates the beneficial elements of both the Google Books service and the library project, emphasizing, *inter alia,* the fact that the search tool is highly innovative.[402] He contended that the fact that copyright affords potential creators exclusive control is only the means to a specific goal: "to expand public knowledge and understanding."[403] Judge Leval with this decision referred to his elaborate interpretation of the fair use doctrine. He placed an emphasis on how copyright is meant to serve the interests of the public.[404] For the judge, transformative use communicates

---

[397] Ibid.

[398] They appealed on five grounds: (1) that the digital copying of entire books is not transformative use in the meaning of *Campbell v. Acuff-Rose Music, Inc.*, (2) that Google's commercial profit motivation precludes fair use, (3) that plaintiff's derivative rights in search functions are in any case infringed, (4) That Google storing of digital copies exposes to hacking risks, and (5) That Google's distribution of digital copies to participant libraries risks loss of copyright revenues from licensing these libraries.
See *Authors Guild, Inc. v. Google Inc.* - 2015 Second Circuit Opinion (see pp.3-4).

[399] Leval, P.N., 1990. Toward a fair use standard. *Harvard Law Review*, *103*(5), pp.1105-1136

[400] *Authors Guild, Inc. v. Google Inc.* - 2015 Second Circuit Opinion, p. 46.

[401] Id. p.2.

[402] Id. p. 7. To emphasize the enhanced search capabilities of the service, he noted that "this identifying information instantaneously supplied would otherwise not be obtainable in lifetimes of searching."

[403] Id. p.12.

[404] Id. p.15.

something new and different from the original or expands its utility, thus serving "copyright's overall objective of contributing to public knowledge."[405] Eventually Judge Leval, like Judge Chin, also found that the Google Books service does not substantially harm the market of the digitized works. Instead, he held that the service amounts to a transformative use of the digitized copies, providing search capabilities that allow the identification of information about originals that is otherwise unavailable.

The final act in the Google Books case is the Guild's petition against the appellate decision to the Supreme Court. The petition was filed in December 2015. The Supreme Court, however, did not grant certiorari, putting an end to the case in April 2016.[406] We will now continue with another project that in the meantime tried to compete with Google Books and demonstrates some similarities: the 'Open Content Alliance' which started more cautiously or conservatively as far as copyright is concerned and unlike Google's project was not involved in litigation. After that we will move to the HathiTrust Digital Library, a project and a lawsuit that are directly associated to Google Books.

6.1.6 The 'Open Content Alliance' and further industry involvement -Yahoo, Microsoft, Amazon and others (2005)

In 2005, Yahoo attempted to lead a competing plan to make digitized books available and searchable online. Yahoo partnered with the Internet Archive and other stakeholders in a project named 'Open Content Alliance'. Full text accessibility to everyone was the announced target and the members cooperating in the project included public universities, like the University of California and the University of Toronto, and other content providers, such as the National

---

In his own words: "[e]ach factor thus stands as part of a multifaceted assessment of the crucial question: how to define the boundary limit of the original author's exclusive rights in order to best serve the overall objectives of the copyright law to expand public learning while protecting the incentives of authors to create for the public good."

[405] Within the decision Judge Leval expressly referred to *the public good* as defining the boundary limits of exclusive rights afforded by copyright. Citing the US Constitution, Article 1(8), Clause 8, he concluded that while copyright surely protects authors "the ultimate, primary intended beneficiary is the public whose access to knowledge copyright seeks to advance by providing rewards for authorship." In all, he identified fair use as a guarantee to the benefit of the public, since it allows unauthorized copyright in circumstances that copyright's very purpose, which is to expand public knowledge, dictates so.
Id. p.17 and pp. 13-15.

[406] Petition before the Supreme Court (No. 15-849): *The Authors Guild, et al., Petitioners v. Google, Inc.* (Docketed: December 31, 2015). The Supreme Court denied the petition.

Archives of the UK and the European Archive.[407] Providing a full text accessible to everyone was clearly a different target than Google's, which even though it scanned full copyrighted material, restricted their availability to the public. The project was announced approximately around the time when Authors Guild filed their lawsuit against Google for copyright infringement. Thus, the fear of liability and the general lack of clarity on the issue was also an obvious concern for the Open Content initiative. The Open Content Alliance partners announced a more cautious approach, seeking the permission of authors and of course digitizing the public domain books of their partners. The approach was *collection focused* with an overall mandate to also *fill gaps* of existing digital collections in a manner complementary to existing projects of partners and of third parties.[408] The alliance planned to operate as follows: the Internet Archive would host the material digitized by the Archive itself or by other partners. Yahoo would partially fund the digitization projects and then index the digitized material. The University of Toronto and other partners such as O'Reilly, for example, would provide book content. There were additional industry stakeholders connected to the project. Besides Yahoo who invested a large sum for digitization and who would offer online searchability, Adobe and Hewlett Packard were involved in the provision of scanning technology. After launching the alliance, Microsoft also got involved in sponsoring the digitization project.

The Open Content Alliance was publicized by the press as a project led by Yahoo, which most probably done so in order to emphasize the rivalry between Yahoo and Google. The initiative, however, was led by the Internet Archive.[409] The announcement that Yahoo published in October 2005 was basically Brewster Kahle's announcement of the project.[410] The Internet Archive, which had already a comprehensive and fully functioning digitization program, set down the targets and standards with an expressed emphasis on openness and collaboration. In the initial announcement, Kahle explained that the three biggest obstacles facing the partners that

---

[407] Suber, P., The Open Content Alliance, SPARC Open Access Newsletter, issue #91, November 2, online available at: https://dash.harvard.edu/bitstream/handle/1/4552008/suber_oca.htm?sequence=1.

[408] Hafner, K., In challenge to Google, Yahoo will scan books, *The New York Times*, October 3, 2015, online available at http://www.nytimes.com/2005/10/03/business/in-challenge-to-google-yahoo-will-scan-books.html.
See also Suber (supra note 407): "The rivalries among Yahoo, Microsoft, and Google are not only real, but interesting enough to soak up news attention. But reporters who focus on the rivalries are missing half the full story. The OCA project and Google Library are compatible and complementary. Both teams take the position that "the more, the merrier" and they're both right to do so. Both teams work to minimize duplication among the scanned books."

[409] Ibid (Suber).

[410] Kahle B., Announcing the Open Content Alliance, Yahoo Search blogbost, October 2, 2005, available at http://www.ysearchblog.com/2005/10/02/announcing-the-open-content-alliance/.

were necessary to overcome were: (i) costs, (ii) rights and (iii) guidelines for sharing.[411] The involved industry partners would be able to offset some of the costs. Also with Yahoo's involvement the added value would be the availability of the scanned works online via the Yahoo search engine. The alliance however did not give Yahoo a monopoly; more search engines would be able to host the digitized content.[412]

When we read about this new collaboration today, it would seem that the involvement of Yahoo followed by Microsoft were indeed in direct competition with Google. This initiative by Google's competitors amounted to an attempt to cover spaces and include partners that Google had perhaps not yet done. [413] With the Google Books project having already a number of outspoken critics, the planning of this project was framed in many ways antithetically to Google's.[414] The newly formed group, however, was not opposed to collaboration with Google. It is worth citing Brewster Kahle's remarks, which we can say represents the Internet Archive: "The thing I want to have happen out of all this is have Google join in. I know we're dealing with archcompetitors, but if there's room for these guys to bend, by the time my kid goes to college, we could have a library system that is just astonishing."[415]

Writing about the then new alliance, Peter Suber provided an extended commentary about the differences between the Open Content Alliance project and Google Books. Before providing the quote, let us note that the commentary came at a stage when the legality of the Google Books project was still unclear and highly contested. Suber comments:

> Unlike Google Library, the OCA [Open Content Alliance] will only scan copyrighted books when it has the copyright-holder's consent. As a result,

---

[411] Id.
[412] Id.
[413] Suber, supra note 407:

> But of course Yahoo and Microsoft are still rivals apart from OCA and still see ways to make money from their participation in OCA even if that has to be long-term and indirect. The only short-term boost they get over Google from this project is the chance to build better relations with the authors and publishers who dislike the Google project. We should never forget, however, that many authors and publishers think the lawsuits against Google are baseless and harmful, approve the opt-out policy, and love the Google project.

[414] Hafner, supra note 408. See also
Roush, W., Digitize This! Yahoo hopes to trump Google with its Open Content Alliance publishing venture, *MIT Technology Review*, October 20, 2005, online available at https://www.technologyreview.com/s/404830/digitize-this/.
[415] Ibid (Hafner).

> publisher groups that criticized Google Library (AAP, AAUP, ALPSP) have endorsed the OCA.
>
> […] all OCA's digitizing, copying, and indexing will be non-exclusive. If other organizations want to copy and host the texts, they may. If other search engines want to index them, they may. Yes, that includes Google.
>
> […]
>
> For publishers of books under copyright, the OCA is opt-in rather than opt-out. This definitely pleases publishers and others with a conservative interpretation of fair use, but it doesn't follow that the Google opt-out policy is unlawful. Nor does the OCA policy weaken the legal case for the Google policy. Both policies might be lawful. If Google wins in court and vindicates the opt-out policy, then book-scanning projects will have two lawful options and it will be interesting to compare them. For some authors and some publishers, the OCA method will look better. But for the book-scanners themselves, for readers, and for many other authors and publishers, the Google method will look better. The advantage of opt-in is strongest in a period of legal uncertainty, like the present. While Google is moiling through lawsuits, paying lawyers, and risking liability, the OCA can proceed without risk.[416]

In sum, this initiative seen together with the Google Books project already suggests the complex range of stakeholders involved. Early visionaries, such as Brewster Kahle with the Internet Archive, had different targets and stakes than the technological partners with whom they inevitably entered into partnerships. Kahle was, and still is, very outspoken about his goal of the project to create openness and to build a common good for future generations. By contrast, the private stakeholders involved remained competitors. Thus, the outcome of such collaborative efforts must be evaluated through the prism of the dominant competition of the private tech players. This can perhaps better explain the reason why despite all initiatives having compatible goals, the actual collaboration of stakeholders in the friendly and idealistic manner proposed by Kahle was less realistic. This maybe also explains how some initiatives, namely Google's, were successful and others, like Yahoo's, were not. Irrespective of their success or failure, they are

[416] Suber supra note 407.

worth looking at together and studied as parallel digital library alliances forming what can be seen as a 'digitization rush' of the time. News reports from these years (approximately from 2005 to 2008) reveal exactly this competitive race-like situation.[417] A 2008 online news article, for example, reports a "race to the shelf" in such explicit terms:

> Internet giants such as Google, Yahoo!, Microsoft, and Amazon are in the middle of nothing short of a modern-day space race: Who can scan the most and the best books in alliance with the biggest and brightest libraries in the U.S. — nay, the world! — while simultaneously providing print on demand, "find in a library," and "buy the book" links as well?[418]

6.1.7 HathiTrust Digital Library (2008)

HathiTrust is a partnership of university libraries that began in 2008 as a collaborative project between what is now the Big Ten Academic Alliance Universities (formerly the Committee on Institutional Cooperation[419]) and the University of California with a goal of creating a common repository for their digitized collections.[420] Many more research libraries partnered along the way, including all Ivy League universities and most state universities, making HathiTrust perhaps the biggest and most important library consortium of its kind.[421] To digitize their collections, both in copyright and public domain materials, the university libraries partnered with Google, as seen above, the Internet Archive, Microsoft and also combined in-house initiatives. The goal of the partnership was "to build a comprehensive archive of published literature from around the world and develop shared strategies for managing and developing their digital and print holdings in a collaborative way."[422] Thus, the digital archive that is being built is co-owned and co-managed. While its primary purpose is preservation, it also aims at assisting scholars with

---

[417] Shields M., Tech giants unite against Google, BBC News, August 21, 2009, online available at http://news.bbc.co.uk/2/hi/technology/8200624.stm.

[418] Ashmore, B., The Race to the Shelf Continues: The Open Content Alliance and Amazon.com, 16(1) *Searcher*, January 2008, online available at http://www.infotoday.com/searcher/jan08/Ashmore_Grogg.shtml.

[419] The academic consortium of the universities in the 'Big Ten Conference' – for a list of the member universities see http://www.cic.net/about-cic/member-universities.

[420] See the information available at the consortium's website at https://www.hathitrust.org/partnership and https://www.hathitrust.org/community.

[421] Ibid.

[422] Ibid – at the 'Mission and Goals' section available at https://www.hathitrust.org/mission_goals.

access in accordance with the applicable copyright and contract rules. The collection currently consists of 13,874,892 volumes (book titles and serial titles), of which approximately 60% are under copyright and the rest is in the public domain.

On the basis of the digital corpus created, a number of services are also on offer. Among its most important contributions, HathiTrust provides access to print disabled groups. As seen from the declared purpose and the current output and standing, the initiative is very clear on its public service commitment. Jeremy York, project librarian at HathiTrust, explains the three overall principles or "overarching considerations" that have the biggest impact on the design and operation of the project: first, scale; second, the connection between preservation and access ("a philosophical belief that the value of preservation is gained through access – that there is no value to a community of preservation without access"); and, third, openness.[423] With respect to scale, the HathiTrust model is based on the collaboration of numerous institutions that all contribute with their own wealth of materials. With respect to preservation and access, HathiTrust provides full access to public domain works, has a fully searchable depository of both public domain and in-copyright works and directs to the location of the physical books in the partner libraries which are fully accessible by patrons even if under copyright. Last, the openness principle refers to content formats, hardware and software and most importantly organizational structure. In sum, the foundational stones of the HathiTrust's philosophy and operations, are based on the notion that "preservation is a social and a collaborative activity."[424]

6.1.8 Litigation against HathiTrust (2011 – 2014)

In 2011, some years after the case that started against Google, Authors Guild also started an action against HathiTrust. Thus, the second American case relevant to digitization becomes the case *Authors Guild, Inc. v. HathiTrust.*[425] As seen above in the Google Books case, Google's distribution of digital copies to participant libraries was also adjudicated. The final decision,

[423] York, J., A Preservation Infrastructure Built to Last: Preservation, Community, and HathiTrust. Proceedings of The Memory of the World in the Digital Age: Digitization and Preservation. An international conference on permanent access to digital documentary heritage, 26-28 September 2012, Vancouver, British Columbia, Canada, edited by Luciana Duranti and Elizabeth Shaffer (UNESCO 2013).
[424] Id. See on the same point Christenson, H., 2016. Building a US Federal Government Documents Collection in HathiTrust. *Collaborative Librarianship*, *8*(3), p.5.
[425] For the full references of the case in all instances see *supra* note 5.

however, excluded secondary liability against Google for the actions of its partner libraries.[426] The two decisive factors with respect to secondary or participatory liability were, first, the fact that the libraries already owned the books that were digitized and, second, that they were bound by contractual obligations to Google, whereby they committed to using their digital copy in a manner consistent with copyright law. In spite of this, libraries that belonged to the HathiTrust consortium were sued independently. This case against HathiTrust almost ran in parallel to the case against Google.[427]

At the first instance the case against HathiTrust ruled strongly in favor of the digital library consortium. In a nutshell, plaintiffs asserted claims for copyright infringement for unauthorized reproduction and unauthorized distribution of books owned by the universities. Quoting the judge, the facts before him at the time were as follows: "the HathiTrust partnership is in the process of creating a shared digital repository that already contains almost 10 million digital volumes, approximately 73% of which are protected by copyright."[428] The district judge, the late Harold Bear Jr., decided that all four factors of the fair use exception were satisfied in the case. Noting that Google's use of the digital works was the matter of a separate lawsuit, Judge Bear defined the scope of this decision to adjudicate on the uses made by the HathiTrust Digital Library.

The copyright infringement claim concerned, first, HathiTrust's 'Orphan Works Project' and, second, HathiTrust's uses of the digitized corpus. The Orphan Works Project aimed at making full copies of orphan works available through HathiTrust to students, faculty and library patrons.[429] When the litigation started, the project was in its initial stages and was still investigating processes to determine which works could be included as orphans. Thus, at the time

[426] Judge Chin's 2013 decision for example held that providing libraries with digital copies of their owned collections is consistent with copyright law as libraries' lawful uses of the works are advanced. Libraries are enabled not only to preserve but also to use the copies in further transformative ways creating their own full-text searchable indexes, innovating with access for groups of print-disabled and so forth.

[427] It is also interesting to note that at the appellate level the two benches overlap as both cases were heard by the Second Circuit. In terms of timing, the HathiTrust case's bench was formed and issued its decision earlier: in 2014 we have the final appellate decision for the HathiTrust case from Judges Walker, Cabranes and Parker (Parker writes for the court) and in 2015 we have the final decision for the Google Books case (analysed above) from Judges Leval (writing for the court), Cabranes and Parker.

[428] *Authors Guild, Inc. v. HathiTrust*– 2012 District Court Opinion, p. 2.

[429] The Orphan Works Project was intended to investigate the available author and publisher information about potential orphan books. If neither the author nor the publisher could be located and contacted for a book and the book was out of print, it would be flagged as potentially orphan. If at any time a copyright owner was identified and located, the book would consequently be removed from the list of orphan candidates. The universities had announced that they would then make these identified orphans available for full view to their communities.

of the decision, the judge considered the claim against the project unripe; it was a claim that the project *would* infringe copyrights of the plaintiffs and others and was therefore deemed inactionable.[430] On the question of the uses of the digitized works, the court concentrated on the three kinds of uses that the HathiTrust Digital Library allowed: (i) full-text searches, (ii) preservation and (iii) access for people with certified print disabilities. What had a huge bearing on the case was the importance of the third use. Blind scholars together with the National Federation of the Blind demonstrated how beneficial the HathiTrust Digital Library had been to them. Thanks to the library, people with disabilities could read digital texts independently through special software, navigate texts, and above all access more material than ever before. Academic participation by print-disabled students was thus revolutionized.[431] For all the three uses, the defendants' main defense was fair use.[432] The court applied the four fair use factors to the case and concluded that:

> The enhanced search capabilities that reveal no in-copyright material, the protection of Defendants' fragile books, and, perhaps most importantly, the unprecedented ability of print-disabled individuals to have an equal opportunity to compete with their sighted peers in the ways imagined by the ADA [American Disabilities Act] protect the copies made by Defendants as fair use to the extent that Plaintiffs have established a *prima facie* case of infringement […][433]

In his decision, the judge explicitly embraced the digitization project and its benefits, taking into account the uses as analyzed. Most notably he expressed: "I cannot imagine a definition of fair use that would not encompass the transformative uses made by Defendants' MDP [Mass Digitization Project] and would require that I terminate this invaluable contribution

[430] *Authors Guild, Inc. v. HathiTrust*– 2012 District Court Opinion, pp. 11-12.
[431] Id., pp. 3-4.
[432] Notably Judge Bear expressly affirmed the possibility for libraries to use the fair use defense under Section 107 of the Copyright Act in addition to the special libraries' exception of Section 108. See the analysis at pp. 12-13 of the Opinion.
[433] See in detail *Authors Guild, Inc. v. HathiTrust*– 2012 District Court Opinion, p.21. Let us note that the American Disabilities Act includes a mandate of equal access to copyrighted information for print-disabled individuals.

to the progress of science and cultivation of the arts that at the same time effectuates the ideals espoused by the ADA [American Disabilities Act]."[434]

At the appellate level, the second circuit confirmed the lower court's decision. This decision not only reaffirmed the fair use defense in favor of HathiTrust but also deemed it as "not disputed that in order for full-text search to be possible the libraries must first create digital copies of entire books."[435] Judge Parker for the appellate court wrote a clear and strong decision in favor of the HathiTrust library. He seems to fully understand the importance of this particular digital library, its comprehensive coverage and the large institutional participation, which at the time of the decision HathiTrust had eighty member institutions. The court held: "HDL [HathiTrust Digital Library] contains digital copies of more than ten million works, published over many centuries, written in a multitude of languages, covering almost every subject imaginable"[436]

Similar to the district court, the usefulness of the HDL for people with disabilities played an important factor, which led to the conclusion that the use is transformative. Another issue discussed for the first time at the appellate level was the number of copies that the University of Michigan and the University of Indiana held for the purposes of HathiTrust. The appellants contended that the copying was excessive: HathiTrust created and maintained digital copies at four different locations; the primary server is located at the University of Michigan, a mirror server at the University of Indiana and two encrypted backup tapes disconnected from the Internet in separate secure places on the University of Michigan campus.[437] The court, however, found that the number of copies stored in the different servers were necessary and reasonable for the purposes of safeguarding the digital corpus from data loss and also balancing the web traffic of users on a single site. Indeed, the storage of copies is essential for preservation, which has an intrinsic value, in order to guarantee access to the material for future generations, especially after the material enters the public domain.[438]

Overall, HathiTrust dodged liability having won the case not only because of its specific project but also for what libraries can generally do with their collections. After this case the following uses that justify digitization by libraries include: first, the availability of full-text

[434] *Authors Guild, Inc. v. HathiTrust*– 2012 District Court Opinion, p.22. Emphasis in the text is added.
[435] *Authors Guild, Inc. v. HathiTrust*– 2014 Second Circuit Opinion, p.18.
[436] Id. p.6.
[437] Id. p.21.
[438] See the analysis on p.31 of the Opinion.

search databases; second, providing access to people with disabilities and, third, preservation. Digitization for such reasons is justified and lawful under the fair use doctrine. The case is revolutionary at least with respect to preservation, pushing the boundaries of the current the US Copyright Act exception under Section 108, discussed in chapter four.[439]

Continuing the timeline, we will proceed with more American initiatives and then move to Europe and the corpus of case-law that has developed on this side of the Atlantic.

6.1.9 World Digital Library (2009)

In a speech delivered to the US National Commission for UNESCO in 2005, James Billington, then US Librarian of Congress, proposed the establishment of a World Digital Library. UNESCO organized a meeting of experts in its headquarters in Paris in 2006.[440] In addition to the Library of Congress and UNESCO, the initiative started with the involvement of five partner institutions, including the National Library of Russia and the Russian State Library, as well as the involvement of the International Federation of Library Associations and Institutions, and institutions from other countries, mainly national libraries around the world.[441] The aim was the digitization of cultural content across the globe, encompassing developing countries with limited digitization capacity, as well as the development of multilingual access to the digitized content that is public and freely accessible. In 2009 the Library of Congress launched the 'World Digital Library.' The library is hosted by the Library of Congress but provides links to its partner institutions' websites directing to content. Therefore, each institution is responsible for the provision of its collections. Compared to the projects examined so far, this project is intended to be more international: represented are 193 countries and 134 languages. Its progress, however, appears to be rather slow if we compare it to the previous efforts. Content is comprised of prints, photographs, newspapers, books, manuscripts, maps, journals, motion pictures and sound recordings and includes in total 13,963 items and 853,216 files.[442]

Compared to the previous US-based initiatives, what is different in the World Digital Library is its institutional dimension: a public sector initiative that is also directly linked to the

[439] See *infra* 4.1.4.
[440] Information on the history of the initiative are online available at https://www.wdl.org/en/background/.
[441] See the list of partner institutions at https://www.wdl.org/en/partners/.
[442] Information about the project's progress found at https://www.wdl.org/en/statistics/ [last visited on May 2nd, 2017].

international arena. Compared to other digital library projects it claims to shift the focus from quantity to quality, emphasizing consistent metadata, multilingualism and the technical development in a collaborative international network. This project, however, seems significantly less developed and perhaps more dependent on funding and willingness of the initiating parties.

6.1.10 Digital Public Library of America (2010)

Last but not least, the latest and very active nationwide initiative to build a digital library for the United States started in 2010 with discussions in academic circles. The Digital Public Library of America (DPLA) began in Cambridge, Massachusetts, with a meeting of interested academics at Harvard University. It developed as a collaboration of scholars from various academic fields. The main base was Harvard's Berkman Center at Cambridge but there was also, for example, a project funded in Berkeley Law that concentrate on the copyright issues. Funders include the Sloan foundation, the Andrew W. Mellon Foundation, the Gates Foundation, the Knight Foundation as well as the Institute of Museum and Library Services and the National Endowment for the Humanities.

The writings of one of its initiators and now member of the board of directors, Robert Darnton, tell the story of the DPLA. Darnton, who at the time was director of the Harvard University Library system, started a public discussion on the possibility of creating a national digital library. We can see the elevation of the discussion and the actual strategy and realization of the set plans by reading his pieces published in the New York Review of Books. On 4 October 2010 Darnton published a piece entitled *A library without walls*. [443] There he revealed "an off-the-record conference at Harvard on October 1st to discuss the possibility of creating a National Digital Library." Since then he wrote four more pieces on the same issue and many more on related topics, such as the future of books and the implications of the Google Books project. On 28 October 2010, he published his talk at the opening of the conference at Harvard: *Can We Create a National Digital Library?*[444] On 24 November 2011, his piece *Jefferson's Taper: A*

[443] Darnton, R., A library without walls, *The New York Review of Books*, October 4, 2010, online available at http://www.nybooks.com/daily/2010/10/04/library-without-walls/.
[444] Darnton, R., Can we create a National Digital Library?, *The New York Review of Books*, October 28, 2010, online available at http://www.nybooks.com/articles/2010/10/28/can-we-create-national-digital-library/.

National Digital Library[445] connected the vision of a national digital library to the sentiments of Thomas Jefferson. Then came his enthusiastic piece on 25 April 2013: *The National Digital Public Library is launched!*[446] The fourth piece appeared in May 2014 and discussed how *A World Digital Library is Coming True!*[447] The discussion had already remarkably shifted from that of a national digital library to a world digital library. But here is what the national digital library is according to Darnton: "It would be the digital equivalent of the Library of Congress, but instead of being confined to Capitol Hill, it would exist everywhere, bringing millions of books and other digitized material within clicking distance of public libraries, high schools, junior colleges, universities, retirement communities, and any person with access to the Internet."[448]

Darnton's pieces mark a time when the creation of a national digital library was first floated as an idea and discussed. DPLA started out slowly but gathered flesh and bones in a collaborative manner directed by elite academic figures who were able to bring together "a group of librarians, foundation heads, and computer scientists." [449] This group of academics was willing to push the agenda, outreach to libraries across the US and talk directly to policy-makers in Washington. They engaged in a highly sophisticated form of lobbying. What is more they launched an initiative, independent of the public or the private sector, which included the involvement of the very academic institutions but in their capacity as a non-profit organization.

---

[445] Darnton, R., Jefferson's Taper: A National Digital Library, *The New York Review of Books*, November 24, 2011, online available at http://www.nybooks.com/articles/2011/11/24/jeffersons-taper-national-digital-library/.
[446] Darnton, R., The National Digital Public Library Is Launched!, *The New York Review of Books*, April 25, 2013, online available at http://www.nybooks.com/articles/2013/04/25/national-digital-public-library-launched/.
[447] Darnton, R., A World Digital Library is Coming True!, !, *The New York Review of Books*, May 22, 2014, online available at http://www.nybooks.com/articles/2014/05/22/world-digital-library-coming-true/.
[448] Darnton, Can we create a National Digital Library?, supra note 444.
[449] Darnton, Jefferson's Taper: A National Digital Library, supra note 445.
See also the History of the DPLA as stated in its webpage, at http://dp.la/info/about/history/:

> The DPLA planning process began in October 2010 at a meeting in Cambridge, MA. During this meeting, 40 leaders from libraries, foundations, academia, and technology projects agreed to work together to create "an open, distributed network of comprehensive online resources that would draw on the nation's living heritage from libraries, universities, archives, and museums in order to educate, inform, and empower everyone in current and future generations." In December 2010, the Berkman Center for Internet & Society at Harvard University, generously supported by the Alfred P. Sloan Foundation, convened leading experts in libraries, technology, law, and education to begin work on this ambitious project. A two-year process of intense grassroots community organization, beginning in October 2011 and hosted at the Berkman Center, brought together hundreds of public and research librarians, innovators, digital humanists, and other volunteers—organized into six workstreams and led by a distinguished Steering Committee—helped to scope, design, and construct the DPLA.

The DPLA is now based in Boston at the premises of the public library. It is still developing, with a significant growth rate, and works on the basis of a strategic plan towards its ambitious goal: "[to] bring […] together the riches of America's libraries, archives, museums, and cultural heritage sites, and making them freely available to students, teachers, researchers, and the general public." [450] Darnton explains that the rationale behind the project goes back to the founding of the United States when Thomas Jefferson preached that "knowledge is the common property of mankind." Since we now have the technical means "to make Jefferson's dream come true", he has consistently urged for the policy will to follow suit.[451] Indeed, it appears that the DPLA already collaborates with most of the initiatives examined so far, notably the Internet Archive and HathiTrust, and plans a more systematic collaboration with the Library of Congress.[452]

6.2 THE EUROPEAN PROJECTS AND LITIGATION

6.2.1 Gallica rejects Google and accelerates the European Digital Library project (2005)

While Google had already become the dominant digital library initiative in the US, Europe started with its own initiatives, which were linked to the public sector. First of all, national libraries engaged in the digitization of their collections that was mostly funded by public foundations, which were followed by public sector programs. Gallica, the digital portal of Bibliothèque Nationale de France (BNF), the national library of France, is one of the best examples that precede many other national and pan-European initiatives. It is online since 1997. The history of the project is quite interesting because it also directly links to the history of Europeana, the pan-European initiative as well as Google Books. The initial idea dating back to 1988 was to create a digital collection, replicating the actual resources, to be accessible through

[450] DPLA Strategic Plan 2015 through 2017 available at http://dp.la/info/wp-content/uploads/2015/01/DPLA-StrategicPlan_2015-2017-Jan7.pdf.
[451] Darnton, A library without walls, supra note 443.
[452] November 29, 2016 announcement of the Library of Congress: Library of Congress, Digital Public Library of America To Form New Collaboration, online available at https://www.loc.gov/item/prn-16-207/.
In his latest related piece in August 2015, Darnton wrote about the possibilities that are open to the Library of Congress, which he thinks thus far "has failed to manage its own information technology to say nothing of developing a national network of electronic resources." See Darnton, R. Great New Possibilities for the Library of Congress, *The New York Review of Books*, August 13, 2015, online available at http://www.nybooks.com/articles/2015/08/13/great-new-possibilities-library-congress/.

portals in library premises.[453] This initial project grew much bigger, which transformed into the creation of a digital library to be accessible on the web "to everyone, everywhere."[454] After it was approached by Google to digitize the library's collection which he declined, Jean-Noël Jeanneney, the director of BNF, wrote a rather impassionate article published in *Le Monde* in 2005. Jeanneney claimed that Google's dominance in the digitization business could result in a triumph of the Anglo-Saxon culture and the 'defeat' of Europe.[455]

France became the main advocate for a European digital library and started the discussions that led to Europeana. Meanwhile, it accelerated its digitization processes and also partnered with publishers and e-retailers in order to offer access to copyrighted works. Today Gallica offers more than 140,000 digitized documents and is one of the best examples of a successful digitization project of a national library.

6.2.2 Europe's New Renaissance plans and the launching of Europeana (2008)

The European Commission showed significant interest in digitization and the prospects of European culture and education. The recommendation to create a 'European Digital Library' is documented in a letter from 28 April 2005 addressed to José Manuel Durão Barroso, the then President of the European Commission. The letter was written by the premiers of France, Poland, Germany, Italy, Spain and Hungary.[456] Europeana was launched in November 2008 in the presence of Barosso and other European Commission officials and the Member States

[453] See the historical details on the webpage of the BNF: http://gallica.bnf.fr/html/und/a-propos

> L'objectif est de proposer aux lecteurs, sur des postes de lecture assistée par ordinateur, un ensemble de documents libres de droits et sous droits constituant la « bibliothèque virtuelle de l'honnête homme » (une volumétrie de 100 000 titres et 300 000 images était prévue pour l'ouverture au public du site François-Mitterrand).

[454] Id.

[455] Jean-Noël Jeanneney, Quand Google défie l'Europe, *Le Monde*, January 22, 2005, online available at http://www.lemonde.fr/archives/article/2005/01/22/quand-google-defie-l-europe-par-jean-noel-jeanneney_395266_1819218.html?xtmc=jeanneney&xtcr=38#.

[456] The letter, signed by Jacques Chirac, Aleksander Kwaśniewski, Gerhard Schröder, Silvio Berlusconi, José Luis Rodríguez Zapatero and Ferenc Gyurcsány, is available at the Digital Single Market webpage of the European Commission at: https://ec.europa.eu/digital-single-market/news/timeline-digitisation-and-online-accessibility-cultural-heritage.

representatives.[457] Europeana was meant to serve as a portal to provide access to Europe's cultural heritage and to ultimately achieve the goal of a European Digital Library.[458]

In April 2010 Neelie Kroes, then Vice President responsible for the EU Digital Agenda, and Androulla Vassiliou, then Commissioner in charge for Education and Culture, set up a scientific committee called the Comité des Sages to reflect on how Europe should actually plan the digitization of European cultural heritage. The committee comprised of three members: Maurice Lévy, (chairman and chief executive officer of the French advertising and communications company Publicis), Elisabeth Niggemann (director-general of the German national library and chair of the European digital library foundation) and Jacques De Decker (Belgian writer and journalist).[459]

The committee's task was to examine how exactly to put the collections of Europe's museums, galleries and libraries online. The committee released an ambitious report in January 2011 entitled "The New Renaissance".[460] In its mission the group referred to Jean Monnet who once said that if Europe were to be reconstructed, he would begin with culture rather than the economy. They made proposals for digitization and online access of European culture and the enhancement and promotion of Europeana in order for the website to become "the reference point for European cultural content online".[461] They specified that EU Member States should receive public funding for digitization on the condition of free accessibility of their digitalized materials. Furthermore, by 2016 the public domain masterpieces should be incorporated into Europeana. The rapporteurs expected that it would be necessary for the Member States to increase their funding for digitization considerably. To this end, they also supported collaborations between private and public sectors. Furthermore, the report suggested solutions for making works covered by copyright available online and called for the prompt adoption of EU rules for orphan works. Among its other recommendations, it is worth noting that the committee also suggested that cultural institutions receive public funding if they digitize public domain material and make it available for access and reuse. They further recommended that

[457] Id.
[458] See the Europeana portal at http://www.europeana.eu/portal/.
[459] EC Press Release, April 21, 2010. *Boosting cultural heritage online: the European Commission sets up a Reflection Group on digitization*, IP/10/456, online available at http://europa.eu/rapid/press-release_IP-10-456_en.htm.
[460] *The New Renaissance* - Report of the 'Comité des Sages' reflection group on bringing Europe's cultural heritage online (January 10, 2011), supra note 18.
[461] Ibid.

cultural institutions should comply with Directive 2003/98/EC on the reuse of public sector information.[462] Indeed, in 2013 the Directive was amended by Directive 2013/37/EU, which extended the scope to include libraries, archives and museums.[463]

While centralized efforts to coordinate digitization continue, one case of small-scale digitization carried out in a university library in Germany reached the courts. This is the first case relevant to digitization and the potential of digital libraries in Europe while across the Atlantic the cases against Google and HathiTrust were ongoing. As part of the timeline we will look at this case followed by two more cases; one case concerns e-lending rights for public libraries and the second case deals with the issue of copying (thus also digitization) of out of commerce works.

6.2.3 The decision against the Technical University of Darmstadt (2014)

The first relevant European court case was initiated in Germany, where the Technical University of Darmstadt (hereinafter the 'TU Darmstadt') digitized a textbook which it held in its library collection and made available on its electronic reading posts on the premises of the library.[464] The book called 'Einführung in die Neuere Geschichte' [Introduction to Modern History] was authored by Winfried Schulze and published by the German publishing house, Eugen Ulmer. Notably, the library had refused the offer of the publishing house to purchase an e-book collection including the book in question. At specific reading points, students were allowed to consult works that belong to the collection. The digital copies available were reflecting the number of copies that the library actually possessed – thus access was allowed to the same number of copies one actually owned by the library to be consulted at any one time. The library users, however, could print all or part of these works or store it on a USB stick.[465]

In view of the above, Eugen Ulmer, the publisher, brought an action against the university and the case was brought before the competent German regional court. The regional court, on the one hand, rejected Ulmer's application to prohibit the digitization and digital use of the work. On the other hand, it granted the request to prohibit users of the TU Darmstadt library

[462] Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the Re-use of Public Sector Information, OJ L 345, 31.12.2003, p.90
[463] Directive 2003/98/EC was amended in 2013 by Directive 2013/37/EU, OJ L 175, 27.6.2013, p. 1–8.
[464] *Technische Universität Darmstadt v. Eugen Ulmer KG*, (Case C-117/13), supra note 8. References hereinafter are to either the Judgment (C-117/13 Judgment) or the Advocate General opinion (C-117/13 Opinion).
[465] C-117/13 Judgment, paras 12-13.

from printing the digital version available or store it in a USB stick.[466] This decision was handed down in 2011. On appeal, the competent German federal court was asked to interpret a specific provision of the Information Society Directive and determine whether the work's availability to purchase or licensing terms played any role in the case, which could prohibit TU Darmstadt from digitizing and offering the work to its patrons.[467]

In 2013 the federal court of Germany sent three questions to the CJEU for a preliminary ruling. First, it asked the European court to interpret Article 5(3)(n) of the Directive, which provides for an exception to the reproduction right and the communication and making available right of copyright-holders that is relevant to the case.[468] Second, it asked whether that specific exception allows Member States to confer on libraries "the right to digitize the works contained in their collections, if that is necessary in order to make those works available on terminals." Third, it asked whether the rights guaranteed under Article 5(3)(n) of the Directive "go as far as to enable users of the terminals to print out on paper or store on a USB stick the works made available there."[469] The German court sent the questions at the end of May 2013. Subsequently the Advocate General issued an opinion in June 2014 and the final decision was handed down in September that the same year. The case was finally resolved also at the national level.[470] The

---

[466] For a detailed account of the facts as the case proceeded from the German system to the CJEU see Linklater, supra note 32. See also Loewenheim, U., 2015. Boundaries of use of electronic terminals. *Journal of Intellectual Property Law & Practice*, *10*(5), pp.384-387.

[467] Directive 2001/29/EC of the Parliament and of the Council of 22 May 2001 on the Harmonisation of Certain Aspects of Copyright and Related Rights in the Information Society, supra note 225. See also the analysis under 4.1.5.

[468] Article 5 paragraph 3 of the Information Society Directive lists a number of exceptions or limitations to both the reproduction and the communication and making available rights that Member States "may provide." Article 5(3)(n) allows an exception for:

> use by communication or making available, for the purpose of research or private study, to individual members of the public by dedicated terminals on the premises of establishments referred to in paragraph 2(c) of works and other subject-matter not subject to purchase or licensing terms which are contained in their collections […]

The establishments referred to in paragraph 2(c) are publically accessible libraries, educational establishments or museums and archives.

[469] C-117/13 Judgment, para. 22. Paragraphs 16-21 review the German federal court's view on the questions. The German court applied the national legal rules and made an initial balancing on the basis of the three-steps test (article 5(5) of the Directive) which has also been incorporated in the national law.

[470] The German courts have decided finally and in favor of the University library. See Pauler, N., and Ettig, D., Germany: Technische Universität Darmstadt wins landmark case with Hogan Lovells – Digital use of copyright protected works in public libraries, *Hogan Lovells Lime Green IP News*, April 20, 2015, blogpost available at http://www.limegreenipnews.com/2015/04/germany-technische-universitat-darmstadt-wins-landmark-case-with-hogan-lovells-digital-use-of-copyright-protected-works-in-public-libraries/

Advocate General's opinion and the final decision of the European court came to the same conclusion. Let us take a look at both.

Advocate General Niilo Jääskinen began his opinion stating that "the main proceedings are a test case."[471] He signaled the significance of the case which goes beyond the specific parties and extends to libraries and publishers throughout Europe, noting how the TU Darmstadt was supported by the German Library Association and the European Bureau of Library, Information and Documentation Associations, while Ulmer was supported by the German Publishers and Booksellers Association.[472] At the same time he was fully aware of the mass digitization litigation in the United States and referred to the legal controversies around the Google Book Search project in the footnote of his opinion.[473] His opinion was delivered in June 2014 when the case against Google had already passed the district court level and was pending at the appellate level. Advocate General Jääskinen, however, unlike the court after him, wanted to distinguish this particular case from cases concerning mass digitization, which according to him, might not lead to the same outcome. In paragraph 38 of his opinion, he specifies that the "specific acts of reproduction" of Article 5(2)(c) and 5(3)(n) "[do not] permit the general digitization of a *collection* [emphasis in the opinion], and the object 'specific acts of reproduction' is thus limited to *individual* 'works and other subject matter.'" [474] Such interpretation was not repeated in the court's decision. Indeed, the Advocate General's position on this point might come across as overly narrow. In any case, mass digitization, or the digitization of an entire collection would raise slightly different legal issues in the European context given the Directive on orphan works, the jurisdictional issues regarding copyright rules in the EU, and the existence of collecting societies. All these factors make the European case different to the US case. Apart from his statement distinguishing mass digitization, Advocate General Jääskinen concluded on the basis of the facts before him that (i) the mere offer of a license does not preclude libraries from digitizing, (ii) that Article 5(3)(n) of the Directive allows for the digitization of library works and making available at specified terminals, and (iii) that this right does not extend to printing or USB storage of the digitized works.[475] His conclusions and the court's conclusions, which we will examine shortly, align. Nevertheless, the court's decision

[471] C-117/13 AG Opinion, para. 4.
[472] Ibid.
[473] Id. footnote #3 of the Opinion.
[474] Id. para. 38.
[475] Id. para. 59.

in this case can be read to be slightly more generous towards libraries. In this particular case, the Advocate General was marginally more cautious in his interpretation of the copyright exceptions laid down in the Directive.[476]

The final decision was handed down by the Fourth Chamber of the CJEU. The court considered the three referred questions and made the following assessment. Firstly, it considered the fact that the publisher had offered a licensing agreement to the university and concluded that even in this case the library was not bound by the offer since they had not actually entered into any agreement. Thus the Article 5(3)(n) exception was clearly not excluded. In fact, the court noted that this exception "aims to promote the public interest in promoting research and private study, through the dissemination of knowledge which constitutes, moreover, the core mission of publically accessible libraries."[477] The mere licensing offer by Ulmer is not enough to negate the exception which has this public interest purpose.[478] Such statements that emphasize the public interest purpose of the exceptions make this decision a stronger statement than the Advocate General's opinion. The wording of the decision indicates that the judges of the Fourth Chamber wanted to make a case for the libraries and highlight the importance and rationale of the copyright exceptions when they favor these institutions.

Even more critical is that the judgment then went on to make specific reference to a *right to digitize*. The court, first of all, found that Member States are allowed to grant publicly accessible libraries covered by the said provision the right to digitize works contained in their collections "if such act of reproduction is necessary for the purpose of making those works available to users, by means of dedicated terminals, within those establishments."[479] In addition, in paragraph 43 of the judgment the court explained in greater detail that the right of libraries to communicate, at specific terminals, the works in their collections "would risk being rendered largely meaningless" if they did not have an ancillary right to digitize.[480] This is already a step further towards digitization, which not solely limited to preservation purposes.[481] Indeed, the even terminology employed - a *right* to digitize - is impressive and clearly strengthens the

[476] See also Linklater, supra note 32, at pp. 819-821 on the slight differences between the AG Opinion and the final Judgment.
[477] C-117/13 Judgment, para. 27.
[478] Id. para. 28.
[479] Id. para. 49.
[480] Id. para. 43.
[481] Note that Article 5(2)(c) of the Information Society Directive allows specific acts of reproduction which are not for direct economic or commercial advantage.

library's position,[482] which under the Directive belongs to a non-obligatory list of exceptions to that Member States exercise. Moreover, the judgment specified that digitization of the work as such is not coupled with an obligation to private compensation. The making available of the digitized work, however, as provided in German law also gives rise to a fair compensation to be given to a collecting society.[483] This is to be assessed at the Member State level, as each national system has its own way to regulate the collecting societies.

Moving to the third question on the printing and storing on USB sticks of the digital versions available at specified terminals, similar to the AG, the court answered in the negative:

> Article 5(3)(n) of Directive 2001/29 must be interpreted to mean that it does not extend to acts such as the printing out of works on paper or their storage on a USB stick carried out by users from dedicated terminals installed by publically accessible libraries covered by that provision. However, such acts may, if appropriate, be authorised under national legislation transposing the exceptions or limitations provided for in Article 5(2)(a) or (b) of that directive provided that, in each individual case, the conditions laid down by those provisions are met.[484]

Both provisions cited in this part of the judgment provide for exceptions with respect to reproductions on paper or other mediums but in that instance the right-holder receives compensation. The answer to this third question was to be expected. If USB storage or printing were to be permitted, then this would go beyond the justification of the listed exceptions.

The overall conclusions that we can draw from this case are necessarily limited to the facts of the case brought before the court and the nature of the CJEU decisions, stemming from references for preliminary ruling. The court interprets the rules of the Directive as they apply to a national case where national rules are also applicable; here, it was German law. National copyright rules are not harmonized at the EU level. Bearing these reservations in mind, the importance of the case should not be discounted. It demonstrates the position of the European court *vis-à-vis* national courts. It also contributes to continuing efforts of the EU legislator to

[482] The case is seen as clearly yielding in favor of the libraries' side in the debate on the balance of rights of copyright-holders and the public interest in access to library collections in the digital age. See Linklater and also Carstens, supra note 32.
[483] C-117/13 Judgment, para. 9.
[484] Id. para. 58.

reform copyright rules at the EU level and take further steps towards harmonization. In conclusion and as far as small-scale digitization is concerned, this judgment contributes to a clear line for EU Member States: they can provide exceptions in favor of libraries and give them an explicit right to digitize works from their collections. On the next level, when it concerns the making available of these digital reproductions, of course additional constraints must be in place.

6.2.4 The e-lending case initiated by the Dutch association of public libraries (2015)

In 2015, the European court received another preliminary ruling request for a case initiated in the Netherlands by the Dutch association of public libraries. This is an association, in which all public libraries in the Netherlands are members: Vereniging Openbare Bibliotheken (hereinafter the 'VOB').[485] The case had started some years earlier, after the regulatory body that sets the price for lending rights in the Netherlands, the Stichting Onderhandelingen Leenvergoedingen, decided that the lending of electronic works is not within the scope of the Dutch law on public lending. This decision, made in 2010, was immediately challenged by the VOB which started legal procedures, first at the domestic level, which ultimately reached the CJEU. VOB initiated the case before The Hague district court against the collecting society that is responsible for collecting remuneration for authors as a result of public lending. VOB asked the domestic court for a declaratory judgment that the lending of electronic books falls within the scope of the lending right. The court was also asked to declare that the lending of electronic books by public libraries, with a fair remuneration to authors, does not constitute copyright infringement. [486] At the same time, VOB was challenging a piece of draft legislation on libraries that provided for the creation of a national digital library and stipulated e-lending fell outside the scope of the legislation already in place.[487] Basically, the VOB consistently sought recognition that e-lending is possible under the existing legal framework that allows for public lending.

This case is very relevant for our timeline since it raised a central issue for the future of digital libraries. One of the main services of libraries is the lending of books. [488] While the legal status of e-lending has been unclear, the Dutch association took the initiative to ask for a

[485] *Vereniging Openbare Bibliotheken v. Stichting Leenrecht*, (Case C-174/15), supra note 9. References hereinafter are to either the Judgment (C-174/15 Judgment) or the Advocate General opinion (C-174/15 Opinion).
[486] C-174/15 AG Opinion, para. 15.
[487] C-174/14 Judgment, para. 23.
[488] We have seen how the lending of electronic works is not as straightforward under the current copyright framework. See *infra* under chapter four: 4.1.6 *Library e-lending*.

declaratory judgment on the issue. VOB had come up with a proposal: an e-lending system that tries to replicate what libraries have always been doing with physical books: it described its lending system as a 'one copy one user' system where the electronic lending mirrors physical lending - thus the copy downloaded by a patron is not accessible to other patrons- and is also lent for a specific, limited period of time. VOB also expressly wished to limit the e-lending possibility to novels to only include collections of short stories, biographies, travel guides, children's books and youth literature.[489]

The questions that the Dutch court referred to the CJEU are framed as ones of interpretation of Directive 2006/115, which regulates renting and lending rights.[490] The Dutch court primarily asked whether the provisions of this Directive are to be interpreted to mean that 'lending' includes the making available of works under copyright (the question is limited to the types of works mentioned above) by "placing a digital copy (reproduction A) on the server of the establishment and enabling a user to reproduce that copy by downloading it onto his own computer (reproduction B)"[491] in a manner that ensures both that the copy will stop being available after a set period and that other users will not be able to simultaneously download that copy while on loan. This was the first question posed to the court. It essentially portrays an e-lending mechanism whereby the lending of physical books is actually replicated as much as possible in the context of e-books. Should the question be answered in the affirmative, it is then inquired on further details with respect to how can Member States regulate e-lending, allowing for example libraries to circulate digital copies after their ownership has been transferred.[492] Essentially, what happens with exhaustion as far as digital versions of a lawfully owned book are concerned? This question is that of remote downloading, thus outside of the premises of a publically accessible establishment: can remote downloading be allowed and for unlimited periods?[493] The hearing of the case took place in March 2016. In June the Advocate General issued his opinion, which was followed by the court's judgment in November.

[489] C-174/15 AG Opinion, para.16.
[490] Directive 2006/115/EC of the European Parliament and of the Council of 12 December 2006 on rental and lending right and on certain rights related to copyright in the field of intellectual property (codified version), supra note 237.
[491] C-174/15 AG Opinion, para. 18.
[492] Ibid.
[493] Ibid.

The Advocate General, Maciej Szpunar, delivered an opinion noteworthy in many respects. His opinion reads as a strong statement in favor of libraries and the transformation they ought to undergo in the digital era. His introduction to the opinion set the tone for what follows:

> Libraries are one of civilization's most ancient institutions, predating by several centuries the invention of paper and the emergence of books as we know them today. In the 15th century, they successfully adapted to, and benefited from, the invention of printing and it was to the libraries that the law of copyright, which emerged in the 18th century, had to adjust. Today we are witnessing a new, digital revolution, and one may wonder whether libraries will be able to survive this new shift in circumstances. Without wishing to overstate its importance, the present case undeniably offers the Court a real opportunity to help libraries not only to survive, but also to flourish.[494]

Subsequently, the opinion referred to the digital revolution, together with the increasing importance and use of digital works – such as e-books and other formats of electronic works. The opinion stressed the need to allow libraries to adapt to the new era, to which if they prove unable "they risk marginalization and may no longer be able to fulfill the task of cultural dissemination which they have performed for thousands of years."[495] Advocate General Szpunar understood that e-lending is key for digital libraries to adapt to the digital age, and at the very beginning of his opinion he cites the relevant European scholarship on the debate.[496]

---

[494] C-174/15, AG Opinion, para. 1.
Eleonora Rosati notes the remarkable language and welcomes the departure of a "*fairly dry style*" of Advocate General Opinions. See Rosati, E., AG Szpunar says that time-limited e-lending is allowed under EU law and interpretation of copyright norms must evolve with technology, *IP Kat Blog*, June 17, 2016, available at http://ipkitten.blogspot.gr/2016/06/ag-szpunar-says-that-time-limited-e.html.

[495] Id. AG Opinion para. 3.

[496] See footnote 3 of the Opinion where he cites the following:

> Davies, P., 'Access v. contract: competing freedoms in the context of copyright limitations and exceptions for libraries' in European Intellectual Property Review, 2013/7, p. 402; Dreier, T., 'Musées, bibliothèques et archives: de la nécessité d'élargir les exceptions au droit d'auteur' in Propriétés intellectuelles, 2012/43, p. 185; Dusollier, S., 'A manifesto for an e-lending limitation in copyright' in Journal of Intellectual Property, Information Technology and E-Commerce Law, 2014/5(3); Matulionyte, R., 'E-lending and a public lending right: is it really a time for an update?' in European Intellectual Property Review, 2016/38(3), p. 132; Siewicz, K., 'Propozycja nowelizacji prawa autorskiego w zakresie działalności bibliotek' in Zeszyty naukowe Uniwersytetu Jagiellońskiego. Prace z prawa własności intelektualnej, 2013/122, p. 54; Zollinger, A., 'Les

In his assessment the Advocate General supported a dynamic interpretation of Directive 2006/115. His interpretation requires an adaptation of the Directive's rules to the digital age. He first went back to the Directive's legislative history and finds "undeniable that, at the time, the EU legislature did not contemplate the inclusion of the lending of electronic books within the concept of lending" [497] Still, he thought that the provisions of the Directive cannot be interpreted today in a way that excludes the lending of e-books and gives three reasons: (i) that rules need to be interpreted in line with technological changes (ii) that copyright should serve the interests of authors and it is indicative that in this case two collective organizations of authors and creators of visual artworks are intervening in support of the library association and the recognition of an e-lending right for the libraries and (iii) that libraries as an institution deserve a favorable interpretation.[498] The issues that the Advocate General raised with his line of argumentation are many. Only the issue of equivalence, for example, between physical books and e-books is a complex one that he addresses in paragraph 31 in a rather casual manner and deserves a much more elaborate analysis. He mentioned that "the lending of electronic books is the modern equivalent of the lending of printed books" and since the difference between e-books and print books is one of format and not content, it is not necessary to justify different treatment by the law.[499] As the AG expressly admits, his interpretation of some of the copyright rules is more 'dynamic' or 'evolving' than literal.[500] It is in any event remarkable, however, how his opinion is written in a manner undoubtedly in favor of the libraries as institutions. The underlying argument is basically that the library is an ancient institution that precedes copyright and serves valuable purposes in the society: cultural preservation and dissemination of knowledge. He is adamant that this institution cannot flourish if it is subject to "the laws of the market."[501] This is how he explains the derogations and privileges in favor of libraries, which primarily encompass the lending rights as codified in the relevant Directive.

Moving to the conclusions, Szpunar answered the first of the preliminary questions in the affirmative. Essentially Directive 2006/115 is to be read to mean that the lending right it provides

---

bibliothèques numériques, ou comment concilier droit à la culture et droit d'auteur' in La semaine juridique. Entreprise et affaires, 2007/25, p. 18.

[497] C-174/15, AG Opinion, para. 25.
[498] Id. paras 26-40.
[499] Id. para 30.
[500] Id. para. 28.
[501] Id. paras 36 and 38.

for "includes the making available to the public of electronic books by libraries for a limited period of time." [502] He continued with the outstanding questions which following the positive answer to the first question could be answered: the question of transfer of ownership requirement for the digital copies to be e-lent (essentially whether exhaustion becomes applicable), and the fact that e-lending must apply only to e-books obtained from lawful sources. The Advocate General clarified that under the Directive's rule the lending right is independent of the exhaustion of the right to distribution.[503] The requirement that the Advocate General set was "that electronic books that are lent should first have been made available to the public by the rightholder or with his consent".[504] As to be expected, he also confirmed that the copy should be acquired by a lawful source.[505]

Lastly, with regard to the distinction that the referring court made by grouping together novels according to collections of short stories, biographies, travel-guides, children's books and youth literature, the Advocate General had already ruled out differentiated treatment of any category of books. He thus considered the questions to include all kinds of books and e-books, which he had explained at the beginning of his opinion.[506]

After this very strong opinion in favor of libraries, and their needs, the final decision of the court was enthusiastically anticipated. In November 2016 the court agreed with the Advocate General; the lending of digital copies is covered by the concept of lending according to Directive 2006/115.[507] The court upheld the conditions, the 'one copy one user' system. The court also agreed with the AG that the lending is independent to the first sale or other form of transferring property within the EU. The court also concurred that in each case the digital copy must be legal.[508] The analysis in the judgment is more elaborate in terms of legal provisions, including international law (the WIPO copyright treaty) and parts of the recitals of the two Directives, 2001/29 and 2006/115.[509] Citing the recital clauses the court solidified the teleological argument that the Advocate General made with regard to the legal rules and their evolution in response to

[502] C-174/15, AG Opinion, para 80.
[503] Id. para. 83. See also Rosati, supra note 494 noting that the AG avoided to reopen the exhaustion question as was discussed extensively in UsedSoft (Case C-128/11) by simply mentioning that there is no reason for digital and hard copies to be treated differently.
[504] Id. para. 85.
[505] Id. para. 86.
[506] Id. para. 22.
[507] C-174/15 Judgment: ruling.
[508] Id.
[509] Id. Paras 3-10.

technological progress. Particularly recital 4 of Directive 2006/115 stating that "Copyright and related rights protection must adapt to new economic developments such as new forms of exploitation" was seen as strengthening such conclusions.[510] A new and quite strong legal argument that the court added was the differentiation of lending from renting. While rental rights explicitly do not apply to digital copies in accordance to the WIPO Copyright Treaty, nothing precludes lending rights to be extended to digital copies.[511] What differentiates lending is the non-commercial character of the making available of a work or a copy.[512] Last but not least, the court stressed the need to ensure the effectiveness of the lending right exception even though exceptions are generally interpreted strictly.[513] With that, the extension of public lending rights to digital copies seemed only logical for the exception to remain relevant in the digital age.

Overall the court confirmed the Advocate General conclusions, while admittedly doing so in a slightly weaker voice. Although not following the same pompous language that was used by the Attorney General, the court indeed implied that it recognizes the special status of libraries in at least two parts of the decision: (i) in paragraph 51 the decision refers to "the importance of the public lending of digital books" and (ii) in paragraph 62, it introduces the notion of "cultural promotion" and the "general interest underlying the public lending exception".

The final outcome of this case is desirable from a policy perspective. The arguments in favor of the library privileges adapting to the digital era are many and both the Advocate General and the court seem to have spent time considering and articulating them. Nevertheless, the line of argumentation and interpretation of existing rules especially as first demonstrated in the Advocate General opinion were not entirely without objection. The French government, for instance, is cited in the opinion for objecting to the broad interpretation of copyright exceptions.[514] The Advocate General in his opinion expressly disagreed and argued that in this case interpreting the copyright exception narrowly so as not to cover e-books, which already covers physical books, would undermine the effectiveness and purpose of the exception.[515]

Before the final decision of the case, academics had also expressed doubts as to whether the claim of an e-lending right is possible under the existing rules. The report prepared by

[510] Id. para. 45.
[511] As the court explains in para. 38 "the EU legislature sought to define the concepts of 'rental' and 'lending' separately. Thus the subject matter of 'rental' is not necessarily identical to that of 'lending'."
[512] Id. para. 29.
[513] Id. para. 50.
[514] See para. 46 of the Opinion.
[515] Id. para. 47.

scholars at the Institute for Information Law at the University of Amsterdam (IViR) to the Dutch government, dated back in 2012, had concluded that e-lending is not covered under the Dutch Copyright Act or the European copyright framework.[516] The report relied on a literal and historical interpretation of the examined rules. In her 2015 article 'A manifesto for an e-lending limitation in copyright', Séverine Dusollier examined the European framework for public lending and the possibility of a copyright exception favoring libraries by replacing the exclusive right of the right-holders to a remuneration right. After examining the law as it is, she argued for a "controlled extension" of the public lending right to also cover e-lending.[517] Dusollier saw that, even though the Directive does not explicitly exclude e-lending, legislative reform would be needed if the target is to support libraries as institutions and give them a better alternative to the e-licensing agreements that publishers currently offer them.[518] According to Dusellier, as the right to make available online is covered by the Information Society Directive and not explicitly by the Directive on rental and lending, e-lending is not covered at all at this point. This is in contrast to the interpretation of Szpunar and that of the court.[519]

In spite of the outcome of the examined case, the issue of e-lending remains a big puzzle for copyright given the general lack of clarity over digital exhaustion, as we also saw in chapter four.[520] Lastly, in view of the confusing framework and exactly because the outcome of the Dutch case has been rather unexpected, its importance in the context of our digitization stories is big. With the legal recognition of the possibility to construct e-lending mechanisms under the current EU copyright framework, libraries have are in a safer position after this judgment and have a stronger incentive to proceed with the digitization of their collections.

[516] van der Noll et al., supra note 48.
[517] Dusollier, supra note 32.
[518] Ibid. On the established licensing practises see also the study of O'Brien, Gasser and Palfrey, supra note 32.
[519] Nevertheless, even if they follow different reasoning, the judges involved in this case and Professor Dusollier ultimately aim for the same outcome: normatively speaking, that e-lending *should* be allowed. The underlying rationale expressed by Dussollier as a normative target is also similar the Advocate General's articulated rationale: the market falls short of advancing the public good, a role that libraries have traditionally played and are capable of carrying on within a good regulatory framework. What differs in the Advocate General's approach, however, is the effort to reach such a result without any reform of the applicable rules. In this sense the Advocate General arguably drafted his own manifesto in favor of e-lending, which in was then followed by the court.
[520] See also among many Rosati, E. 2015. Online copyright exhaustion in a post-All posters world. *Journal of Intellectual Property Law & Practice*, p.122.

6.2.5 The decision on French legislation about books out of print (2016)

The final case to be included in this timeline is *Marc Soulier & Sara Doke v. Premier ministre & Ministre de la Culture et de la Communication*.[521] This case concerns legal developments in France, where new rules applicable to the category of out of print books would have direct impact on digitization projects of libraries and other cultural heritage institutions.

In 2012 the French legislator passed a new law for out of print books: *Loi n° 2012-287, du 1er mars 2012, relative à l'exploitation numérique des livres indisponibles du XXème siècle* (Law No 2012-287 of 1 March 2012 on the digital exploitation of out of print 20th century books). The new piece of legislation gave collecting societies the right to authorize reproduction and communication to the public in digital form for the category of out of print books. As defined by this law, the category is to include books "published in France before 1 January 2001 which is no longer commercially distributed by a publisher and is not currently published in print or in a digital format'."[522] A public database was to be created and made openly available "free of charge, through an online, public communication service" and be operated and updated under the responsibility and supervision of the National Library of France.[523] The particular provision allowing for exceptional authorization to reproduce goes as follows:

> When a book has been registered in the database referred to in Article L.134-2 for more than six months, the right to authorize its reproduction and performance in digital format shall be exercised by a collecting society governed by Title II of Book III of this Part and approved for that purpose by the Minister for culture.
>
> With the exception of the case provided for in the third subparagraph of Article L.134-5, the reproduction, on a non-exclusive basis and for a renewable period of five years. […][524]

[521] *Marc Soulier & Sara Doke v. Premier ministre & Ministre de la Culture et de la Communication* (Case C-301/15), supra note 10. Hereinafter references to the Judgment (C-301/15 Judgment) or the Advocate General Opinion (C-301/15 AG Opinion).
[522] See C-301/15 Judgment para 15, referring to Article L. 134-1 of the French Intellectual Property Code.
[523] Ibid, Article L. 134-2.
[524] Ibid, Article L. 134-3.

Subsequent provisions provided for the possibility for authors to oppose their right to exercise authorization through an approved collecting society. This right remains also after the expiry of the six-months period (L.134-3) but without any right of compensation. The law also specified the rights of publishers to similar opposition and provides for the specific terms of notification and for more elaborate timelines for the exploitation of the books for which there is no opposition from authors or publishers, as well as the distribution of future income collected.

The above legal provisions were implemented by decree 2013-182, against which two authors, Marc Soulier and Sara Doke, initiated legal proceedings.[525] They requested the annulment of the decree before the Conseil d' État which then sent a preliminary request to the CJEU on the interpretation of Information Society Directive. In particular, the issue at hand was whether the new law amending the French intellectual property code was in conformity with the Directive. Advocate General Melchior Wathelet first issued his opinion on 7 July 2016 and the Third Chamber of the court issued its decision on 16 November that the same year. Both the Advocate General and the Third Chamber held that the new legislation was not in conformity with the Directive. First the Advocate General rephrased the question referred to the court while invoking the applicable provisions of the Directive:

> […] the referring court asks […] whether national legislation that gives approved collecting societies the right to authorize, in return for remuneration, the reproduction and performance in digital format of 'out-of-print' works is compatible with Article 2(a) of Directive 2009/29, which establishes an exclusive right of reproduction for authors, and with Article 5 of that Directive, which allows Member States to provide for exceptions or limitations to that right.[526]

[525] While the case was still at the national level, the Conseil d' État found that the new law was consistent to the French constitution and in particular the right to property. Under the set framework, the interference with the property rights of authors was according to the French highest court not disproportional. The French court made a cost-benefit analysis on the one hand weighing the importance of making available 'written heritage which has become inaccessible for want of commercial distribution to the public' and on the other the rights of the authors. It also took into account the several stages and measures of publicity in place. Yet the French court referred to the European one for the question of compatibility with the European Directive, which brought us to the opinion of the Advocate General and the court decision.

[526] C-301/15 AG Opinion, para. 23.

The Advocate General explained that the exploitation of the digital version of a book involves both reproduction and making it available to the public thus also constitutes communication in the sense of Article 3(1) of the Information Society Directive. Therefore, the compatibility of the French provisions with both Articles 2(a) and 3(1) should be assessed.[527] For that he looked at the scope of the two articles and stated that they require prior and expressed consent from the right-holders for any reproduction or communication to the public ('high level of protection'). He also considered whether the possibility of opposition and withdrawal as well as the right to remuneration that the French law reserved for right-holders of out of print books are enough to fall within the scope of these provisions. Again, he answered negatively. Last but not least, he clarified how the fact that the work is not distributed commercially does not alter his conclusions.[528] Furthermore, the Advocate General concluded that the new French exception to the right of reproduction and communication is not authorized by the Directive. Indeed, Article 5 of the Directive provides for an exhaustive list of exceptions and limitations. On top of that, there is the three-steps test under Article 5(5). The Advocate General maintained that Member States should not extend the scope of the set exceptions at the expense of legal certainty.[529] He also considered the Orphan Works Directive and found that it to be another explicit legal exception limited to orphan works but cannot cover the out of print works as defined by the French law. He actually argued that the conditions under which orphan works can be exploited cannot be more restrictive than those applicable to out of print works.[530] By the same token, the mechanism proposed by the French law cannot be interpreted simply as a management arrangement.

At the end of his opinion Advocate General Wahthelet added a very interesting part where he replies to submissions from SOFIA (Société Française des Intérêts des Auteurs de l'écrit), the French, the German and the Polish governments. This part is quite relevant to our discussion on digitization. The intervening parties introduced a Memorandum of Understanding signed on 20 September 2001 *on key principles on the digitization and making available of out-of-commerce works*.[531] This document was the result of a stakeholder dialogue initiated by the Commission. The purpose of the dialogue and the document that resulted had been to support

[527] Ibid, paras 24-25.
[528] Ibid, para 44.
[529] para. 31.
[530] Para. 52.
[531] Memorandum of Understanding Key Principles on the Digitisation and Making Available of Out-of-Commerce Works, September 20, 2001, online available at http://ec.europa.eu/internal_market/copyright/out-of-commerce/index_en.htm.

large-scale digitization of out of commerce works. Arguably the French law under scrutiny in this case was doing exactly that. Yet, according to the Advocate General, the Memorandum is not legally binding and thus cannot limit the scope of Articles 2(a) and 3(1) of the Information Society Directive. Furthermore, even within the realm of the Memorandum, any agreements between right-holders and collecting societies must be voluntary. Thus under the existing law, Member States cannot allow opt-out arrangements that are not explicitly authorized by legislative exceptions at the European level.

Paul Keller commenting on the opinion said that it "effectively undermines the idea that extended collective licensing can serve as a solution for the copyright problems created by mass digitization."[532] This observation is fair. It is also surprising that both the Advocate General and later the court disregarded the Commission's promises to make the digitization of out of commerce works easier as documented in soft-law documents and elaborate preambles of its latest relevant Directives.[533] Placing the case in its historical context, however, we should keep in mind the announced reforms to the EU copyright laws which presumably the European judges did not want to predict. Making clear that the current state of the law, having a *numerous clausus* of exceptions, does not allow the organizing mass digitization projects through extended collective licensing, the role of the judges in this case can be interpreted as perhaps emphasizing the need for reform.[534]

The final decision of the Third Chamber did not add much to the Advocate General's opinion. Citing the Berne Convention and the international obligations under the Information Society Directive, the decision solidified the conclusion already made by the Advocate General and emphasized that national legislation must not bring back, even indirectly, any sort of formality for the exercise of the exclusive rights of right-holders.[535] Indeed, when the court discussed the scope of the respective articles of the Directive, it emphasized the high and broad level of protection demanded in view of Berne Convention as well as the WIPO Copyright

[532] Keller, P., Advocate General Wathelet: Extended Collective Licensing is NOT the answer for mass digitisation!, July 14, 2016 blogpost, available at http://www.communia-association.org/2016/07/14/advocate-general-wathelet-extended-collective-licensing-not-answer-mass-digitisation/.
[533] See the above-mentioned Memorandum of Understanding and also the preamble of the Orphan works Directive.
[534] Keller, supra note 532: "While the immediate effect of such a ruling might be negative in the short run (it might limit access to French out-of-print books that are currently available), it could have a positive long-term impact. In the wake of the AG opinion the very real possibility of such a ruling could sway the Commission (and Member States that are reluctant to update exceptions, such as France) toward deciding on an approach based on an update of the relevant exceptions. This possibility alone makes it worth taking a closer look at AG Wathelet's opinion."
[535] C-301/15 Judgment, paras 3-9.

treaty.[536] The decision is rather brief compared to the Advocate General's opinion but is better connected to a particular *aquis* of copyright decisions which reinforce the need of prior consent of authors when it comes to reproduction of their works.[537] The court concluded that "subject to the exceptions and limitations laid down exclusively in Article 5 of Directive 2001/29, any use of a work carried out by a third party without such prior consent must be regarded as infringing the copyright in that work."[538]

Interestingly, the conclusion of this case reminds us of the final dismissal of the Google Books Settlement. The European court seems to refrain from a step that would go 'too far' from the default copyright rule of right-holder consent. While these rules remain in place, let us conclude the timeline with an overview of the progress that Member States have made in the digitization process under the supervised of the Commission.

6.2.6 Progress until today: The EU Commission supervises the members' digitization progress

Over the past number of years, the European Commission and Member States have invested millions of Euros in supporting cultural heritage institutions (museums, archives and libraries) to digitize their collections and make them available on the Internet. Between 2005 and 2009, the Commission invested one 149 million Euros in the 'eContentPlus' program alone.[539] Since 2008 there has been significant progress both in the Member State engagement with digitization and in EU legislation. Having said that, the overall work is still in progress, not only in view of legal obstacles but also in view of other practical, mostly financial, obstacles. The Commission set up a 'Member States Expert Group on Digitisation and Digital Preservation' (MSEG) in 2007. The

[536] Ibid, paras 31 and 50.
[537] Ibid, para. 33. With respect to the consent to reproduction the decision makes reference to: *Infopaq International*, C-5/08 (paras 57 and 74), *Football Association Premier League and Others*, C-403/08 and 429/08 (para. 162). On consent for the communication to the public see: *SCF Consorzio Fonografici*, C-135/10 (para 75) and *Svensson and Others*, C-466/12 (para. 15).
[538] C-301/15, para 34.
For this prior consent to be valid, the court emphasized the need that authors are *actually and individually* informed. Ibid, paras 43-43: "It does not follow from the decision to refer that that legislation offers a mechanism ensuring authors are actually and individually informed. Therefore, it is not inconceivable that some of the authors concerned are not, in reality, even aware of the envisaged use of their works and, therefore, that they are not able to adopt a position, one way or the other, on it. In those circumstances, a mere lack of opposition on their part cannot be regarded as the expression of their implicit consent to that use. That is all the more true considering that such legislation is aimed at books which, while having been published and commercially distributed in the past, they are so no longer."
[539] *The Cost of Digitising Europe's Cultural Heritage*, supra note 320.

group's mandate has been extended until 2016, and it looks like it will be further extended until 2020.[540] The group has reported on the progress of the Member States, particularly on the implementation of 'Commission recommendation on the digitisation and online accessibility of cultural material and digital preservation'.[541] In June 2016 the group published its 2013-2015 report according to which:

> Overall, almost all Member States have achieved good progress with the digitization of cultural material, reporting continuity of plans that have been established in the past few years, or new developments […]. Different approaches in planning digitization were again reported, with schemes ranging from national strategies (10 MS) supported by national funding programmes or implemented through domain-specific digitization plans, to domain-specific initiatives (6 MS)led by Ministries or by national institutions, to regional schemes or even planning based on strategies of individual institutions.[542]

The financing of digitization is usually public, but also comes from a mix of public and private funds, with some Member States being more successful than others in securing public-private partnerships; the report mentions the UK, France and the Netherlands, as examples. Some libraries have partnered with major private partners, such as Google and Proquest. Other countries, such as Spain, and Finland, have also relied on national sponsoring from foundations.[543] An interesting finding that emerges from the Member States reports is that there seem to be obstacles in ensuring that public domain material remain in the public domain after their digitization.[544] Among the libraries or portals that have been listed as successful paradigms are the Koninklijke Bibliotheek's and University Library's databases in the Netherlands, also allowing for free reuse of digital books, the French BNF, the Polish cultural portal Polana, and

[540] See the meetings of the 18th meeting of the MSEG (22-23 November 2016) available at https://ec.europa.eu/digital-single-market/en/news/18th-meeting-member-states-expert-group-digitisation-and-digital-preservation-mseg.
[541] *Cultural heritage, Digitisation, online accessibility and digital preservation,* Report on the Implementation of Commission Recommendation, 2011/711/EU 2013-2015, online available at http://ec.europa.eu/information_society/newsroom/image/document/2016-43/2013-2015_progress_report_18528.pdf.
[542] Ibid, p. 6.
[543] Ibid, p. 10 et. seq.
[544] Ibid, p. 22.

the Swedish Digisam. Last but not least, some best practices from various Member States are promoted as good examples. These include the British Library Labs initiative funded by the Andrew W. Mellon Foundation, which promotes experimentation with the collections and data of the British library, or the Dutch Open Cultuur Data, launched by the Rijksmuseum Amsterdam to promote creative reuse of virtual content.[545] The digitization projects of Scandinavian countries have also been innovative and very successful.[546]

If one compares the findings of the latest report to the findings after the 2013 national reports, it seems that substantial progress has been made. On the basis of the national reports submitted in late 2013 and early 2014, the Commission concluded that on average 12% of library collections had been digitized while the percentage for certain media, such as films for example, was even smaller (less than 3%).[547] Even in the latest report, however, and despite the progress of the last two years, the expert group cautions that "the overall picture of cultural heritage digitization remains fragmented and patchy, widely dependent on cultural institutions' initiatives or funding, with a limited overview of digitization activities across sectors and borders."[548]

6.3 LESSONS TO BE LEARNED FROM THE TIMELINE

The timeline above started in Illinois. It moved to the north of the Pacific coast and then to the mainland of the US, including Michigan, Washington DC and other 'stops', and ultimately the east coast, Cambridge and Boston in Massachusetts. Crossing the Atlantic it included Paris, Brussels and three judicial decisions at the Luxemburg court initiated in Germany, the Netherlands and France. From this outline of the most important digitization initiatives in the US and Europe, including the relevant case-law, we can draw some initial conclusions. Most important is how all the different projects and approaches seem to share the same objective: the

[545] Id. See more details for the projects at their webpages: http://labs.bl.uk/ and http://www.opencultuurdata.nl/.
[546] Among the examples of best practices are projects of the Finish National Gallery, and the National Museums of World Culture in Sweden. Notably Norway engaged in digitization of libraries' entire collections, aiming to place online every work ever written in Norwegian. In view of the healthy collective rights management system in these countries, it has been easier for their public sectors to negotiate and plan the digitization and making available of entire collections compensating the right-holders.
[547] Data available at the Commission's webpage: https://ec.europa.eu/digital-single-market/en/news/2013-national-reports-digital-preservation.
[548] *Cultural heritage, Digitisation, online accessibility and digital preservation,* Report on the Implementation of Commission Recommendation, supra note 541, p. 73.

creation of a digital library. Looking in particular at the language used for the announcements of and self-descriptions of each of the initiatives, one can see different expressions of the same core vision. The language changes depending on the ambition: a national library, a regional library or a universal library.[549] Nonetheless, the term digital library is used loosely to refer to some kind of digital platform, online accessible, offering access to complete works and metadata, or sometimes only to metadata (as illustrated, for example, by Google Books when it comes to copyrighted content).

A first conclusion to draw regards the definition of digital libraries for legislative purposes: It might be worth considering a sharp and narrow definition of what constitutes a digital library if we are to apply special rules or copyright exceptions for their benefit. Among the various stakeholders, who have engaged with digitization projects, not all of them have the same motivations or purposes. For example, insofar as we are to retain the original copyright rationale, the encouragement of learning, copyright exceptions for digital libraries shall be applicable to institutions devoted to such purpose. From a normative perspective it is arguable that not all stakeholders engaged with digitization are to benefit from legal privileges for preservation or access purposes. The concluding chapter of the thesis will elaborate further on this point as it makes suggestions for normative directions.

Reviewing the existing initiatives together with the case-law that we examined, one can see a number of stakeholders which include non-for-profit and for profit private parties, academic libraries and public libraries. Comparing the situation of the United States and Europe, first of all we can detect a different balance in the driving forces. In the US, private initiatives seem to dominate thus far; historically, there were non-profit private efforts that paved the way. These were quickly followed by for-profit and mixed initiatives that involved different, mostly private, stakeholders. Lately, it seems that public sector has started to get more involved, mostly collaborating with various private stakeholders. By contrast, the continental European stories are either directly or indirectly linked to the public sector, at the national and the supranational (European Union) level. This is an important observation that also links to regulation. In the European context, one sees more elaborate regulatory intervention with new legislation passed or

[549] Among many news articles on the 'universal library' ambitions associated particularly with Google Books, see for example Toobin J., Google's Moon: The quest for the universal library, *The New Yorker*, 5 February 2007 [online], available at http://www.newyorker.com/magazine/2007/02/05/googles-moon-shot or Singer P., Whither the dream of the universal library?, *The Guardian*, 19 April 2011 [online], available at https://www.theguardian.com/commentisfree/2011/apr/19/moral-imperative-create-universal-library.

discussed mostly at the EU level and aimed at addressing legal issues around digitization. We have already mentioned the Orphan Works Directive and the now proposed amendments towards the modernization of EU copyright rules which includes concrete rules applicable to out of commerce works. These are examples of how legislative solutions have been faster in Europe than in the US. This is partly due to the immediate involvement of the public sector and also partly due to the differences between the two legal cultures more generally.

Another point in the comparison is the scope of the projects and the challenge posed to existing copyright rules and the resulting litigation. As we saw, litigation arose in the courtrooms of both jurisdictions. The scope of the copyright challenges, however, was arguably different, which further differentiated the scope and the breadth of litigation issues. The first two non-profit endeavors, Project Gutenberg and the Internet Archive, understood the copyright status of many works as a clear burden. Hence, they started working with public domain material while advocating against the lengthy copyright term and a general framework that make digitization of copyrighted works disproportionately difficult. The Google Books and HathiTrust projects, however, had a broad scope to include the scanning of copyrighted works. It is worth acknowledging the 'progressive' ethos that most of the American endeavors share. Even though they dealt with it in a radically different manner, Project Gutenberg and the Internet Archive (together with the Open Content Alliance), on the one hand, and Google Books and together HathiTrust, on the other, all shared the same disapproval of copyright obstacles that they encountered. The subsequent projects (principally the DPLA and less even the World Digital library project) cannot be categorized in the same way. Having a more academic or institutional flavor, these two better fit the European paradigm in many ways. Precisely due to the connection to the public sector, European efforts began in a more 'conservative' manner, which proceeded hand-in-hand with legislative measures. For instance, when the European Commission announced that it will promote digitization projects in the EU Member States, it first passed the new Directive dealing with the issue of orphan works. Similar to attempts to address continuing problems with the European digitization endeavors, the EU concurrently proposed new exceptions to copyright laws that would also allow cultural heritage institutions to digitize out of commerce works. In view of the involvement of the public sector in the European digitization efforts, the challenge of the legislative framework has been expectedly much more modest.

Going beyond the above first conclusions, the following section makes a further observation with regard to the difficult position that libraries find themselves in when competing with the various other stakeholders interested in digitization and the creation of different forms of digital libraries. Ultimately, the section argues how this period of digitization projects together with the various cases before the American and European courts perhaps reflect new period of 'copyright wars.'

6.4 DIGITIZATION AS THE CAUSE FOR ANOTHER ROUND OF 'COPYRIGHT WARS'?

'Copyright wars' is a phrase frequently used to describe the conflicts between various stakeholders during the short - three centuries old - history of copyright laws. Both historians and lawyers have employed war metaphors to describe the evolution of copyright doctrine. Indicatively William Patry had used the phrase copyright wars to refer to the challenges that new technology and new business models cause to copyright as a system.[550] In his work Patry teaches us that metaphors matter and they have mattered a lot during the relatively brief history of copyright. Adding a historical perspective, Peter Baldwin wrote a book recently dedicated to the *Three centuries of transatlantic battle*, where he describes the conflicts in the rationale and principles that the continental European and the American copyright systems initially represented.[551] The war metaphor has also been employed to portray of the industry battles and legislative action against piracy. Interestingly, all the above references to war imply much more than legal controversies: the controversies, expressed in legal terms, are ultimately deeply cultural. [552]

The phenomenon of digitization, as it actually unfolded with the examined initiatives and litigation that they sparked, could be interpreted as another copyright 'battleground'. Wars are

[550] Patry, W., 2009. *Moral panics and the copyright wars*. Oxford University Press.
[551] Baldwin, P., 2014. *The copyright wars: three centuries of transatlantic battle*. Princeton University Press. According to Baldwin, the US, which used to be a major copyright pirate in the eighteenth and nineteenth century, eventually yielded to the European protectionist copyright rationale and became the world's intellectual property policeman.
[552] See, for example, Yu, P.K., 2003. The escalating copyright wars. *Hofstra L. Rev.*, *32*, p.907, concentrating on the battles against online piracy. Also: Wu T., Leggo my Ego – Google Print and the Other Cultural War, *Slate*, October 17, 2005 [online], available at http://www.slate.com/articles/news_and_politics/jurisprudence/2005/10/leggo_my_ego.html. Wu's small article is utmost relevant to this discussion as it refers to Google print. For Wu, who wrote this back in 2005, Google started a cultural war on authorship and how it is to be understood and defined in the future.

usually over scarce goods like gold is. Scarcity makes goods extremely valuable, and therefore it is relatively straightforward that competing stakeholders over scarce goods arise. While digitized books need not bring about such issues of scarcity, copyright laws seem to have created the circumstances for the rise of competitive interests and claims. Thus, even if at first glance the competition that started over book digitization, which by definition creates non-scarce goods, might seem paradoxical, the fact that the digitization wars were basically expressed as copyright litigation explains this paradox. As we have seen, copyright has been controlling the making of copies of books and until today retaining their scarcity. Copyright laws insist on retaining books' scarcity even when technology allows for mass and cheap copying. Here is where the legal battles were grounded. Parties benefiting by the current copyright framework have been willing to fight these wars having long invested interests in the rules as they were drafted in the pre-digital era and as evolved at times when technological evolution was either underestimated or perhaps misinterpreted.

Our claim here is that the relevant case-law, as we saw it in the timeline that preceded, forms part of the so-called 'copyright wars'. In view of the value that the various parties saw in digitization, the phenomenon has provided for another occasion for copyright territory to be claimed. As copyright affords exclusive rights, the parties that manage to win the copyright claims in this context gain the exclusive control over the valuable resource winning a 'copyright war'. This might be beneficial for parties with clear private interests, as for example the authors, the publishers or the technology companies involved. The same benefit - of claiming copyright over a digitized corpus – might, however, be less relevant for libraries. Libraries care not about the exclusive control over books but about preserving them, systematizing them in collections and, more importantly, about providing access. Thus, for libraries, the fact that this 'war' is being fought in copyright terms is not necessarily as beneficial as it is for the other players involved. We will continue this discussion in the concluding chapter when we evaluate the overall position of libraries amidst the phenomenon of digitization. The argument is made on the basis of a mismatch between the potentials that libraries have and the leverage that the legal framework allows them in practice. Insisting only on copyright claims is actually undermining the role of libraries as institutions dedicated to a public service that goes beyond market considerations that copyright serves.

Before arriving to these conclusions the next chapter will add to the discussion on the adverse consequences that digitization has on books and libraries. In particular, chapter seven examines, firstly, the transformation of the book as a medium, secondly, a failure relating to book unavailability or inaccessibility and the phenomenon of shadow libraries online, and thirdly, - connected to the second point - the persistent inequality in access to books.

CHAPTER SEVEN - ADVERSE CONSEQUENCES OF DIGITIZATION ON BOOKS AND ON LIBRARIES

This chapter examines the consequences that the phenomenon of digitization seemingly has both vis-à-vis books and the libraries. The first section is devoted to the transformation that books experience in the digital era, as a result of digitization and the loss of books' physicality. This affects the common understanding of ownership (and property rights) on books held by individuals or institutions. It also discusses the issue of increased commercialization as a result of the digitization. Is then the book becoming less of a merit good?

With the proliferation of media content, especially available online, the reader's expectation for book availability is arguably higher. There are studies, however, indicating that there are real gaps in book availability, accessibility and reuse as a result of copyright. Together with these studies this chapter will discuss the phenomenon of file sharing of literary content and of the so-called 'shadow libraries.' Thus, the next section of this chapter is devoted to the failure of book unavailability or inaccessibility online and the phenomenon of shadow libraries. Ultimately the unavailability of books, or availability in shadowed places operating on the margins of the law, brings back the initial thoughts about unequal access to books and the 'needs of the poor.'

For some digitization could also signal the end for books and libraries. However, such a 'disaster scenario' is mostly dismissed as exaggerated. Thus, the chapter concludes with a tribute to the relevance of the physical book and the physical library, whose role is far from obsolete. On a more pragmatic account digitization brought a number of changes affecting the book sector, including both market players (publishers and book stores) and non-market players such as libraries. The book remains as the common resource for these players. Thus, the effects that digitization has on books and how they are experienced are relevant for all. Time will indicate how important the transformation will be for our future experiences in accessing content, in reading and in creating further.

7.1 BOOK 2.0: THE TRANSFORMATION OF THE BOOK IN THE DIGITAL ERA

With digitization, the book undergoes a radical transition in its form. The effects on the original paper books can be minimal or substantial; for example, disbanding pages usually makes the digitization process faster and easier but at the same time damages the original.[553] This is not, however, the kind of transformation we discuss here. We want to explore the conceptual changes that digitization results in as regards the medium - how is it experienced by readers but also by machines. Indeed, the reading experience is transforming. Maurizio Borghi and Stavroula Karapapa talk about the 'empowered reading environment' that digitization creates.[554] Reading books in a digital form can become interactive where words and information are linked to other relevant digital resources and programs. This interactivity is heightened when the experience takes place online. Borghi and Karapapa further observe that this transition from the physical to the digital version also converts works into data. Books can be used less for reading purposes and more to feed machines with data. As the uses of books change, we might wonder whether, despite the same content, paper books and electronic books are not substitutes for one another. Arguably from the perspective of the human reader, they are substitutes since the content remains the same. The reading environment might offer additional features but the overall transformation can be considered moderate. From the perspective of the machine, reading the two different forms of the good are not perfect substitutes.[555]

Whether paper books and electronic books are substitutes can be seen as a two-pronged question of principle and fact. One part of this question is whether books and electronic books are different versions of the same good. A second, less conceptual part of the same question is the actual perception on the market: do readers understand two versions of the same content, one printed on paper and one in electronic form, as the same good? A third and connected question is normative: Should electronic books be treated differently by law and if yes, then why? Even if

[553] See Ghiasi M., Books in the digital revolution, January 1, 2011 blogpost, available at http://ghiasi.org/2011/01/books-at-the-digital-revolution/.
[554] Borghi, M. and Karapapa, S., supra note 35, at p.10.
[555] Resolving the substitutability issue is important for a number of legal reasons. These questions are, for example, relevant to the application of competition law on markets of paper books and e-books. They are also relevant with regard to pricing and taxation. If the paper and electronic versions are perfect substitutes, then under competition law, they are part of the same market and should also be priced and taxed in the same way. If they are not perfect substitutes, then all the above conclusions are subject to greater debate. The two versions of the same content could actually compete in the market. To borrow the wording of Landes and Posner, the demand curve of a given book will be negatively sloped when there are good [but not perfect] substitutes. Thus also pricing is affected.
Posner and Landes, supra note 68, p. 39.

books and e-books are perceived as different goods, principled reasoning might, for example, dictate that they should be still treated the same, or *vice versa*.

We will make two disclaimers before we move forward. First of all, this dissertation does not attempt to give definite answers to any of the above questions on whether books and e-books are in principle different and whether they are perceived as such in the market. With the digital libraries as the central focus, the third question is ultimately more important for our purposes: the normative arguments for the legal treatment of books and e-books especially as library recourses. The question of substitutability posed here complements the overall analysis of books and their transformation in the digital era and will ultimately constitute the framework in order to answer the questions specific to libraries. The second disclaimer is that the use of the term electronic books, or e-books, is meant to include both digitized paper versions of books and the so-called born-digital versions. Most books today are born in digital form and are (or not) printed subsequently. The analysis refers to both mechanical digitization of preexisting paper versions and the reproduction of electronic copies. In both cases it is essentially the act of copying the same content that matters.[556]

7.1.1 It's the end of property as we know it

A critical conceptual difference that digitization brings about is the loss of physicality. This is important from a legal perspective, as property has traditionally been attached to some form of physicality. The concept of ownership over a book has always been connected to its physicality. It is not the content that we own when we buy a new book, which belongs to the author at least as a matter of attribution. It is the particular physical copy that we own.

In 2001 Thomas Merrill and Henry Smith wrote that *property has fallen out of fashion*.[557] They discussed how the law and economics school contributed to an overall alienation from the *in rem* character of property. They claim that Coase's impact on the modern understanding of property rights is such that they are now commonly understood as nothing more than a bundle of

[556] There is no doubt, for example, that even if paper and digital copies of books are substitutes, rare antique books and their copies are not.

[557] Merrill, T.W. and Smith, H.E., 2001. What happened to property in law and economics?. *The Yale Law Journal*, *111*(2), pp.357-398.

rights.[558] Nonetheless, these authors insist on a qualitative difference between *in rem* rights and *in personam* rights.[559] To understand property as an institution and therefore a system of property rights, rather than individual cases or single disputes, they claim that the *in rem* aspect of property needs to be restored.[560] The *in personam* model that has prevailed is helpful to understand two-party models. If the *in rem* aspect of property rights is neglected, then the bigger picture is missing. In a recent book Aaron Perzanowski and Jason Schultz look at how the traditional *in rem* dimension of property has in reality been lost on more occasions than we spend time to think about. This is the case because of the transformation of products from physical, 'analogue', to digital but also in view of the technology creep extends to physical objects. They think of two situations: (1) licenses replacing ownership, for example in the cases of e-book purchases from Amazon or music playlists on i-Tunes and (2) Digital Rights Management protection embedded in physical objects, such as smart TVs or smart coffee machines that are programmed to work with only the same brand of coffee beans.[561] Perzanowski and Schultz focus on how the digital distribution of goods together with a combination of factors, predominantly technological progress (like cheap and easy digital copying, remote storage enabled by cloud computing and so on) contribute to a fluidity in the concept of ownership. They carry out a comprehensive analysis to apprehend "how did our rights in media goods become so unstable and insecure."[562] One way to view it is that the digital marketplace offers better price discrimination and today's consumers have a range of choices along a spectrum of full ownership of goods to "conditional, impermanent access to digital and physical goods."[563]

---

[558] Ibid, p.365: "The conception of property as an infinitely variable collection of rights, powers and duties has today become a kind of orthodoxy."
The authors attribute this to a particular influence that Coase's 1960 article, *The Problem of Social Cost,* had on scholarship that followed. They think that to a certain extent Coase over-interpreted and also that his particular examples of two-party disputes made it easier for him to "overlook the differences between in personam and in rem rights" (p.371). This observation is based more on the nuisance disputes of the Problem of Social Cost. His comparatively less cited article *Federal Communications Commission* (1966) deals with a topic that does not involve both property and tort law as the previous article but a more straightforward allocation of property rights on a resource (broadcast spectrum). Merrill and Smith claim, however, that Coase's conception of property rights as a mere bundle of rights is also detectible in this article.
Among the many scholars the authors cite see: Penner, J.E., 1995. The "Bundle of Rights" Picture of Property, 43 UCLA L. Rev. 711, at pp. 733-38. Last, the same concern to Merrill and Smith has been expressed by Heller, M.A., 1999. The boundaries of private property. *The Yale Law Journal*, *108*(6), pp.1163-1223.
[559] Ibid. (Merrill, T.W. and Smith) p. 364. See also, their reference to older works such as Hohfeld, W.N., 1913. Some fundamental legal conceptions as applied in judicial reasoning. *The Yale Law Journal*, *23*(1), pp.16-59.
[560] Ibid, p.379.
[561] Perzanowski A. and Schultz J., 2016, supra note 42.
[562] Ibid, p.3.
[563] Ibid, p.6.

Perzanowski and Schultz stress that when making non-ownership choices, we need to be aware of the list of substantive rights we lose, or risk losing. This essentially goes back to the discussion of Merrill and Smith on the consequences of weakening the concept of property. All four authors agree that the lack of clarity concerning rights demand closer legal attention. Part of the work that Perzanowski and Schultz made is a comprehensive analysis of an observation made by legal scholars some time ago: the shift from ownership to licensing in the digital era.[564] Most important for this dissertation is their focus not only on the impacts on the individual consumer but also on educational and cultural institutions. In their chapter on digital libraries, Perzanowski and Schultz look at the licensing of e-book collections and how e-lending has developed in the US context where the principle of digital exhaustion is not applicable to digital goods.[565] They also look at the rising privacy concerns and the dangers posed to readers' anonymity, as well as concerns with regard to archiving and preservation. In other words, the dangers that libraries face in view of the transition to digital collections are not negligible: (i) the legal status and permanency of their e-book collections is unclear so long as access to these collections is based on licensing schemes in stead of full ownership of the electronic collections; (ii) without ownership, and the corresponding rights, on their electronic collections, libraries' archiving and preservation services are also discounted; and (iii) once e-book distribution to patrons takes place via intermediary platforms, readers' anonymity is not guaranteed. The difficulties with applying the legal concept of exhaustion to digital goods is, therefore, another example of how the transformation of the book in the digital era has real consequences that ultimately affect both what readers can or cannot enjoy and what libraries can and cannot do with their electronic collections.

7.1.2 Does digitization increase commercialization of books?

It is unclear whether digitization has increased the commercialization of books. In a compelling article, M. Bjørn von Rimscha and Sarah Putzig investigate the tendencies towards

[564] See for example the same observation in O'Brien, Gasser, and Palfrey, supra note 32; Müller, supra note 32; Chiarizio, An American tragedy: e-books, licenses, and the end of public lending libraries, supra note 32; Afori, The Battle Over Public E-Libraries–Taking Stock and Moving Ahead, supra note 32.
[565] See also chapter four *infra* (Section 4.1.6).

commercialization which can perhaps be attributed to digitization.[566] Commercialization is defined as the increasing orientation towards market demand and profit, rather than normative values associated with a self-referred cultural market.[567] As products of a cultural market books have been traditionally treated as merit goods and enjoyed benefits, such as reduced tax rates and in some jurisdictions fixed pricing.

With greater detail, von Rimscha and Putzig claim that there is some evidence that due to digitization, the commercial character of the book market is stronger. Before digitization, books were explicitly treated, even within the market by publishers and booksellers, as cultural goods. The self-conception of the book industry as a cultural industry is not without criticism. The industry has used the 'cultural card' when lobbying for beneficial legal treatment, for example by demanding taxation or antitrust exemptions.[568] Von Rimscha and Putzig also point to an irony and a change in the industry with the entry of digitization:

> Even if the shelves of an ordinary bookshop are dominated by rather trivial mystery thrillers or mommy porn, the self-conception of many in the industry is that of a cultural industry. Even more than their counterparts in the film industry, book managers consider their product as art not to be judged solely by economic measures. However, some of the new players in the market are of a new pedigree. With a background in IT or retail logistics they might not care all that much about the cultural aspects of the trade. Still, digitization brings more power to these actors, thus shifting the priorities in the industry.[569]

The same authors identify two central ways in which digitization transforms the book market. The first way is the direct communication with the customers, whose preferences are not only tracked but also targeted by digital marketing. This might in fact disrupt diversity in the market. Secondly, digitization reduces barriers to entry into the publishing market, with self-

[566] von Rimscha, M.B. and Putzig, S., 2013. From Book Culture to Amazon Consumerism: Does the Digitization of the Book Industry Lead to Commercialization?. *Publishing research quarterly*, *29*(4), pp. 318-335.
[567] Ibid, p. 321.
[568] Luey, B., 1994. Priorities for publishing research: a multidisciplinary agenda. *Publishing Research Quarterly*, *10*(3), pp. 31-35.
[569] von Rimscha, M.B. and Putzig, S., supra note 566, p.319.

publishing being a viable option for authors. Thus, even if market demand of certain titles leads commercial publishers to concentrate their investment in certain genres, authors of less popular or less commercial genres and themes can find alternative outlets in the digital sphere. There are more effects on the overall book market structure which we can briefly summarize as follows: (i) retailers are bypassed due to online sales which allow immediate paring of supply and demand; (ii) pricing depends on output (since the cost of extra digital copies is low, the higher the demand and the output, the lower the prices); (iii) while competition at the product level decreases, platform competition increases[570] and (iv) the rise of converged and networked market structures disrupts the authority of 'legacy actors' who are faced with new competitors from the technology sector (see the bar as set by Amazon, but also i-tunes).

The empirical study that von Rimscha and Putzig conducted focused exactly on the question of whether the priorities of the industry have changed because of digitization and the redefinition of the publisher's and the bookseller's connection to customers. Their study concentrated on the German and Swiss markets.[571] The German market in particular is quite relevant, which is also the second biggest book market in the world. Based on their study, they concluded that online retailers in 2012 were clearly gaining ground lost from 'brick and mortal retailers'. Further conclusions included that digitization has clear effects on the distribution chain and that there are sufficient signs that digitization leads to commercialization.[572]

Such studies reinforce the idea that the dematerialization that digitization has real effects not only on the market but also on the notion of the book and how it might be shaped in the future. This is much more of a cultural shift than a market disruption. As we saw, the loss of physicality has broader implications related to our relationship with digital goods and our ability to possess them in the same manner we possess physical property. Markets have been based on property structures and ultimately digitization not only disrupts the distribution chains in a changing book market but also challenges the very foundations of that market which is grounded in property.

A notable effect to libraries when it comes to increased commercialization is perhaps the rise of competition as result of private actors. Libraries are understandably less inclined to

[570] Ibid. Also, including the lock-in effects due to platform switch costs.
[571] Ibid. Indeed, they disclaimed that their findings cannot simply apply to Anglo-Saxon markets without taking into consideration the differences in the book market cultures.
[572] Ibid, at p.331: digitization allows a stronger market orientation and thus can lead to commercialization.

compete with the book market actors as they serve different purposes. The competition, however, might be inevitable when the services that book retailers wish to offer start to resemble those of the library. To use an obvious example, Amazon's lending library on Kindle is basically offering a service traditionally associated with libraries. Let us note that there is a distinction between the transformation of the book as a medium and the ways it is distributed. Amazon's model for example has challenged both small and big bookstores, perhaps challenging big retail bookstores more than smaller ones. Online shopping and home delivery of physical books is arguably a bigger threat to bookstores than the threat of digitization. Libraries, on the other hand, cannot easily be grouped together as their functions and services differ to those on the book market. Access to electronic books, however, blurs the lines between bookstores, online retailers and libraries when the latter offer e-books to their patrons. In the case of electronic books, all these actors may be directly competing. Unrestricted e-lending programs from libraries could cannibalize on sales of e-books in the market. Nevertheless, bookstores and libraries also need to remain distinct in our analysis in view of their different objectives. For libraries books and e-books are not a commercial product, which they are for bookstores. Therefore, the effects of digitization on each of these two actors should be studied separately.

For the purposes of this dissertation we shall remain focused on libraries and books in libraries. It is there that books, in physical and digital form, retain their characteristics as merit goods even if the nature of the book changes in market structures. The balance between the public purpose that books serve when accessible in libraries and their role as commercial products is of course still relevant, especially when we look at how many private stakeholders were involved in the digitization initiatives. In the concluding chapter we will inquire into possible motivations particularly for the involvement of the technology companies. We will see how the digitized book which for libraries remains a merit good, is commercially relevant for multiple uses in the digital environment.

7.2 A SIGNIFICANT FAILURE: BOOKS UNAVAILABLE AND THE PHENOMENON OF SHADOW LIBRARIES

While disaster scenarios of the death of the book or the libraries are less realistic the challenges remain. Adding to the ones discussed already there is a notable failure with respect to book availability in the digital age that strikes us as contradictory: while digitization allows for easier copying and distribution, books are still considered by many to be unavailable or inaccessible online. This links to the phenomenon of piracy and the rise of the so-called shadow libraries. Book piracy is not a new digital phenomenon. It is a complex historical phenomenon that can be traced back to the early days of book trade in the 15th century. Pirate publishers would put into circulation books from different jurisdictions, books for which the right to publish was afforded only to printers in the big cities, censored books and so forth. According to Balázs Bodó, an economist that has researched the phenomenon of book piracy, pirates at that time usually responded to market inefficiencies created by publishing monopolies in the first book markets around Europe.[573] In his paper, *Short History of Book Piracy*, he sees substantial similarities between book piracy in the early ages of publishing and online piracy today. He argues that the power relations among firms and cultural markets, including the position of dominant guilds whose interests the first copyright laws protected, are all factors associated with book distribution, piracy and copyright enforcement. This has been the case from the early ages until today.

When thinking about the history of copyright law and the history of book piracy, it becomes apparent that the rationale and behavior of earlier book pirates is similar to today's book piracy. Book pirates tend to emerge when legal supply and demand is either non-existent and inefficient or otherwise problematic. Recent economic studies have pointed to the unavailability of book content online, particularly when it comes to out of print books. The expectation that books are available is a recent phenomenon that is related to the promises of the digital age. Online content is massive but at the same time Internet users also expect it to be unlimited. Google and Amazon readers get easily frustrated if a book not born in the digital era is unavailable or not found online in some format, often the e-book, which is usual for born digital and digitized books, or available to be shipped quickly by online retailers, like Amazon.

Below, we will take a look at Paul Heald's recent studies aiming to prove that copyright protection has resulted in what he describes as a *shocking unavailability* of books. Subsequently

[573] Balázs, Coda: A short history of book piracy. Media piracy in emerging economies, supra note 45, p.399.

we will look at another study that shows how books under copyright are less likely to be reused online.

7.2.1 Gaps in the availability of books

Paul Heald has conducted research on the issue of the demand for out of print books and their "shocking unavailability."[574] In his paper, *How Copyright Keeps Works Disappeared,* Heald demonstrates how the availability of books since the 1980's (sometimes orphan works or out of print works) is far lower than in the 1880's (which at the time of his study were obviously in the public domain).[575] Heald endeavored to test the assumption that works that are not protected by copyright are underexploited. His study suggests, otherwise. On the basis of his findings he argued, first, that the copyright status and legal availability of works are highly correlated and, second, that copyright law seems to "deter distribution and diminish access."[576] Heald's claim is bold, which goes as follows:

> Copyright correlates significantly with the disappearance of works rather than with their availability. Shortly after works are created and propertized, they tend to disappear from public view only to reappear in significantly increased numbers when they fall into the public domain and lose their owners. For example, more than twice as many new books originally published in the 1890's are for sale by Amazon than books from the 1950's, despite the fact that many fewer books were published in the 1890's.[577]

Heald took a random sample of new book titles available on Amazon.com for which he determined copyright status by locating them in the Library of Congress records. From a large sample of almost 7,000 titles, he could verify the copyright status of 2,266 non-duplicate titles (as not all books are registered and also many books in his random sample were foreign, non-English titles). From the final sample of 2,266 titles he found that 72% were published prior to 1923 and had entered the public domain but 28% were still under copyright. He also estimated

[574] Heald, The Demand for Out-of-Print Works and Their (Un)Availability in Alternative Markets, supra note 46.
[575] Heald, How copyright keeps works disappeared, supra note 46.
[576] Ibid, p. 829.
[577] Ibid, p. 830.

the Amazon titles by percent per decade and his findings led to the following conclusions: instead of a smooth decline in availability of older works, he found that the curve from the decade 2000-2010 compared to the decade 1800-1810 "declines sharply and quickly, and then rebounds significantly for books currently in the public domain initially published before 1923."[578] In view of the steep decline of copyrighted books over time and before they enter the public domain, he assumed that publishers are not selling their books on Amazon more than a few years after their initial publication. Combining his data with additional factors, such as the estimated number of books published per decade, publishing business models, tax laws and copyright case-law, he concludes that copyright law is a determining factor for the disappearance of books, which reappear when they fall into the public domain:

> In a world without copyright, one would expect a fairly smoothly downward sloping curve from the decade 2000-2010 to the decade of 1800-1810 based on the assumption that works generally become less popular as they age (and therefore are less desirable to market). If age were the only factor, one would expect to see fewer titles available from each successively older decade. Instead, the curve declines sharply and quickly, and then rebounds significantly for books currently in the public domain initially published before 1923. Since age should be a factor that depresses availability, the most plausible conclusion from the data is that the expiration of copyright makes older works reappear. A corollary hypothesis is also supported by the data: Copyright helps make books disappear.[579]

Another conclusion, which is related to the first, was that no serious alternative marketplace for out of print books has developed so far. To verify the lack of alternative markets for out of print works, Heald looked at whether titles unavailable on Amazon are available in eBook format. He found that while works from the years 1913-22 are popular and successfully published in eBook form, copyrighted bestsellers in eBook format are far fewer in the 1930-2009 period (with a rise after 2010). He also looked at the market of used books, namely Google

[578] See figures 1 and 2 on pp. 839 and 841.
[579] Ibid. p. 841.

Books and data gathered from the Chicago library system, which roughly verify the same story: availability of eBooks after 1923 and before the millennium seems to have dropped significantly. Interestingly Heald drew a comparative analysis to the music industry and did not find the same results. He concluded that "the lack of availability of books from the post-1923 portion of the 20th century is startling."[580]

In a more recent study, Abhishek Nagaraj used evidence from a baseball magazine digitized by Google to demonstrate how copyright rules impact the *reuse* of content online. The study focused on reuse in Wikipedia.[581] Nagaraj explains the reason he chose the *Baseball Digest* is because Google digitized all existing issues of the digest, from 1940 to 2008. The key with this collection is that because of a failure to renew copyright at the time when renewal was still a requirement for protection under US law, the issues published before 1964 are already part of the public domain. Thus, Nagaraj was able to collect data and compare the reuse of the copyrighted issues in contrast to the ones in the public domain, tracking citations in Wikipedia. His finding was that information from the copyrighted issues of the magazine were far less likely to be used compared to information in the public domain: "after digitization, citations of out-of-copyright publication-years increased by about 135% compared to citations to in-copyright publication-years, even after controlling for publication-year and calendar-year fixed effects."[582] In addition, he shows how Wikipedia pages affected by copyright appeared to have significantly lower traffic compared to those that cited the public domain material. The author differentiates however between the impact on images and on text. In view of the fair use exceptions, the impact on the reuse of text is less compared to that for images. This study differs to Heald since it focuses on the impact of copyright rules even when content becomes available (in this case all the digitized materials were available online). The results are complementary, however, since both scholars point to a gap created by copyright rules.

The online piracy of books and other literary content is arguably linked to the phenomena that Heald and Nagaraj studied. Ultimately, we can see that if copyright makes content unavailable or it becomes a disincentive to reuse, the online community tends to disregard it. This shows that to a certain extent copyright rules have resulted in a failure in the digital age.

[580] Ibid. p. 861.
[581] Nagaraj A., 2016. Does Copyright Affect Reuse? Evidence from the Google Books Digitization Project. May 8, 2016, available at: https://ssrn.com/abstract=2810761.
[582] Ibid.

Signs or results of this failure are phenomena of book piracy and, furthermore, of shadow libraries operating online.

7.2.2 Book piracy and the phenomenon of shadow libraries

Let us take a closer look at the phenomenon of file sharing and piracy for literary works. First, there have been a number of loud and celebrated cases of online activism with regard to access to academic scholarship and the Open Access movement.[583] Perhaps most known is the fatal story of the federal prosecution against Aaron Swartz, the young activist of freedom of information online and Open Access.[584] In 2011, Swartz was charged with federal charges for computer fraud and illegal access and downloading content from JSTOR through MIT's network. It is unclear whether Swartz intended to create a digital library with the content he allegedly downloaded and make it available online. While facing serious criminal charges, Aaron Swartz committed suicide in January 2013 at the age of 26, leading to a public outcry against the severe enforcement of the Computer Fraud and Abuse Act in the United States and prompted rationalization of both copyright laws and their enforcement.[585]

---

[583] See among many: Swartz A. 2008. Guerilla Open Access Manifesto, available at https://archive.org/stream/GuerillaOpenAccessManifesto/Goamjuly2008_djvu.txt.

[584] Ibid, quote from the text of the Guerilla Open Access Manifesto:

> There is no justice in following unjust laws. It's time to come into the light and, in the grand tradition of civil disobedience, declare our opposition to this private theft of public culture.
> We need to take information, wherever it is stored, make our copies and share them with the world. We need to take stuff that's out of copyright and add it to the archive. We need to buy secret databases and put them on the Web. We need to download scientific journals and upload them to file sharing networks. We need to fight for Guerilla Open Access.
> With enough of us, around the world, we'll not just send a strong message opposing the privatization of knowledge — we'll make it a thing of the past.
> Will you join us?
> Aaron Swartz
> July 2008, Eremo, Italy

[585] Constant, S.A., 2013. The Computer Fraud and Abuse Act: A Prosecutor's Dream and a Hacker's Worst Nightmare-The Case Against Aaron Swartz and the Need To Reform the CFAA. *Tul. J. Tech. & Intell. Prop.*, *16*, p.231; See also Gould, R., 2013. Aaron Swartz's Legacy. Academe: Magazine of the American Association of University Professors, Vol. 100, No. 1, 2014, online available at https://ssrn.com/abstract=2207299; Eckersley P. 2013. Farewell to Aaron Swartz, an Extraordinary Hacker and Activist, Electronic Frontier Foundation post, online available at https://www.eff.org/deeplinks/2013/01/farewell-aaron-swartz.
There has also been a legal reform proposal to amend the said Computer Fraud and Abuse Act with the introduction of a bill named after the late Aaron Swartz, also known as "Aaron's Law" Available at https://lofgren.house.gov/images/stories/pdf/aarons%20law%20-%20lofgren%20-%20061913.pdf. See also Lessig L. Aaron's Laws: Law and Justice in a Digital Age. Harvard Law School Appointment Lecture available at http://today.law.harvard.edu/lessig-on-aarons-laws-law-and-justice-in-a-digital-age-video/.

More recently, in March 2016, the news reported on the case of Alexandra Elbakyan, another young (27 years old then) "hacktivist." Alexandra Elbakyan, originally from Kazakhstan, founded "Sci-Hub", an online repository containing nearly 50 million academic articles apparently uploaded after being accessed through .edu proxies.[586] Elbakyan, like Swartz, has been vocal against the enclosure of scholarship. She ended up getting sued by Elsevier for copyright infringement in the New York court system. The fines she faces are far from negligible. The US District Court, Southern District of New York, granted Elsevier preliminary injunction in October 2015. Judge Sweet for the US District Court, notes in the injunction that "[i]t is highly likely that the Defendants' activities will be found to be willful - Elbakyan herself refers to the websites' activities as "pirating" […] - in which case they would be liable for between $750 and $150,000 in statutory damages for each pirated work."[587] Elbakyan's case led to a new round of public debate about accessibility of scholarship online and how views and acts of digital natives come into direct contradiction with the current copyright framework.[588]

The above cases indicate a general dissatisfaction with access to knowledge, and in particular access to academic resources. There is widespread understanding that such systematic online sharing of content, including copyrighted content without required consent, is filling a real gap in availability. The research conducted by Balázs Bodó, who gathered data directly from one of the biggest online pirate library, or 'shadow' library, is illuminating. Bodó examined the case of Gigapedia, which later became Library.nu, the leading online pirate library for English-language works before its shutdown in 2012.[589] Aleph is the name that the author uses for the Russian-language pirate library, to which Gigapedia's collections migrated. The empirical work he conducted on the supply of works in Aleph indicates that downloading works from the pirate

[586] Rosenwald M.S., This student put 50 million stolen research articles online. And they're free. The Washington Post, March 30, 2016, online available at https://www.washingtonpost.com/local/this-student-put-50-million-stolen-research-articles-online-and-theyre-free/2016/03/30/7714ffb4-eaf7-11e5-b0fd-073d5930a7b7_story.html?hpid=hp_hp-more-top-stories_journalwars-1140am%3Ahomepage%2Fstory.

[587] See: *Elsevier Inc. et al v. Sci-Hub et al*, No. 1:2015cv04282 - Document 53 (S.D.N.Y. 2015).

[588] Waddell K. The Research Pirates of the Dark Web. *The Atlantic*, February 9, 2016, online available at https://www.theatlantic.com/technology/archive/2016/02/the-research-pirates-of-the-dark-web/461829/.
See also: McNutt M. My love-hate of Sci-Hub, *Science* 29 Apr 2016: Vol. 352, Issue 6285, pp. 497.

[589] Bodó, B., 2014. In the Shadow of the Gigapedia-The Analysis of Supply and Demand for the Biggest Pirate Library on Earth, online available at: https://ssrn.com/abstract=2616633 [Forthcoming as The shadow of Gigapedia: quantitative analysis of shadow library usage. in: Karaganis (ed): *Shadow libraries*]
See also: Bodó, Libraries in the post-scarcity era, supra note 45.

collections correlates with availability.[590] Bodó and his research team reviewed the legal availability of titles offered in the Aleph catalog, using Amazon.com as representative of legal availability in print form and worldcat.org for e-library availability. They found that Aleph is arguably filling a gap in the global scientific publishing ecosystem and mostly serving readers in developing countries where the catalogue titles were not available legally or where legal e-library subscriptions are not affordable in the academic environments of those countries.

The study's data indicate that more than 55,000 publishers are represented in Aleph's collection and the biggest western academic publishers are dominant, including Springer, Cambridge University Press, Routledge, Wiley, Oxford University Press. The works coming from the biggest publishers, however, do appear to be the most downloaded. Furthermore, the largest collections, divided into sectors, are not necessarily the most downloaded, with mathematics and philosophy being apparently in highest demand.[591] The study takes account of the share of downloads by date of publication and then measures the availability of the catalog's titles in Amazon.com and worldcat.org. From the data collected it concluded that while print availability of the titles is high, reaching at times 83%, electronic availability is low with huge gaps that seem to improve for the most recently published works. Another interesting finding is the very low eBook availability for natural sciences and mathematics' titles, which constitute "the core of Aleph's collection."[592] Also, from the legal sources researched, it is concluded that only subscription-based e-library availability could compete with the supply of Aleph in terms of electronic access. These subscriptions are normally institutional subscriptions and are prohibitively high for individuals surfing online. The study also finds a correlation between price and demand in pirate catalogues.[593]

From their research on the demand question (data acquired from the administrators of a commercial mirror of Aleph), it can be derived that the primary audience for Aleph sites comes within academic communities. In addition to this, some countries appear to be heavier downloaders than others. Taking into account that most content available in Aleph sites is in English and Russian (with German content apparently growing) the study ultimately claims that Aleph plays multiple roles in the scientific publishing ecosystem. For example, it fills a

[590] Id. Bodó, 2014. In the Shadow of the Gigapedia-The Analysis of Supply and Demand for the Biggest Pirate Library on Earth, p. 16 (SSRN version, supra note 589).
[591] Ibid, p.8
[592] Ibid, p. 11.
[593] Ibid. p.14.

developmental gap between downloaders from "countries that have strong western ties and an even stronger desire to catch up with the West." These are countries in central and Eastern Europe and countries in deep financial crisis, mostly the European south. Another connecting factor in the findings is the comparatively poorer academic infrastructure of these countries.

Most importantly, Bodó argues that there are lessons to be learnt from online pirate libraries.[594] In the digital era, when the book takes digital form and is therefore no longer scarce, shadow libraries directly compete not only with the market for books but also with libraries themselves. Shadow libraries such as Aleph are, according to Bodó, set out to completely remove sources of book scarcity; both physical obstacles that technology promises to remove with digitization and legal ones like copyright restraints. Legalization of file-sharing for books is the direction where his argument is heading since the evidence he gathered indicates a clear correlation between downloading and digital unavailability in legal markets as well.[595]

7.2.3 How (un)availability of books leads us back to the discussion on the needs of the poor

Previous chapters in the study portray a series of complex considerations behind the legal rules of copyright that are applicable to books. Copyright has been conceptualized as rewarding the unique moments of inspiration of authors, as creating a market for intellectual production, as a second best solution to incentivize authorship and as the outcome of a delicate economic balancing between incentives and access. After the book's commercial lifespan before and after copyright expires, some books find themselves in dark places, often orphaned or out of print.

---

[594] Bodó, Libraries in the post-scarcity era, supra note 45:

> Book pirates compete with some of the core services of libraries. And as is usually the case with innovation that has no economic or legal constraints, pirate libraries offer, at least for the moment, significantly better services than most of the libraries. Pirate libraries offer far more electronic books, with much less restrictions and constraints, to far more people, far cheaper than anyone else in the library domain. Libraries are thus directly affected by pirate libraries, and because of their structural interdependence with book markets, they also have to adjust to how the commercial intermediaries react to book piracy. Under such conditions libraries cannot simply count on their survival through their legacy. Book piracy must be taken seriously, not just as a threat, but also as an opportunity to learn how shadow libraries operate and interact with their users. Pirate libraries are the products of readers (and sometimes authors), academics and laypeople, all sharing a deep passion for the book, operating in a zone where there is little to no obstacle to the development of the "ideal" library. As such, pirate libraries can teach important lessons on what is expected of a library, how book consumption habits evolve, and how knowledge flows around the globe.

[595] See also, the IVIR project *Copyright in an Age of Access: Alternatives to Copyright Enforcement feasibility of legalizing file-sharing* – publications available at https://www.ivir.nl/projects/copyright-in-an-age-of-access-alternatives-to-copyright-enforcement/.

When copyright expires all books enter into the public domain. The public domain is mostly celebrated as a space of freedom and useful resources but still a place that includes well-known and much read works together with a big number of less popular or less read ones. The existence of libraries adds another layer to the complexity of the ways in which the legal system treats books. While copyright allows for the treatment of books as market goods, such as meat or potatoes, special exceptions allow for the operation of libraries where the same good enjoys a life beyond the market. Indeed, the last section of our first chapter discussed the historical connection between public libraries and 'the needs of the poor.'[596] In the same chapter we also pointed to the paradox of an official rhetoric that copyright laws have created and maintained, which is the encouragement of learning.

It is on the basis of this paradox that, I believe, we must interpret phenomena of piracy discussed in the previous sections. File sharing, activist efforts to illegally download content and systematic shadow libraries are all connected to the one discussion: the oxymoron of witnessing unavailability of content in the digital age not because of technical issues but because of legal obstacles. Given the technical possibility to overcome access barriers, it is to be expected that shadow library phenomena will continue to proliferate. In this situation, one can see the role of traditional libraries as a potential counter-balance. Publically accessible libraries have full potential to contribute to more equal access to knowledge for those who cannot afford market access. While this gap is currently filled by a number of websites operating in the shadows of the Internet, it can also be more systematically addressed by allowing libraries to provide more access and better services in the digital age. The birth of the 'digital library' directly linked to the institution of a physical library has been the primary focus of this dissertation. One can have legitimate hopes that libraries can legally fill the access gaps as they adapt to the digital age, digitize their collections and develop practices allowing online access to a broader readership, hopefully a public readership without barriers.

The following and concluding chapter of the dissertation concentrates on such hopes related to an enhanced role for libraries in the digital environment. Technically, digitization promises and can perhaps deliver the realization of dreams in a world library accessible to all, possibly also in egalitarian terms close to what Aaron Swartz or Alexandra Elbakyan imagined.

[596] *Infra* 1.5 *Public libraries and 'the needs of the poor'.*

As we have seen,the reality has, thus far, been more antagonistic and is defined by the restrictions of the copyright framework already in place.

7.3 ARE BOOKS OR LIBRARIES BECOMING OBSOLETE?

In view of the transformations that the book is experiencing in the digital age, the distinction between an 'old' and a 'new' world is usually employed. This is a theme that we have seen throughout the thesis as we constantly differentiate between physical or paper books and digitized or e-books and between 'traditional' or 'brick' libraries and digital libraries. Could the transformation signal something far more radical for the medium and institution of the 'old' world? As per the title of this section: do these developments signal the end of physical libraries? Can the Internet and the availability of materials in digital form ultimately replace the need for books and/or libraries in the traditional sense? At least from what we see today, the direction has been on the contrary. Below we will first discuss the matter vis-à-vis books and subsequently vis-à-vis the libraries.

There are arguments that the younger generations of digital natives have a different book culture whereby the physical book will be replaced with tablets and other screens.[597] Besides the obvious environmental concerns about the sustainability of paper production I think that the current state of the market does not indicate the end of the paper book but only a reduction in its consumption. There is also an interesting comeback to physicality, with Amazon for example opening its first physical stores in 2017.[598] The store chain, called 'Amazon Books' is advertised as 'the physical extension to Amazon.com'. Thus ultimately the digital and the physical components of the book experience seen to be more complementary than competitive. In fact, the e-book market which we saw rapidly expand is reportedly declining lately. Notably in 2011 Amazon.com officially announced that it sells more kindle books than print books.[599] In 2012

[597] Palfrey, and, supra note 19.
[598] Alter A. and Wingfield N., A Trip Through Amazon's First Physical Store, *The New York Times*, March 10, 2016, online available at http://www.nytimes.com/2016/03/12/business/media/a-virtual-trip-through-amazons-physical-store.html?_r=0. See also Amazon's announcement at https://www.amazon.com/b?node=13270229011.
[599] The Amazon press release, *Amazon.com Now Selling More Kindle Books Than Print Books* (May 19, 2011) is available at: http://phx.corporate-ir.net/phoenix.zhtml?c=176060&p=irol-newsArticle&ID=1565581&highlight. See also Cain Miller C. and Bosman J., E-Books Outsell Print Books at Amazon, *The New York Times*, May 19, 2011, online available at http://www.nytimes.com/2011/05/20/technology/20amazon.html.

Amazon.co.uk made the same announcement.[600] Since 2015, however, several sources noted a decline in e-book sales and even talked about a growing 'digital fatigue'.[601] The conclusion based on figures provided by publishers associations and reported on in the news, are nevertheless contested, particularly in view of the pricing and antitrust wars that the industry has been fighting recently.[602] Whether 'digital fatigue' is an actual phenomenon or a temporary observation bound to change as generations of digital natives take over the markets is perhaps the million-dollar question to ask for the future of books in general.

The 'come-back' of physicality is also noted with respect to libraries. Recently, there has been rebooted attention given to libraries which suggests that the reports on library's death have been greatly exaggerated. "The reports of my death have been greatly exaggerated" has been the title of an event organized by the American Institute of Architecture in New York in October 2016. The theme of the event was the libraries of the future. It included various workshops discussing the various projects on renovations and building new libraries throughout the United States.[603] The clear message of this gathering of architects, librarians, representatives of several academic institutions and also representatives of sponsoring foundations, was that libraries are far from dead in the digital age: the physical space of a library continues to *thrive and evolve*. At the same time, there has been a significant amount of rethinking of library spaces in public libraries and academic libraries, because of the transformation in the ways in which patrons access - and expect to access – knowledge in the digital era. While tons of information is available on the web, which makes the need for physical research browsing library shelves less of a priority, libraries are far from obsolete as spaces that offer enhanced information service and resources, specialized material, organized material, Internet access and an environment that allows studying and social interaction.

---

[600] Malik S., Kindle ebook sales have overtaken Amazon print sales, says book seller, *The Guardian*, August 6, 2012, online available at http://www.theguardian.com/books/2012/aug/06/amazon-kindle-ebook-sales-overtake-print.

[601] Milliot J., As E-book Sales Decline, Digital Fatigue Grows, *Publishers Weekly*, June 17, 2016, online available at http://www.publishersweekly.com/pw/by-topic/digital/retailing/article/70696-as-e-book-sales-decline-digital-fatigue-grows.html; Flood A., Ebook sales falling for the first time, finds new report, The Guardian, February 3, 2016, online available at https://www.theguardian.com/books/2016/feb/03/ebook-sales-falling-for-the-first-time-finds-new-report.

[602] See for example the following contestation: Ingram M., No, e-book sales are not falling, despite what publishers say, Fortune, September 24, 2015, online available at http://fortune.com/2015/09/24/ebook-sales/.

[603] *The Reports of My Death have Been Greatly Exaggerated: Libraries for the Future*, AIA New York, Center for Architecture, October 29, 2016, program available at http://cfa.aiany.org/index.php?section=calendar&evtid=9882.

There are interesting stories of community engagement and advocacy related to libraries that prove the institution's relevance as well as the social demand for the maintenance and boom of traditional library services. New York, for example, is a fascinating place to study community engagement involving libraries. I find recent stories of community mobilization in different boroughs of the city around issues affecting their local libraries quite illuminating. Firstly, the saga around the New York Public Library's reform plans, which included moving books from the main Manhattan building. The Library reconsidered its plans after significant disagreement from patrons who also a filed complaint against the plans.[604] Another controversy has unfolded over plans to sell the Downtown Brooklyn Heights Library.[605] A new library is currently under construction at Hunters Point in Queens, as a result of a fruitful community engagement, supported by the city and the architects that undertook the project.[606] Lastly, there is an ongoing fierce debate in the American public sphere today with regard to the dangers that libraries face if their federal funds are cut as a result of political decisions.[607]

Most importantly, these examples are not standalone or exceptional cases in the US or beyond. The discussion about the future of libraries is live in diverse forums worldwide. To give only a couple of examples one can observe the same energy in the discussions within the International Federation of Library Associations and Institutions. In its activities, publications and support to its member states, IFLA has embraced the digital shift. Libraries have expressed loudly and effectively that they understand and, indeed, experience technological progress, which they see as opportunity rather than a threat. Among the various publications published by the IFLA, the recent trend report, updated within 2016 is indicative.[608] Among the many issues discussed are connectivity and the conversation between libraries' physical and digital services. Importantly, the IFLA focuses on different parts of the world, and adapts the policy discussions

[604] The 'Citizens Defending Libraries Complaint' filed with the courts of the Supreme Court of the State of New York. For a summary of the happy ending of the saga see Pogrebin R., For the Mid-Manchattan Library, a Redesign for the Future, The New York Times (November 16, 2016), at: https://www.nytimes.com/2016/11/17/arts/design/for-the-mid-manhattan-library-a-redesign-for-the-future.html.
[605] Citizens Defending Libraries Blogspot, Love Brooklyn Libraries Files Complaint With Attorney General Eric Schneiderman That Brooklyn Public Library Understates Capital Funds In Order to Sell Heights Library (Sunday, January 31, 2016), online available at http://citizensdefendinglibraries.blogspot.com/2016/01/love-brooklyn-libraries-files-complaint.html.
[606] Dunlap D.W., A New Queens Building That Can't Be Overlooked: A Library, The New York Times (July 20, 2016), at: https://www.nytimes.com/2016/07/21/nyregion/a-new-queens-building-that-cant-be-overlooked-a-library.html.
[607] Templeton, supra note 16.
[608] IFLA Trend Report 2016 - Update available at https://trends.ifla.org/update-2016.

and recommendations according to the specific needs of each library system and infrastructure, such as libraries in rural areas in developing countries.[609] The IFLA is also actively participating in the United Nations 2030 Agenda for Sustainable Development, where access to knowledge is given proper attention.[610] In chapter four, we also saw how the IFLA has been advocating for a WIPO treaty on library exceptions.

For the purposes of this section let us add two examples from the European context, and more specifically from the European South highlighting the relevance of libraries and their future in economies hit by the financial crisis. Our first example is the current focus on the future of libraries in Italy. A ministerial decree on the new single platform, which connects collections from forty-six libraries and more than one hundred archives creating the Italian Digital Library, was announced in March 2017.[611] This has been a centralized effort coordinated by the Italian Ministry of Culture that includes various European projects involved in digitizing collections and is co-financed by EU and Italian funds. The physical libraries on the ground have been absolutely central throughout this enterprise. In addition to this, there are contemporary discussions about institutional collaborations and 'open digital collections' in many fora and contexts.[612] Another hopeful story is that of the latest developments with regard to the National Library of Greece. After periods of underfunding, the library has been revived through generous sponsorship. The library has relocated to new premises in Athens and introduced a number of new research programs, cultural activities and the reviving digitization projects.[613]

This all shows that libraries as an institution are alive and well and do not face imminent extinction. It is, however, an institution conscious of the transitional phase to the digital age and the various challenges. These recent stories are vivid examples of both citizenry and public engagement that positively affect the future of brick libraries, which they use and will continue using in the foreseeable future. The stories also underscore the importance of physical premises

---

[609] Ibid.
[610] See Libraries, Development and the United Nations 2030 Agenda, available at https://www.ifla.org/node/11013.
[611] La Stampa – Technologia, Nasce la Digital Library d'Italia: online 101 archivi e 46 biblioteche Il ministro Franceschini annuncia una piattaforma digitale unica per i vari istituti italiani: sarà finanziata con 2 milioni di euro, March 10, 2017, online available at http://www.lastampa.it/2017/03/10/tecnologia/news/nasce-la-digital-library-ditalia-online-archivi-e-biblioteche-FtHZ2lMVzgwErudo5RSMbJ/pagina.html.
The portal to the digital catalogues ("Internet culturale" Cataloghi e collezioni digitali delle biblioteche Italiane) is available online at http://www.internetculturale.it/opencms/opencms/it/.
[612] La Biblioteca Aperta - Tecniche e Strategie di Conivisione Convegno Biblioteche 16-17 Marzo 2017.
[613] On the recent relocation and future plans, see the announcement from the sponsoring Stavros Niarchos Foundation Cultural Center, available at https://www.snfcc.org/about/vision/the-national-library-of-greece/.

and physical books for the patrons. Despite the challenges, libraries continue to form an integral and valuable part of civic life in rural areas where the need for community spaces and communication hubs is vital as well as in urban spaces where the social and educational role of libraries is essential for large groups of more and less vulnerable members of a diverse society.

CONCLUSION – WHAT'S THE RUSH?

Cioè, la funzione ideale di una biblioteca è di essere un po' come la bancarella del bouquiniste, qualcosa in cui si fanno delle trouvailles, e questa funzione può essere permessa solo dalla libera accessibilità ai corridoi degli scaffali.

Umberto Eco[614]

(i.) Preliminary Thoughts

Libraries or Google - *or* why did the thesis focus on libraries?

With a study centered on libraries, this dissertation aimed at complementing an emerging literature on book digitization and the creation of digital libraries. It tells the story of the transition from the physical to the digital library, which not only has digital collections of books but is also fully transforming into an institution operating in the digital space and seeking to offer its services as it did in the physical space. There are many actors involved in this transition but the thesis chose to, first, tell the story from the perspective of the libraries and, second, focus its assessments and now its suggestions on the libraries. The justification for this choice can be found in libraries' longstanding institutional mandate as reserved spaces outside of the book market, which I explore under Chapter three. While we ought to constantly evaluate the need to maintain 'old' institutions, my understanding is that the *moral cost* of eliminating libraries is still significant in view of both persistent inequalities and a labeling of books as merit goods. The

[614] Umberto Eco, De Bibliotheca (1981): *That is to say, the ideal function of a library is to be somewhat like the bouquinist stand (antique bookshop), a place where one uncovers treasures, and this role can only be realized through free access to the stacks of shelves.* [own translation].

thesis places an emphasis on libraries' transformation and not on the corresponding transformation of authorship and publication in the digital age. This is only a primary phenomenological observation, or one view of the cathedral.[615] More views include: the authors', the publishers', and the readers'. In other words, the present analysis is an attempt of a coupling of legal norms, namely copyright exceptions for libraries, and their historical, social and now technological environment: a historical inequality in access to knowledge (Chapter 1), the role that libraries have in society, as reserved spaces outside of markets contributing to book de-commodification (Chapter 3) and a fundamental change that I think the digital age presents (Chapter 5). I argue that this latter change, or evolution, is critical. So critical, that justifies the unpacking of an old balance of interests (before digitization was cheap and widespread) to a new one.[616]

This area of research is still relatively new and comes, I think, with great potential. First of all, it is highly topical. Secondly, it relates to even more contemporary and revolutionary directions that technology is taking us. To give an example, the Google Books saga has not ceased to appear in public debates. While at the final stage of my drafting the thesis this year, I was surprised to read two articles on the issue published in April, one appearing in *Wired* (April 11, 2017) and the second in *The Atlantic* (April 20, 2017). The first, written by Scott Rosenberg, is entitled "How Google Book Search got lost".[617] Rosenberg observes a decline in Google's "drive and ambition" with respect to the Google Books project and updates us on the current status of the project; apparently Google is still scanning but at a much slower pace. From the people he interviewed, he cites "a Google engineer who has worked on Books for a decade and now leads its team", Stephane Jaskiewicz, who said:

> We're not focused on shiny features and things that are very visible to users. It's more like behind the scenes work and perfecting the technology — acquiring

[615] Calabresi, G. and Melamed, A.D., 1972. Property rules, liability rules, and inalienability: one view of the cathedral. Harvard Law Review, pp.1089-1128.
[616] Irrespectively of whether the old balance was efficient or not as regards the production, distribution, and accessibility of books.
[617] Rosenberg S., How Google Book Search got lost, *Wired* [online], April 11, 2017, available at https://www.wired.com/2017/04/how-google-book-search-got-lost/.

> content, processing it properly so that we can view the entire book online, and adjusting the search algorithm.[618]

Rosenberg notes how the Google team realized that without a change in the copyright legal framework they would not be able to give any further access to human users than they provide now - snippets of the millions of books that still do not rest in the public domain. Thus they focused on technological functions instead, such as the search functions. However, where did this leave machine users? Here Rosenberg makes the link between machine learning and the research potential that could be realized if researchers, not only in Google but also perhaps in academia, had at their disposal the large corpus of digitized books to experiment with. One might need to explore at greater length the motivation of the other parties whose involvement we have seen throughout the thesis. To a certain extend, I assume that the motivation behind the involvement of libraries in digitization does not require much explanation, especially when libraries maintain their role in preserving and providing certain access to the books in their collections. The motivation for other actors might perhaps be less clear, Google, for example, is without doubt one of the most important players to focus on when it comes to book digitization and its uses.

The second article is written for *The Atlantic* by James Somers and its title is "Torching the modern-day library of Alexandria". Like Rosenberg, Somers wonders what happened with the Google Book project. He discusses how Google initiated its digitization project, how it got involved in litigation and wonders where are the books now.[619] Somers appears more frustrated than Rosenberg that the digitized books are not accessible for copyright reasons:

> It was strange to me, the idea that somewhere at Google there is a database containing 25-million books and nobody is allowed to read them. It's like that scene at the end of the first Indiana Jones movie where they put the Ark of the Covenant back on a shelf somewhere, lost in the chaos of a vast warehouse. It's there. The books are there. People have been trying to build a library like this for ages—to do so, they've said, would be to erect one of the great humanitarian artifacts of all

[618] Id. (the quote is taken from the article).
[619] Somers, Torching the modern-day library of Alexandria, supra note 17.

> time—and here we've done the work to make it real and we were about to give it to the world and now, instead, it's 50 or 60 petabytes on disk, and the only people who can see it are half a dozen engineers on the project who happen to have access because they're the ones responsible for locking it up.
>
> I asked someone who used to have that job, what would it take to make the books viewable in full to everybody? I wanted to know how hard it would have been to unlock them. What's standing between us and a digital public library of 25 million volumes?
>
> You'd get in a lot of trouble, they said, but all you'd have to do, more or less, is write a single database query. You'd flip some access control bits from off to on. It might take a few minutes for the command to propagate.[620]

The press has covered the Google Books story extensively while the decade-long litigation was ongoing but the fact that the publication of these two articles came in April 2017 - two years after the legal battles have terminated – is significant. So too is the timing of one more publication: an article appeared in the New York Times Magazine in December 2016, entitled "The Great A.I. Awakening: How Google used artificial intelligence to transform Google Translate, one of its more popular services – and how machine learning is poised to reinvent computing itself."[621] As its title suggests, the article discusses developments in artificial intelligence and the transformation of Google Translate. This piece connected the dots for me: Google's interest in enhancing its algorithms, especially when it comes to the processing of natural language and machine learning, might very well relate to its owning a great corpus of digitized books from the collections of some of the world's greatest libraries. This was also hinted at in the aforementioned Rosenberg's piece.

One message that we can take from the above is the ongoing interest in the topic. A second and perhaps more compelling issue is the question posed by the first two articles: what happens with all the digitized books at Google's disposal? If the dream was to create a universal library why are only computer engineers and machines able to 'read' them and to access this corpus in its entirety? Here is the central theme of the thesis conclusion: What are the different motivations

[620] Ibid.
[621] Lewis-Krausdec, The Great A.I. Awakening, supra note 58.

behind each of the parties involved in the phenomenon of digitization rush? As the title puts it, *what was the rush* for each of them? If this is a quest for control over resources (the digitized books), arguably, the party that controls the resources is in the more powerful position. Among the various stakeholders that claim control over the resources, both in legitimate and illegitimate ways, the thesis places a vote of confidence in libraries. There is an additional reason for this, besides the obvious compatibility of their institutional mandate with such a gatekeeping role. They are also trusted institutions dedicated to public service and widespread access to knowledge and, therefore, can counter-balance the rise of information superpowers, like Google. Without doubting the tremendous contributions that the company, together with other technology giants, makes to innovation, it is important to retain safeguards against the creation of bottlenecks and against private control of a public domain (here I use the term loosely, taking into account that also the copyrighted works that are digitized will eventually become part of the public domain).[622]

(ii.) The research question

Answering the research question

In the historical and legal context that the chapters of this dissertation have considered, the research question that we posed is whether the copyright framework in place sustains the role of libraries as the institutional guardians of our literary treasures in the age of digitization. In other words, *what is the position of libraries in the middle of the so-called 'digitization rush'*?

Considering that we live in a time when the information flows and technological connectivity is both fast and cheap, the potential for wide access to knowledge is unprecedented. Assuming that digital technology has the potential to fulfill the promise of a new renaissance, the question is whether libraries can have a central role in this renaissance. On the basis of the thesis

[622] On a side note, the ongoing debate about Google's funding of research, including on intellectual property, reminds us of the debate on how to achieve the delicate balance between public interest and corporate strategies that might also be relevant to our discussion on placing trust in particular institutional players. See Google Academics Inc., Google Transparency Project, 14 July 2017, available at http://googletransparencyproject.org/sites/default/files/Google-Academics-Inc.pdf.
Much of the scholarship cited in the report relates to topics that have been relevant to this thesis. See also the debate that sparked after the publication of the WSJ article: Mullins B. and Nicas J., Paying Professors: Inside Google's Academic Influence Campaign, *The Wall Street Journal*, 14 July 2017 [online], available at https://www.wsj.com/articles/paying-professors-inside-googles-academic-influence-campaign-1499785286.

findings, the present and concluding chapter provides a negative response. As the law stands now, the position of the libraries is not strong and, on the contrary, I further argue that it necessitates further support that could perhaps take the form of more exceptions and limitations to copyright in their favor.

Structure of this conclusion

The chapter is divided into two parts: The first is methodological; it reviews the arguments made in parts I and II of the thesis and summarizes their findings in order to respond to the research question (8.1 Thesis Findings). The second makes normative suggestions on the basis of the overall assessment of libraries' position while setting a clear target: to privilege libraries as institutions and allow them the space to experience their own digital renaissance (8.2 Normative Suggestions). In order for them to do so, this thesis makes the argument for exceptions to copyright which allow digital libraries to be created (digitization should be permissible) and to operate as public service institutions (access by users should also be permitted, perhaps on a user-group scaled basis). The proposals are built on the basis of key findings of the thesis chapters, with particular emphasis on chapters four and six.

8.1 THESIS FINDINGS

8.1.1 The phenomenon of a 'digitization rush'

Reflecting on all digitization initiatives as mapped in chapter six of the thesis, as well as the power relations between the stakeholders involved, we can observe a phenomenon of competition between the various players who aimed to capture what they identify as a new value. The competing projects of Google and then Yahoo and their partners in the Open Content Alliance are illustrative. Because the additional value is computational, it is not surprising that technology companies play a prominent role in the competition. Indeed, digitization has revealed a new value that is worth capturing: that of a digital corpus of books whose content is machine-readable.

Around the new value created by digitization, that of the digitized book sit both old and new interests. If we stick with the property metaphor, the new phenomenon of a 'digitization rush' can be described as a competition over new resources and ultimately over new territories.

As seen in the included timeline, the existence of a new valuable resource was well understood by a number of actors who engaged in digitization projects coming both from the private and the public sector. These actors include (i.) non-profit organizations, such as Project Gutenberg and Internet Archive and subsequently the DPLA; (ii.) academic institutions, such as all those institutions that partnered with Google, including those institutions that have been particularly taking leading roles in driving digitization projects - Stanford University with the Digital Library Project, the HathiTrust partners with the leadership of University of Michigan and Harvard University initiating the DPLA; (iii.) private technology companies – with Google playing the leading role; (iv.) national and international public institutions, from the less extended World Digital Library project of UNESCO and the Library of Congress to the elaborate efforts of the EU resulting in Europeana; and last but not least libraries, as the Bibliothèque Nationale de France and others (also HathiTrust perhaps better fits this category). In addition to these actors, there are more stakeholders to be taken into account that have traditionally embedded interests in books. These include (v.) authors and their associations and (vi.) publishers and their associations. Both groups are institutionally recognized by the existing copyright framework to have absolute claims over the books' value, at least for a long period of time. To the above list, one could perhaps add also (vii.) unauthorized actors who see value in 'freeing' access to materials by defying existing copyright rules: the 'shadow libraries' and all those involved in the online sharing of digital files of books, as we discussed in chapter seven of the thesis.

From the groups of actors enumerated above, non-profit organizations, academic institutions, private companies, public institutions, libraries, authors and publishers had some embedded rights and interests in books and, with digitization, rushed to claim the same or similar interests in the digitized books. We have already made the claim that digitization does not result in merely a copy of a book; it creates an entirely new value. It seems that all the above actors understood this - only some were much quicker than others. Those who were quickest soon found themselves at the receiving end of lawsuits, as defendants of their digitization projects while the slower ones found themselves in the position of the complainants claiming that digitization infringes their rights on the books. Indeed, it was mostly certain authors and publishers that led the wars in the courts in the above cited cases of against Google and also against HathiTrust and the Technical University of Darmstadt.

Depending on their motives and mandates each of the actors has been either competing against the other actors, active or potential ones, to capture this value first and secure some kind of exclusivity or is acting less competitively driven by other mostly altruistic or public service purposes. Even though not all actors are *stricto sensu* competing, precisely because some have engaged in digitization for alternative motives and are not claiming exclusivity, it is helpful to conceptualize their interaction as one of *lato sensu* competition over the new resource. Even the non-profit initiatives and the public initiatives, which are not competing for a primary or secondary market that the new value hypothetically creates, they are still competing for visibility and use in order to stay relevant, public funding or donations to cover their costs and so forth. In this sense it remains accurate to describe the above setting as one of competition between all the relevant actors in capturing the new value of digitized material. This is in a nutshell a first level of competition that this thesis observes.

At a second level, more institutional players are added, this time 'competing' (again *lato sensu*) to provide protection or guardianship to rights and right-holders that they consider relevant. While non-profits, academic institutions and libraries, private companies and public institutions have all been competing to capture the new value of digitized literary works, more institutions have joined in an attempt to claim guardianship over the works that are being digitized as well as to protect the rights of the authors as creators and rights holders of these works. This second level of competition is still competition over the value of this new resource only this time instead of engaging in digitization some players simply tried to either prohibit others from digitizing or to define the rules under which digitization can take place. Thus three more actors come into play: first, courts, second, legislators and regulators, and third, the greater public. Ultimately all these players, those who immediately tried to capture the value that digitization brought about and the ones that tried to regulate it aimed at the same goal: to define the rules applicable to digitization and therefore to define how digital libraries are to be created. In this manner they would benefit from capturing or controlling the new value that digitization has created.

8.1.2 What's the rush?

One cannot help but wonder, what are the different motivations of each of the parties involved. Indeed, there are significant technological advantages that the involved parties gained by

engaging in digitization projects that go beyond copyright and other questions related to the book market. Even though copyright has been at the center of the legal debates so far, it is worth looking at what is at stake here beyond copyright and beyond the books themselves. We already looked at the example of Google. Sitting on its California goldmine, Google first of all uses the digitized corpus to feed its search engine. There are many more uses however that perhaps help us better explain what other motivations the company could have to engage in digitization and fight the legal battles it fought.

As discussed in the beginning of this chapter, Google today appears to have a significant advantage not only in search but also in data mining, language processing and particularly translation services, which may give the company an important precedence on machine learning applications.[623] There are plausible links between the recent advancement in language processing including translation services and Google owning a significant and rich digital corpus of books. Such advantages at an early stage, especially in the hands of strong monopolies can prove crucial for future technological applications. One can think of many important applications related to language processing and translating algorithms. Translation is crucial for every dimension of communication at a global level, touching areas from global trade to international and global law.

It is very difficult, however, to make predictions about the future of such technologies as the contemporary and much discussed advances in machine learning and artificial intelligence. All the more, it is very difficult to find connections and draw conclusions on the basis, for example, of copyright legal battles. This becomes today the next challenge for legal scholars who study the intersection of law and technology. Nevertheless, the complex legal portrait we have discussed thus far gives us a hint as to possible connections between copyright battles and what else could be at stake in the context of digitization. The possibility that in the near future we would have to deal with the coming *digital dark ages,* as Richard Whitt warns in a recent publication, is a good example of imminent priorities.[624] While our analysis in this thesis has been doctrinal and relies heavily on the interpretation - and criticism - of copyright rules, the related law and technology questions I believe lead the way for a next frontier of research.

[623] Lewis-Krausdec, The Great A.I. Awakening, supra note 58.
[624] Whitt, R. Through A Glass, Darkly" Technical, Policy, and Financial Actions to Avert the Coming Digital Dark Ages, 33 Santa Clara High Tech. L.J. 117 (2017).

8.1.3 The position of libraries in the middle of the rush: Comparing Parts I and II of the thesis

Studying the phenomenon of digitization at a descriptive level and focusing on the legal and policy issues around it, one cannot help but notice the mismatch between the situation before and after mass digitization was possible. Part I of the thesis, on the one hand, wants to portray a rather functional story of the historical evolution of copyright laws as applied to books and to libraries. The story portrayed in part II, on the other hand, is disappointing. Despite the vast potential that digitization opens up for libraries, legal obstacles are the main reason why libraries are having a hard time realizing that potential. Libraries that initiated digitization projects faced a complex situation as regards the copyright clearance of a substantial part of their collections. Even worse, libraries found themselves involved in serious legal battles and fighting for their privileges and rights while struggling to compete with other players in what we called a 'digitization rush'. In other words, a general observation coming from the structure of the thesis is that the context that the first part concentrates on and the reality of the law as examined in the second part, are surprisingly different. Let us examine this in more detail.

The first part of the thesis has been contextual and described how copyright laws applied to physical books and to libraries. It has provided a historical account on why books were treated differently than other market products and on how the book market was created with the assignment of property rights over intellectual works. It has shown that one of the main reasons for the introduction of copyright was to fight censorship. Finally, it has focused on the historical function of libraries, providing, in particular, an account on the creation of the first public libraries for 'the poor'.

Copyright created a market for books. While the book became a commodity in the market, libraries, in contrast, remained a respected institution outside the book market where books were freely accessible. Thus, part one tells a success story for libraries: accumulating a wealth of physical books and having become gradually accessible to a broad public, libraries are well respected for the positive externalities they generate. They play an important role in de-commodifying books and equalizing access to knowledge. This is why there are considerable moral costs associated with the elimination of libraries, similar to the moral cost that comes with poor or inexistent public education or the public healthcare system. Libraries also play an important long-term role as memory institutions because they preserve materials for future generations.

The space given to libraries to operate outside or on the margins of the book market is in line with the original purpose of copyright, which has been the encouragement of learning. Indeed, if the first priority of the law is to incentivize creation not for its own sake but to encourage learning and the advancement of knowledge in the broadest sense, then the space given to libraries is appropriate. The evolution of copyright rules, however, in particular copyright maximization as experienced in the past decades, have betrayed the original purpose of the law and created inefficiencies in the form of unnecessary barriers to access which also affect libraries. Chapter four exemplified the inefficiencies resulting from an excessive copyright protection, with regard to orphan works and to works out of print. In view of the copyright length extension, the book market inflated to the point that books with little commercial value would still be 'locked' in property rights and become fairly inaccessible. This situation directly affects libraries that hold such books in their collections. They cannot freely make (digital) copies of either orphan works or out of print works in their collections. The majority of these books have no active right-holders to negotiate and seek permissions from. Nevertheless, libraries cannot yet treat them as books belonging to the public domain.

Indeed, chapter four makes the transition to the challenges that the existing legal framework poses in the digital age; challenges that were already in place but are heightened in view of digitization. To be more precise, digitization brings about a lot of potential for library collections that cannot easily be realized due to excessive copying restrictions (such as the above-mentioned failures with regard to orphan works and out of print works). Looking also at the issue of e-lending, we saw more challenges in providing access – not only preservation. Indeed, chapter four and various sections in Part II is where the thesis switches to a more negative narration of a series of problems that arise in particular at the age of digitization.

Part II, focusing on digitization, is a story of great technological potential but also serious obstacles. While the expectations for access are getting ever higher, the availability of books is lower than expected. Shadow libraries have quickly taken the role of access providers while institutions struggle to transform into digital libraries. Even though the Internet did not kill libraries, as some had feared or predicted, the potential of the digital age for online access to knowledge is not so easy to realize. Chapter six demonstrated how libraries engaged in important digitization projects perhaps taking legal risks as, indeed, some of them led to litigation. The relevant case-law, mainly the American case against HathiTrust and the European case against

TU Darmstadt, reveal some of the most pressing legal issues for libraries. Furthermore, the European case initiated by the Dutch library association confirms how confusing the current framework is when it comes to e-lending. In addition, the litigation in chapter six shows how delicate the balance of interests is; this was perhaps to be expected if we think of the picture that chapter four had already portrayed, including the various discussions for legislative reforms.

Therefore, I conclude that there are still legal obstacles in place impeding the role of libraries as they struggle to remain relevant also in the digital era. These obstacles are well understood both by scholars and by policy-makers. This explains an emerging scholarship in the field, very rich in its analysis of certain issues - such as the various studies and suggestions regarding orphan works. This further explains recent reform efforts as we saw them for example under chapter four.[625]

My overall conclusion on the basis of this study is that libraries operate on *unclear legal grounds* in their transition to the digital age. To make an assessment of their position compared to the rest of the stakeholders also interested in digitization and in some form of digital library creation we return to the normative benchmark set in the beginning of the thesis: Can libraries promote *learning* when they face difficulties in digital preservation and access? Our assessment here is relative to the position of private stakeholders (mainly technology companies) who can perhaps better afford the risk of litigation when they engage in mass digitization and to those who defy copyright rules completely and operate shadow libraries. These stakeholders appear to be in a better position if they wanted to provide access and encourage learning. Yet, it is not usually part of their direct mandates. The argument of this dissertation has been that the position of libraries needs strengthening as a matter of institutional trust in view of what I understand as an effort to mitigate the moral cost of eliminating libraries. I think that their role in mitigating market failures and enabling access to knowledge, thus encouraging learning on egalitarian terms, is and is perceived as crucial. Thus, among the many actors that claim the role of the institutional guardians of our literary treasures in the digital age, libraries' dedication to public service makes a compelling case for their legislative support.

[625] European legislation is trying to create more exceptions that will benefit libraries, as cultural institutions, and give them more space for maneuver with regard to the digitization of orphan works (passed legislation) and works out of print (proposed legislation). The US courts, rather than legislators, have been trying to achieve similar goals employing the fair use doctrine. Besides digitization, electronic access is based on even less solid legal ground. We have seen how both access to dedicated terminals within library premises and e-lending schemes were also litigated.

8.2 NORMATIVE DIRECTIONS

In this context, policy-makers need to make choices as to the future of the digital library. Perhaps before that, they need to decide which of the involved players are institutionally equipped to enjoy legal privileges. Here is a first argument *a maiore ad minus* in favor of libraries and regarding digitization: if Google is now *de facto* in possession of a large corpus of digitized materials, then libraries as institutions dedicated to the preservation and dissemination of knowledge for centuries, if not millennia, should be able at the very least to digitize and safeguard their own cultural record. This dissertation argues in favor of libraries enjoying *digitization rights without any restrictions*. It furthermore proposes privileges when it comes to the providing of *access*. If trusted with such a mandate, libraries can play the key role as an access-facilitator in the digital environment.

In addition, to give real libraries the chance to experience their own renaissance, instead of struggling to compete with other players interested in the digitization of books for their own purposes, we need to define 'digital libraries'. A digital library can be defined narrowly by the combination of its digital components – the digital corpus – and its institutional components – the performance of library functions and services. Then, it might be questionable whether all the digitization initiatives, as we saw them under the timeline, indeed constitute digital libraries. Perhaps the biggest doubt lies with Google Books, which does not offer the library functions and services that digitization projects run by libraries do. In contrast to the Google example, HathiTrust is a digital library project that can arguably be seen as a model for the combination of digital and institutional components. While, I believe, there is merit in sustaining the institutional or professional distinction - as per sustaining librarianship and the involvement of information scientists, a narrow definition entails a slippery slope: to exclude innovative private digitization initiatives from a revisited legal privilege. Last but not least, there is significant promise in the collaboration between libraries in the creation and management of their digital collections (see the HathiTrust model) as well as aggregation of different digitization initiatives under an interoperable model (see the Internet Archive, DPLA, Europeana models and other initiatives of large or small scale).[626] Thus, the suggestions that follow are meant to apply to individual libraries, library consortia and other digitization initiatives that fulfill a threshold of

[626] On this point see also Wu M., 2016. Collaborative Academic Library Digital Collections Post-Cambridge University Press, HathiTrust and Google Decisions on Fair Use, 1 Journal of Copyright in Education and Librarianship.

systematization and public service mandate analogous to what this thesis reads as the one usually associated to the libraries as (brick) institutions.

8.2.1 Suggestions to promote book digitization, libraries and *their* digital renaissance

From the many potentially different ways one could promote both book digitization, thus ensure digital preservation and possibilities for access to future generations, and the position of libraries, this final section examines the option of legislative reform. Another way to strengthen the libraries' position and also promote digitization would be to rely on courts when they adjudicate cases such as the ones analyzed in the thesis:

i. *Authors Guild v. Google*
ii. *Authors Guild v. HathiTrust*
iii. *Technische Universität Darmstadt v. Eugen Ulmer*
iv. *Vereniging Openbare Bibliotheken v. Stichting Leenrecht*
v. *Marc Soulier & Sara Doke v. Premier ministre & Ministre de la Culture et de la Communication*.

As per the outcome of these cases, judges arguably embrace the potential of digitization and have mostly acted favorably towards libraries as clearly demonstrated in the case against HathiTrust and the Dutch *Vereniging Openbare Bibliotheken* case. In these cases, as well as in the TU Darmstadt case, the judiciary granted the vote of confidence to libraries, which is in line with the argument of this thesis. However, as the latest French case, *Marc Soulier & Sara Doke v. Premier minister & Ministre de la Culture et de la Communication*, suggests there are limits to judiciary discretion when it comes to interpreting the copyright framework in place. The same applies to the fair use doctrine, which would be applicable in future American cases. Thus, there are limits also in expecting that libraries, which are often underfunded and usually risk-averse institutions, will keep pushing the boundaries of copyright for the obstacles they face, relying solely on favorable precedent in case they get sued. By contrast, as illustrated in the Dutch case, *Vereniging Openbare Bibliotheken,* libraries are themselves willing to clarify the limits of the legal provisions, by choosing to initiate proceedings rather than rest in the defense position.

Nevertheless, perhaps a more drastic measure to strengthen the position of libraries would be to carry out legislative reform.[627] While concluding, here I will argue in favor of this method for a number of reasons. Legal reform would enhance clarity, which is very much needed for libraries at this stage. It would also bring legal certainty. On a relatively optimistic note, James Boyle recently wrote that "it is possible to break the logjam apparently dictated by collective action problems, ideologies of maximalism, impenetrable subject matter, revolving doors and so on…It is not impossible for us to reform copyright law to make it a little more rational – just very very hard."[628] Momentarily putting aside the question of how hard reforms would be, I will attempt to put forward two suggestions – a specific suggestion regarding digitization (input) and a more general suggestion with regard to access (output). The first contends that physical libraries should be granted the right to digitize their collections. Secondly, once they have created digital collections, libraries should be given more flexibility to provide access to patrons. Let us examine the suggestions in more detail.

There are a number of choices that policy-makers and legislators need to make when thinking about reforming the rules applicable to libraries. They have a range of options: (i) allow libraries to operate as they have been doing so far on the basis of existing copyright exceptions and limitations, but not introduce new exceptions to permit the digitization of collections or the provision of digital access; (ii) introduce specific exceptions for libraries and allow them to digitize their collections under specific requirements – this is the current European policy, which started with the Orphan Works Directive and will most probably continue covering also works out of commerce; (iii) take a step further and allow not only the digitization of collections but also mechanisms for libraries to provide access to digitized material. Again, the example of the EU is indicative for the directions that the third option can take, as shown by the e-lending schemes discussed by the CJEU in the context of *Vereniging Openbare Bibliotheken.*

The first of these options would perhaps not fit the direction that the law seems to be taking in the digital age. Allowing digitization for purposes of preservation is the closest option to the direction of copyright reforms as they are being discussed today: reforms which introduce copyright exceptions and limitations to better fit the reality of the digital age at least with respect

[627] On the potential but also the limitations of copyright reform see in particular Boyle's thoughts on the Hargreaves Review of Intellectual Property in the UK: Boyle J., (When) Is Copyright Reform Possible? in Okediji, R.L. ed., 2017. *Copyright law in an age of limitations and exceptions*. Cambridge University Press.
[628] Ibid, p. 233.

to preservation. This is the case for example with the discussed revision of the library and archives exceptions in the US Copyright Act (17 U.S.C. Section 108).[629] In September 2017, the US Copyright Office published a discussion document that includes a 'model statutory language.'[630] Notably proposed Section 108 (c), under the subtitle Preservation and Security, states:

> An institution eligible under subsection (a) may reproduce each copy or phonorecord of a work currently in the collection of that institution as many times as is reasonably necessary for preservation and security.
> In addition –
> (1) A reproduction made under this subsection, if made from a work not disseminated to the public, may be made available to the public only – (A) on the premises of the eligible institution; (B) by lending a physical copy or phonorecord to a user; or (C) by providing access to a digital, non-physical copy or phonorecord to a single user at a time, for a limited time.
> (2) such copies and phonorecords, if made from works lawfully disseminated to the public, may be made available only on the premises to employees of the eligible institution, but not to members of the public.

I think that, indeed, the benefits of granting digitization rights for the purposes of preservation outweigh the costs imposed on right-holders especially for the categories of orphan and out-of-print works. While for books that are still commercially available, more careful economic assessment is required, one can expect that the adoption of such rule would force the value of physical and digital versions of books to merge into one for new collections acquired by libraries. In a sense, this readjustment of value is already happening in many book markets. [631]

[629] Library of Congress & Copyright Office Notice of Inquiry, Section 108: Draft Revision of the Library and Archives Exceptions in U.S. Copyright Law, Docket No. 2016 – 4, available at https://www.gpo.gov/fdsys/pkg/FR-2016-06-07/pdf/2016-13426.pdf.
[630] Section 108 of Title 17, A discussion document of the register of copyrights (September 2017), available at https://www.copyright.gov/policy/section108/discussion-document.pdf.
[631] See Katz, A. 2016. Copyright, Exhaustion, and the Role of Libraries in the Ecosystem of Knowledge. 13(1) I/S: *A Journal of Law and Policy for the Information Society* 81. Katz offers a rich analysis of how to reconcile digital exhaustion with technology potentially enabling profitable price discrimination schemes to match heterogeneous demand for books. Katz argues that even if perfect price discrimination

For the above reasons, my proposal would be that libraries should be vested with the right to digitize their collections. Once the legal right is clear, thus with a greater degree of legal certainty as per digitization, the debate can be refocused on the practical questions of digitization costs as well as the technological protection measures required.

Going a step further, I would propose that reforms should not only focus on input (digitization) but also output (access). There can be different models for providing access to digitized copyrighted works. Taking part in an ongoing discussion, I want to place on the table a user-group approach whereby restrictions to access the digitized versions can be adjusted to a flexible scale of different sets of 'readers': readers inside or outside the premises of a library, readers with or without a library membership (thus there is a need for clearer lending rights and rules regarding digital donations), institutional readers (including clearer rules for digital interlibrary loans), readers that can claim another copyright exception (i.e. claiming fair use) and special groups such as the blind and visually impaired communities that can now work together with a number of institutions, including libraries, towards the practical implementation of the Marrakesh treaty. The proportionality assessments included in the fair use or the three steps test criteria are presumably slightly different for each different group.

Access-centric options are arguably the most favorable for libraries and could prove transformative for the institution. User-centric options are arguably most favorable for patrons as regards their needs and demands for digital access to existing collections. I also believe that promoting exceptions that enhance access to knowledge would better adhere to the normative benchmark, the encouragement of learning, which we have discussed during the introduction and also in chapter one of this thesis. Thus, an access-centric exception in favor of digital libraries could be interpreted as a social policy choice. It would allow the institution of the public library to function in similar fashions in the digital, as it also does in the physical world.[632] An access-centric exception would, furthermore, assist libraries to claim an infrastructural role in the digital

becomes possible it is not necessary desirable, as it requires data collection and surveillance to readers. Libraries remain all the more important as they indirectly assist price discrimination in the market, enabling both the bundling and the sharing of content while avoiding the costs that direct price discrimination would have.

[632] For an account of the various justifications for copyright exceptions and limitations see Samuelson P., Justifications for Copyright Limitations and Exceptions, in Okediji, R.L. ed., 2017. *Copyright law in an age of limitations and exceptions*. Cambridge University Press.

environment, similar to the one we have seen them occupy in the physical space.[633] Under a more favorable regulatory framework, libraries could become a new digital infrastructure in addition to the existing content layer of the Internet.[634] Chapter five of the thesis referred to the ambitious ideas of the technologically assisted compilation of human knowledge as conceptualized and articulated even by visionaries of the pre-digital era.[635]

Summing up, in order to achieve these goals, I contend that we need to 1. shift our focus away from copyright obstacles as regards digitization (input) and 2. allow libraries to experiment with schemes providing access to their digital collections, reflecting various situations such as on-site consultation as well as interlibrary loans and lending (output). Ultimately, access-centric and user-centric exceptions would also exempt libraries from the ongoing 'copyright wars' on the basis of their public service mandate.[636]They would encourage the treatment of libraries not only as memory institutions engaging in the collection and preservation of our literary wealth, but primarily as access institutions.[637] In the words of Umberto Eco it is important to have libraries with open rather than hidden material, to have *access* to library stacks and to be able to browse through full library collections in order to discover books "whose existence we never suspected, only to discover that they are of extreme importance to us."[638]

---

[633] See *infra* 3.2 and 3.3.
Exactly on this point see Mattern S., Library as infrastructure – Reading room, social service center, innovation lab. How far can we stretch the public library?, Places Journal, June 2014 [online], available at https://placesjournal.org/article/library-as-infrastructure/?gclid=EAIaIQobChMIo5K8_fi41QIVybobCh3lsQ8PEAAYASAAEgLyefD_BwE.

[634] Reference is to Benkler's tri-layered portray of the digital network.
Benkler, Y., 2006. *The wealth of networks: How social production transforms markets and freedom*, supra note 291. Similarly see also Zittrain, J., 2008. The future of the internet--and how to stop it. Yale University Press.

[635] *Infra* at section 5.3.

[636] See *infra*, section 6.4. There we argued that, for libraries, the fact that this 'war' is being fought in copyright terms is as beneficial as it is for the other players involved.

[637] See also: Pessach, G., 2008. [Networked] Memory Institutions: Social Remembering, Privatization and its Discontents. Cardozo Arts & Ent. LJ, 26, p.71.

[638] Höfer, C., Libraries, Schirmer/Mosel (2005) translating Umberto Eco, De Bibliotheca (1981).
In this essay, written on the occasion of the twenty-fifth anniversary of the Milan Public Library, Eco described what constitutes a 'bad library' (la biblioteca cattiva) in twenty-one points most of which present difficulties of different sorts in accessing books or in having the space and time to consult them.

BIBLIOGRAPHY

LEGISLATION (Chronological order)

- June 1643: An Ordinance for the Regulating of Printing, in Acts and Ordinances of the Interregnum, 1642-1660, ed. C H Firth and R S Rait (London, 1911), pp. 184-186. British History Online http://www.british-history.ac.uk/no-series/acts-ordinances-interregnum/pp184-186
- An Act for the Encouragement of Learning, by Securing the Copies of Maps, Charts, and Books to the Authors and Proprietors of Such Copies, During the Times Therein Mentioned., Copyright Act of 1790 online available at https://copyright.gov/about/1790-copyright-act.html
- 17 U.S.C. § 107 - U.S. Code - Title 17. Copyrights
- Berne convention for the protection of literary and artistic works, of September 9, 1886 (as revised, Rome 1928, Brussels 1948, Stockholm, 1967)
- TRIPS: Agreement on Trade-Related Aspects of Intellectual Property Rights, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1C, 1869 U.N.T.S. 299, 33 I.L.M. 1197 (1994)
- Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals With Regard to the Processing of Personal Data and on the Free Movement of Such Data, OJ L 281, 23.11.1995
- WIPO Copyright Treaty, December 20, 1996
- Agreed Statements concerning the WIPO Copyright Treaty, adopted by the Diplomatic Conference on December 20, 1996, online available at http://www.wipo.int/treaties/en/text.jsp?file_id=295456
- Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the Harmonisation of Certain Aspects of Copyright and Related Rights in the Information Society. OJ L 167, 22/06/2001 P. 0010 – 0019
- Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the Re-use of Public Sector Information, OJ L 345, 31.12.2003, p.90 (Directive 2003/98/EC was amended in 2013 by Directive 2013/37/EU, OJ L 175, 27.6.2013, p. 1–8)
- Directive 2006/115/EC of the European Parliament and of the Council of 12 December 2006 on Rental Right and Lending Right and On Certain Rights Related to Copyright In The Field of Intellectual Property (codified version)
- Directive 2012/28/EU of the European Parliament and of the Council of 25 October 2012 on Certain Permitted Uses of Orphan Works, OJ L 299, 27.10.2012
- Marrakesh Treaty to Facilitate Access to Published Works for Persons Who Are Blind, Visually Impaired or Otherwise Print Disabled, WIPO Treaties, June 27, 2013 (adopted), June 30, 2016 (entered into force)

PROPOSED LEGISLATION/ NOTICES OF INQUIRY/ POLICY DOCS (Chronological order)

- Report from the Select Committee on Public Libraries; together with the Proceedings of the Committee, Minutes of Evidence and Appendix, ordered by the House of Commons, to be printed 23, July 1849. Record available online at the HathiTrust catalogue: https://catalog.hathitrust.org/Record/011254579
- Memorandum of Understanding Key Principles on the Digitisation and Making Available of Out-of-Commerce Works, September 20, 2001, online available at http://ec.europa.eu/internal_market/copyright/out-of-commerce/index_en.htm
- Letter to the Presidency of the Council of the European Union and to the European Commission signed by Six Heads of State and Government (Jacques Chirac, Aleksander Kwaśniewski, Gerhard Schröder, Silvio Berlusconi, José Luis Rodríguez Zapatero and Ferenc Gyurcsány), April 28, 2005, available at the Digital Single Market webpage of the European Commission at: https://ec.europa.eu/digital-single-market/news/timeline-digitisation-and-online-accessibility-cultural-heritage
- Response by the Cornell University Library to the Notice of Inquiry Concerning Orphan Works, 70 Fed. Reg. 3739 January 26, 2005, Comment OW0569, available at http://www.copyright.gov/orphan/comments/OW0569-Thomas.pdf
- Orphan Works Act of 2006, H.R. 5439, 109th Congress, 2nd session, 22 May 2006
- Orphan Works Act of 2008, H.R. 5889, 110th Congress, 2nd session, 24 April 2008
- Shawn Bentley Orphan Works Act of 2008, S. 2913, 110th Congress, 2nd session, 24 April 2008
- Crews K., WIPO Study on Copyright Limitations and Exceptions for Libraries and Archives, SCCR/17/2 (August 26, 2008), online available at http://www.wipo.int/edocs/mdocs/copyright/en/sccr_17/sccr_17_2.pdf, at p. 25 et. seq.
- EC Press Release, April 21, 2010. *Boosting cultural heritage online: the European Commission sets up a Reflection Group on digitization*, IP/10/456, online available at http://europa.eu/rapid/press-release_IP-10-456_en.htm
- Report of the 'Comité des Sages' reflection group on bringing Europe's cultural heritage online, The New Renaissance, January 10, 2011, online available at https://ec.europa.eu/digital-single-market/en/news/timeline-digitisation-and-online-accessibility-cultural-heritage
- Vuopala, A., Assessment of the orphan works issue and costs for rights clearance (2010, report for the European Commission), available at http://ec.europa.eu/information_society/activities/digital_libraries/doc/reports_orphan/anna_report.pdf
- *The Cost of Digitising Europe's Cultural Heritage - A Report for the Comité des Sages of the European Commission*, Prepared by Nick Poole, the Collections Trust, November 2010, available at http://nickpoole.org.uk/wp-content/uploads/2011/12/digiti_report.pdf

- Memorandum of Understanding on key principles on the digitization and making available of out-of-commerce works, September 20, 20111, available at http://europa.eu/rapid/press-release_MEMO-11-619_en.htm
- The Case for a Treaty on Exceptions and Limitations for Libraries and Archives: Background Paper by IFLA, ICA, EIFL and INNOVARTE, WIPO Doc. SCCR/23/3 (Nov. 15, 2011)
- Berkeley Digital Library Copyright Project Comments in Response to the U.S. Copyright Office Notice Inquiry on Orphan Works and Mass Digitization (Feb. 4, 2013), available at https://www.law.berkeley.edu/files/Berkeley_Dig_Lib_Copyright_Proj__REPLY_Comments_to_Copyright_Office_OW_NOI_FINAL_with_appendix.pdf
- The International Federation of Library Associations and Institutions, IFLA Principles for Library eLending (revised August 16, 2013), online available at: http://www.ifla.org/files/assets/hq/topics/e-lending/principles-for-library-elending-rev-aug-2013.pdf
- Library of Congress & Copyright Office Notice of Inquiry, Mass Digitization Pilot Program; Request for Comments, Docket No. 2015–3, available at https://www.copyright.gov/fedreg/2015/80fr32614.pdf
- *Cultural heritage, Digitisation, online accessibility and digital preservation,* Report on the Implementation of Commission Recommendation, 2011/711/EU 2013-2015, online available at http://ec.europa.eu/information_society/newsroom/image/document/2016-43/2013-2015_progress_report_18528.pdf
- Crews K., WIPO Study on Copyright Limitations and Exceptions for Libraries and Archives: Updated and Revised, SCCR/30/3 (June 10, 2015), online available at http://www.wipo.int/edocs/mdocs/copyright/en/sccr_30/sccr_30_3.pdf
- US Copyright Office, June 2015, Orphan Works and Mass Digitization: A Report of the Register of Copyrights, available at http://www.copyright.gov/orphan/
- Library of Congress & Copyright Office Notice of Inquiry, Section 108: Draft Revision of the Library and Archives Exceptions in U.S. Copyright Law, Docket No. 2016 – 4, available at https://www.gpo.gov/fdsys/pkg/FR-2016-06-07/pdf/2016-13426.pdf
- Section 108 Study Group report, online available at http://www.section108.gov/docs/Sec108StudyGroupReport.pdf
- Proposal for a Regulation of the European Parliament and of the Council Laying Down Rules on The Exercise of Copyright and Related Rights Applicable to Certain Online Transmissions of Broadcasting Organisations and Retransmissions of Television and Radio Programmes, COM/2016/0594 final - 2016/0284 (COD) (14.9.2016), text available at http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016PC0594
- Proposal for a Directive of the European Parliament and of the Council on Copyright in the Digital Single Market, COM/2016/0593 final - 2016/0280 (COD) (14.9.2016), text available at http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016PC0593
- Proposal for a Directive of the European Parliament and of the Council on Certain Permitted Uses of Works and Other Subject-matter Protected by Copyright and Related Rights for the Benefit of Persons Who are Blind, Visually Impaired or Otherwise Print Disabled and Amending

Directive 2001/29/EC on the Harmonisation of Certain Aspects of Copyright and Related Rights in the Information Society, COM/2016/0596 final - 2016/0278 (COD) (14.9.2016), text available at http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016PC0596

Proposal for a Regulation of the European Parliament and of the Council on the Cross-border Exchange between the Union and Third Countries of Accessible Format Copies of Certain Works and Other Subject-matter Protected by Copyright and Related Rights for the Benefit of Persons Who are Blind, Visually Impaired or Otherwise Print Disabled, COM/2016/0595 final - 2016/0279 (COD) (14.9.2016), text available at http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016PC0595

- 18th Meeting of the Member States Expert Group on Digitisation and Digital Preservation (MSEG), 22-23 November 2016, Event Reports available at https://ec.europa.eu/digital-single-market/en/news/18th-meeting-member-states-expert-group-digitisation-and-digital-preservation-mseg
- The International Federation of Library Associations and Institutions, IFLA Trend Report 2016 - Update available at https://trends.ifla.org/update-2016
- Libraries, Development and the United Nations 2030 Agenda, available at https://www.ifla.org/node/11013
- The International Federation of Library Associations and Institutions (IFLA), 17 April, 2017, Draft guidelines for library services to people experiencing homelessness, online available https://www.ifla.org/files/assets/lsn/projects/homeless/draft-guidelines-homeless-call-for-review.pdf

CASE-LAW (INCL. BRIEFS & *AMICI CURIAE*) (Chronological order)

- *Abrams v. United States*, 250 U.S. 616 (1919)
- *Williams & Wilkins Co. v. United States*, 487 F.2d 1345 (1973)
- *Williams & Wilkins Co. v. United States*, 420 U.S. 376 (1975)
- *Campbell v. Acuff-Rose Music*, 510 U.S. 569 (1994)
- *A&M Records, Inc. v. Napster, Inc.*, 239 F.3d 1004 (2001)
- *Kelly v. Arriba Soft Corp.*, 336 F.3d 811 (9th Cir. 2003)
- *Eldred v. Ashcroft*, 537 US 186 (2003)
- Brief of *Amici Curiae :* George A. Akerlof, Kenneth J. Allow, Timothy F. Bresnahn, James M. Buchanan, Ronald H. Coase, Linda R. Cohen, Milton Friedman, Jerry R. Green, Robert W. Hahn, Thomas W. Hazlett, C. Scott Hemphill, Robert E. Litan, Roher G. Noll, Richard Schmalensee, Steven Shavell, Hal R. Varian, and Richard J. Zeckhauser (May 20, 2002), in support of Petitioners in *Eldred v. Ashcroft*, 537 US 186 (2003)
- *Bill Graham Archives*, 448 F.3d (2006)
- *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1168 (9th Cir. 2007)
- *Infopaq International*, (Case C-5/08)
- *Football Association Premier League and Others*, (Cases C-403/08 and 429/08)
- *SCF Consorzio Fonografici*, (Case C-135/10)
- UsedSoft (Case C-128/11)
- *Svensson and Others*, (Case C-466/12)
- *Authors Guild et al. v. Google Inc*., 770 F. Supp. 2d 666 (S.D.N.Y. 2011) [in the text cited as: *Authors Guild etal. v. Google Inc*. – 2011 Opinion Rejecting the Settlement]
- *Authors Guild Inc. v. Google Inc.* Docket:
    - Publishers' initial complaint, 19 Oct. 2005, available at https://www.authorsguild.org/wp-content/uploads/2008/10/McGraw-Hill-v.-Google-10192005.pdf
    - Authors' Guild initial class action complaint, 20 Sept. 2005, available at https://www.authorsguild.org/wp-content/uploads/2008/10/Authors-Guild-v-Google-09202005.pdf
    - Filing #59, Second Amended Class Action Complaint (October 31, 2008)
- *Authors Guild, Inc. v. HathiTrust, 902 F. Supp. 2d 445, 448 (S.D.N.Y. 2012)* [in the text cited as: *Authors Guild, Inc. v. HathiTrust– 2012 District Court Opinion*]
- *Authors Guild, Inc. v. Google Inc*., 954 F. Supp. 2d 282 (S.D.N.Y. 2013), District Court's Summary Judgment on Google's fair use defense [in the text cited as: *Authors Guild, Inc. v. Google Inc*. – 2013 Summary Judgment]
- *Georgia State University v. Becker*, 863 F. Supp. 2d 1190, 1209 (N.D. Ga. 2012)

- *Technische Universität Darmstadt v. Eugen Ulmer KG*, (Case C-117/13), ECLI:EU:C:2014:2196
- *Capitol Records LLC v. ReDigi Inc.*, case number 1:12-cv-00095, in the U.S. District Court for the Southern District of New York
- *Capitol Records LLC v. ReDigi Inc.*, 934 F.Supp.2d 640 (2013)
- Brief of *Amici Curiae:* American Library Association, Association of College and Research Libraries, Association of Research Libraries, and Internet Archive, *Capitol Records LLC v. ReDigi Inc.*
- *Authors Guild, Inc. v. HathiTrust, 755 F.3d 87 (2d Cir. 2014)* [also cited as: *Authors Guild, Inc. v. HathiTrust– 2014 Second Circuit Opinion*]
- *Cambridge Univ. Press et al v. Patton*, 769 F.2d 1232 (11th Cir. 2014)
- Pending on appeal before the Eleventh Circuit (Case No. 16-15726: *Cambridge University Press, Oxford University Press, Inc., and Sage Publications, Inc. v. J. L. Albert, et al*.)
- *Authors Guild et al. v. Google, Inc*., No. 13-4829 (2d Cir. 2015) [in the text cited as: *Authors Guild, Inc. v. Google Inc*. – 2015 Second Circuit Opinion]
- Ranks/Microsoft, Case C-166/15, ECLI:EU:C:2016:762
- *NUV v. Tom Kabinet*, Court of Appeal of Amsterdam, ECLI:NL:GHAMS:2015:66, (20 January 2015)
- *Vereniging Openbare Bibliotheken v. Stichting Leenrecht*, (Case C-174/15), ECLI:EU:C:2016:856
- *Marc Soulier & Sara Doke v. Premier ministre & Ministre de la Culture et de la Communication*, (Case C-301/15), ECLI:EU:C:2016:878
- Petition before the Supreme Court (No. 15-849): *The Authors Guild, et al., Petitioners v. Google, Inc.* (Docketed: December 31, 2015)
- *Elsevier Inc. et al v. Sci-Hub et al*, No. 1:2015cv04282 - Document 53 (S.D.N.Y. 2015)
- *The Chancellor, Masters & Scholars of the University of Oxford & Ors. v Rameshwari Photocopy Services & Anr.* (CS(OS) 2439/2012) 16 Sep 2016, Delhi High Court
- *The Canadian Copyright Licensing Agency ("Access Copyright") v. York University* (2017 FC 669)

BOOKS (alphabetical order)

- Arms, W.Y., 2000. Digital libraries. MIT Press
- Baldwin, P., 2014. The copyright wars: three centuries of transatlantic battle. Princeton University Press
- Benkler, Y., 2011. The penguin and the leviathan: How cooperation triumphs over self-interest. Crown Business
- Benkler, Y., 2006. The wealth of networks: How social production transforms markets and freedom. Yale University Press
- Bishop, A.P., Van House, N.A. and Buttenfield, B.P., 2003. Digital library use: Social practice in design and evaluation. MIT Press
- Black J., Hashimzade N., and Myles G. A Dictionary of Economics. Oxford University Press (ed. 4)
- Borelli A., Pastore Passaro M., Beebe D., (ed.) 2006, Selected writings of Girolamo Savonarola: religion and politics, 1490-1498, Yale University Press.
- Borges, Jorge Luis.1941. The library of Babel
- Borghi, M. and Karapapa, S., 2013. Copyright and mass digitization. Oxford University Press
- Borgman, C.L., 2000. From Gutenberg to the global information infrastructure: access to information in the networked world. MIT Press
- Boyle, J., 1996. Shamans, software, and spleens. Harvard University Press
- Brown S. and Sibley D. 1986, The theory of public utility pricing. Cambridge University Press
- Calabresi G. 2016. The Future of Law and Economics. Essays in Reform and Recollection. Yale University Press
- Caves, R.E., 2000. Creative industries: Contracts between art and commerce. Harvard University Press
- Chowdhury, G. and Chowdhury, S., 2002. Introduction to digital libraries. Facet Publishing
- Cohen, J.E., 2012. Configuring the networked self: Law, code, and the play of everyday practice. Yale University Press
- Crawford, A. 2015. The meaning of the library: a cultural history, Princeton University Press
- D'Angelo, E., 2006. Barbarians at the gates of the public library: how postmodern consumer capitalism threatens democracy, civil education and the public good. Library Juice Press, LLC
- Deazley, R., 2004. On the origin of the right to copy: Charting the movement of copyright law in eighteenth century Britain (1695-1775). Bloomsbury Publishing
- Derclaye, E. ed., 2010. Copyright and Cultural Heritage: Preservation and access to works in a digital world. Edward Elgar Publishing
- Umberto Eco, De Bibliotheca (1981)

- Ferguson, S.J. and Hebels, R., 2003. Computers for Librarians: An Introduction to the Electronic Library. Elsevier
- Frischmann B. 2012. Infrastructure – The social value of shared resources. Oxford University Press
- Grimmelman J., Patterns of Information Law - Intellectual Property Done Right, Fall 2016 Course pack project [online], available at http://james.grimmelmann.net/ipbook/
- Hansen, D. 2016. Digitizing Orphan Works: Legal Strategies to Reduce Risks for Open Access to Copyrighted Orphan Works. Kyle K. Courtney and Peter Suber, eds., Harvard Library, available at http://nrs.harvard.edu/urn-3:HUL.InstRepos:27840430
- Höfer, C., Libraries, Schirmer/Mosel (2005) translating Umberto Eco, De Bibliotheca (1981)
- Karaganis J. ed. (2018). *Shadow Libraries: Access to Educational Materials in the Digital Age*. MIT Press. [forthcoming]
- Levine, D. and Boldrin, M., 2008. Against intellectual monopoly. Cambridge University Press
- Levine M., Adler R.C., Bonfiglio J., Eden K., and Hall B.S., Finding the Public Domain: Copyright Review Management System Toolkit, DOI: http://dx.doi.org/10.3998/crmstoolkit.14616082.0001.001
- Locke J. 1692, Letter to Edward Clarke, *in* Benjamin Rand, ed. The correspondence of John Locke and Edward Clarke 366, Harvard University Press (1927)
- Locke, J. [1689] *The Two Treatises of Civil Government* (Hollis ed.)
- Milton J. 1644, Areopagitica (Jebb ed.)
- Morozov, E., 2014. To save everything, click here: The folly of technological solutionism. Public Affairs
- Murison, W.J., 1988. The public library: its origins, purpose and significance. London: C. Bingley
- Palfrey, J.G. and Gasser, U., 2013. Born digital: Understanding the first generation of digital natives. Basic Books.
- Palfrey, J., 2015. BiblioTech: Why libraries matter more than ever in the age of Google. Basic Books
- Patry, W., 2009. Moral panics and the copyright wars. Oxford University Press
- Parker H., 1643, To the High Court of Parliament: The humble Remonstrance of the Company of Stationers, London.:Stationers' Company (London, England)
- Pawley, C., 2010. Reading places: Literacy, democracy, and the public library in Cold War America. Univ. of Massachusetts Press
- Perzanowski A. and Schultz J., 2016, The End of Ownership: Personal Property in the Digital Economy, MIT Press
- Pigou, A.C., 2013. The economics of welfare. Palgrave Macmillan
- Posner, R.A. and Landes, W.M., 2003. The economic structure of intellectual property law. Harvard University Press

- Reagle, J.M., 2010. Good faith collaboration: The culture of Wikipedia. MIT Press
- Rose, M., 1993. Authors and owners: The invention of copyright. Harvard University Press
- Sherman, B. and Bently, L., 2003. The Making of Modern Intellectual Property Law: the British Experience, 1760 - 1911. Cambridge University Press
- Smith A., Lectures On Jurisprudence, ed. R. L. Meek, D. D. Raphael and P.G. Stein, vol. V of the Glasgow Edition of the Works and Correspondence of Adam Smith (Indianapolis: Liberty Fund, 1982)
- Towse, R. 2001. Creativity, Incentive and Reward: An Economic Analysis of Copyright and Culture in the Information Age, Edward Elgar Publishing
- Trevelyan H., (1875), The works of Lord Macaulay Complete (Vol. 8), Bombay, Longmans, Green, and co.
- Waldron J. 1988. The right to private property, Clarendon Press
- Wells H.G. 2016 (reprint). World Brain. Read Books Ltd.
- Wright A. 2014. Cataloging the World. Oxford University Press
- Young, G.M. ed., 1967. Macaulay: Prose and poetry. Harvard University Press
- Zittrain, J., 2008. The Future of the Internet and How To Stop It. Yale University Press

Articles/ Journals (alphabetical order)

- Afori, O.F., 2013. The Battle Over Public E-Libraries–Taking Stock and Moving Ahead. *IIC-International Review of Intellectual Property and Competition Law*, *44*(4), pp.392-417
- Baldonado, M., Chang, C.C.K., Gravano, L. and Paepcke, A., 1997. The Stanford digital library metadata architecture. *International Journal on Digital Libraries*, *1*(2), pp.108-121
- Band, J., 2009. The long and winding road to the Google Books Settlement. *J. Marshall Rev. Intell. Prop. L.*, *9*, p.i
- Becker, L.C., 1992. Deserving to own intellectual property. *Chi.-Kent L. Rev.*, *68*, p.609
- Black, A., 2000. Skeleton in the cupboard: social class and the public library in Britain through 150 years. *Library History*, *16*(1), pp.3-12
- Benkler, Y., 1998, September. The commons as a neglected factor of information policy. In *26th Annual Telecommunications Research Conference* (Oct. 3-5, 1998), online available at http://www.benkler.org/commons.pdf
- Benkler Y. and Clark D. 2016, The Internet: Introduction, *Daedalus* 145(1): 5-8
- Bodó, B., 2011. Coda: A short history of book piracy. *Media piracy in emerging economies*, p.399
- Bodó, B., 2014. In the Shadow of the Gigapedia-The Analysis of Supply and Demand for the Biggest Pirate Library on Earth, online available at: https://ssrn.com/abstract=2616633 [Forthcoming as The shadow of Gigapedia: quantitative analysis of shadow library usage. in: Karaganis (ed): *Shadow libraries*]
- Bodó, B., 2015. Libraries in the post-scarcity era. *in* H. Porsdam (ed.), Copyrighting creativity: creative values, cultural heritage institutions and systems of intellectual property Farnham: Ashgate, pp. 75-92
- Boldrin, M. and Levine, D., 2002. The Case against Intellectual Property. *The American Economic Review*, *92*(2), pp.209-212
- Boyle, J., 2003. The second enclosure movement and the construction of the public domain. *Law and contemporary problems*, *66*(1/2), pp.33-74
- Boyle, J., 2003. Foreword: The opposite of property?. *Law and contemporary problems*, pp.1-32, at p. 9
- Boyle J., (When) Is Copyright Reform Possible? in Okediji, R.L. ed., 2017. *Copyright law in an age of limitations and exceptions*. Cambridge University Press
- Brennan, T.J., 1992. Copyright, property, and the right to deny. *Chi. Kent L. Rev.*, *68*, p.675
- Breyer, S., 1970. The uneasy case for copyright: A study of copyright in books, photocopies, and computer programs. *Harvard Law Review*, pp.281-351
- Brin, S. and Page, L., 1998. The anatomy of a large-scale hypertextual Web search engine. *Computer networks and ISDN systems*, *30*(1-7), pp.107-117

- Calabresi, G. and Babbitt, P., 1978. *Tragic Choices: The Conflicts Society Confronts in the Allocation of Tragically Scare Resources*. W.W.Norton & Company
- Carstens, A.M., 2015. Technische Universität Darmstadt v. Eugen Ulmer KG. *American Journal of International Law*, *109*(1), pp.161-167
- Carter, S.L., 1992. Does it Matter Whether Intellectual Property is Property. *Chi.-Kent L. Rev.*, *68*, p.715
- Chiarizio, M., 2013. An american tragedy: e-books, licenses, and the end of public lending libraries. *Vand. L. Rev.*, *66*, p.615
- Child, J.W., 1990. The moral foundations of intangible property. *The Monist*, *73*(4), pp.578-600
- Christenson, H., 2016. Building a US Federal Government Documents Collection in HathiTrust. *Collaborative Librarianship*, *8*(3), p.5
- Coase, R.H., 1937. The nature of the firm. *economica*, *4*(16), pp.386-405
- Coase, R.H., 1959. The Federal Communications Commission. *Journal of Law and Economics*, pp.1-40
- Cohen J., Copyright, Commodification and Culture: Locating the Public Domain, in Hugenholtz, P.B. and Guibault L. (eds), 2006. *The future of the public domain: identifying the commons in information law*. Kluwer Law International, pp. 121-166
- Cohen, J., 2011. Copyright as Property in the Post-Industrial Economy: A Research Agenda. *Wis. L. Rev.*, p.141
- Constant, S.A., 2013. The Computer Fraud and Abuse Act: A Prosecutor's Dream and a Hacker's Worst Nightmare-The Case Against Aaron Swartz and the Need To Reform the CFAA. *Tul. J. Tech. & Intell. Prop.*, *16*, p.231
- David P. 2004. The End of Copyright History? *Review of Economic Research on Copyright Issues*, 1(2), pp. 5-10
- Davies, P., 2013. Access v contract: competing freedoms in the context of copyright limitations and exceptions for libraries. *EIPR*, *35*(7), pp.402-414
- Dusollier S., 2014, A manifesto for an e-lending limitation in copyright. *Journal of Intellectual Property, Information Technology and E-Commerce Law* 3
- Easterbrook, F.H., 1990. Intellectual property is still property. *Harv. J. L. & Pub. Pol'y*, *13*, p.108
- Elkin-Koren, N., 1996. Cyberlaw and social change: A democratic approach to copyright law in cyberspace. *Cardozo Arts & Ent. LJ*, *14*, p.215
- Elkin-Koren N., Copyright in a Digital Ecosystem: A User Rights Approach, *in* Okediji, R.L. ed., 2017. Copyright law in an age of limitations and exceptions. Cambridge University Press
- Feather, J., 1980. The book trade in politics: The making of the copyright act of 1710. *Publishing history*, *8*, p.19
- Floridi, L. (1995) The Internet: Which Future for Organized Knowledge—Frankenstein or Pygmalion? International Journal of Human-Computer Studies 43: 261-274

- Fisher, W., 2001. Theories of intellectual property. *New essays in the legal and political theory of property*, *168*(1)
- Fox, E.A. and Marchionini, G., 1998. Toward a worldwide digital library. *Communications of the ACM*, 41(4), pp.29-32
- Garrett, J., 1991. Missing Eco: On Reading "The Name of the Rose" as Library Criticism. *The Library Quarterly*, *61*(4), pp.373-388
- Geddes R., Public Utilities, in *Encyclopedia of Law and Economics* 1161, 1165-67 (Boudewijn Bouckaert & Gerrit De Gees eds., 2000)
- Ginsburg C.J., 2015, Google Books and Fair Use: From Implausible to Inevitable? *The Media Institute*, October 19, 2015 [online], available at http://www.mediainstitute.org/IPI/2015/101915.php
- Gordon, W.J., 1982. Fair use as market failure: a structural and economic analysis of the" Betamax" case and its predecessors. *Columbia Law Review*, *82*(8), pp.1600-1657
- Gould, R., 2013. Aaron Swartz's Legacy. *Academe: Magazine of the American Association of University Professors*, Vol. 100, No. 1, 2014, online available at https://ssrn.com/abstract=2207299
- Greenstein S., Internet Infrastructure, in *The Oxford Handbook of Digital Economy* (Martin Peitz and Joel Waldfogel eds.)
- Grimmelmann J. 2012. The Orphan Wars, *EDUCAUSE Review*, 47(1) [online], available at http://www.educause.edu/ero/article/orphan-wars
- Hahn, T.B., 2011. Mass digitization. *Library Resources & Technical Services*, 52(1), pp.18-26
- Hall E., Adventures in ancient Greek and Roman Libraries. in Alice Crawford (ed.) 2015. *The meaning of the library: a cultural history*, Princeton University Press
- Hanratty, E., 2005. Google Library: Beyond Fair Use?. *Duke L. & Tech. Rev.*, *2005*, pp.10-26
- Hansen D. 2011. Orphan Works: Definitional Issues, Berkeley Digital Library Copyright Project, White Paper No.1 [online], available at http://ssrn.com/abstract=1974614
- Hansen D. 2012. Orphan works: mapping the possible solution spaces (Berkeley Digital Library Project, White Paper No 2 [online], available at http://ssrn.com/abstract=2019121
- Hansen D. 2012. Orphan works: causes of the problem, Berkeley Digital Library Copyright Project, White Paper No 3 [online], available at http://ssrn.com/abstract=2038068
- Hansen D., Hinze G., and Urban J. 2013. Orphan Works and the Search for Rightsholders: Who Participates in a 'Diligent Search' Under Present and Proposed Regimes, Berkeley Digital Library Copyright Project, White Paper No. 4 [online] available at http://ssrn.com/abstract=2208163
- Hansen D., Hinze G., and Urban J. 2013 What Constitutes a Diligent Search Under Present and Proposed Orphan Works Regimes, Berkeley Digital Library Copyright Project, White Paper No. 5, 2013 [online] available at http://ssrn.com/abstract=2229021
- Hansen D., Hashimoto K., Hinze G., Samuelson P., and Urban J. 2013. Solving the Orphan Works Problem for the United States, 37 *Columbia J. L & Arts* 1

- Heald, P. 2013. The Public Domain. Illinois Public Law Research Paper No. 14-21. available at *SSRN*: https://ssrn.com/abstract=2362983
- Heald P. 2013, How Copyright Keeps Works Disappeared, *Illinois Public Law*, Research Paper No. 13-54
- Heald P., 2014. The Demand for Out-of-Print Works and Their (Un)Availability in Alternative Markets, *Illinois Public Law*, Research Paper No. 14-31
- Heller, M.A., 1999. The boundaries of private property. *The Yale Law Journal*, *108*(6), pp.1163-1223
- Hohfeld, W.N., 1913. Some fundamental legal conceptions as applied in judicial reasoning. *The Yale Law Journal*, *23*(1), pp.16-59
- Holt, L.E. and Holt, G.E., 2010. Public library services for the poor: Doing all we can. Publication of the *American Library Association*
- Huang, O., 2006. US copyright office orphan works inquiry: Finding homes for the orphans. *Berkeley Technology Law Journal*, pp.265-288, at p. 268.
- Hughes, J., 2009. Locke's 1694 Memorandum (and More Incomplete Copyright Historiographies). *Cardozo Arts & Ent. LJ*, *27*, p.555
- IVIR project *Copyright in an Age of Access: Alternatives to Copyright Enforcement feasibility of legalizing file-sharing* – publications available at https://www.ivir.nl/projects/copyright-in-an-age-of-access-alternatives-to-copyright-enforcement/
- Jenkins, J., 2014. Last sale? Libraries' rights in the digital age. *College & Research Libraries News*, *75*(2), pp.69-75
- Kahle, B. *Ethics of Digital Librarianship*, Thinking Machines (now brewster@archive.org), online available at https://archive.org/about/ethics_BK.php
- Kahn R. and Cerf V., 1999. What is the Internet (and what makes it work), paper online available at http://www.policyscience.net/cerf.pdf
- Karaganis J. 2017, Digital Piracy. *in* Handbook of Digital Media and Social Practice. Springer, pp. 1-9
- Katz, A., 2012. The Orphans, the Market, and the Copyright Dogma: a modest solution for a grand problem. *Berkeley Technology Law Journal*, *27*(3), pp.1285-1346
- Landes, W.M. and Posner, R.A., 1989. An economic analysis of copyright law. *The Journal of Legal Studies*, *18*(2), pp.325-363
- Leiner, B.M., Cerf, V.G., Clark, D.D., Kahn, R.E., Kleinrock, L., Lynch, D.C., Postel, J., Roberts, L.G. and Wolff, S., 2000. A brief history of the Internet, online available at the Internet Society webpage: http://www.internetsociety.org/internet/what-internet/history-internet/brief-history-internet#Timeline
- Lemley, M.A. 1996, Romantic Authorship and the Rhetoric of Property. *Tex. L. Rev.*, *75*, pp.873-887
- Leval, P.N., 1990. Toward a fair use standard. *Harvard Law Review*, *103*(5), pp.1105-1136

- Liebowitz, S.J., 1985. Copying and indirect appropriability: Photocopying of journals. *Journal of political economy*, *93*(5), pp.945-957
- Liebowitz, S.J., 2004. Will MP3 downloads annihilate the record industry? The evidence so far. *in Intellectual property and Entrepreneurship* (pp. 229-260). Emerald Group Publishing Limited
- Liebowitz, S.J. and Margolis, S.E., 2005. Seventeen famous economists weigh in on copyright: The role of theory, empirics, and network effects. *Harv. J. of L. & Tech.*18(2)
- Linklater, E., 2015. Make me an offer I won't regret: Offers to license works on acceptable terms cannot block libraries' "right" to digitize for access on dedicated terminals: Technische Universität Darmstadt. *Common Market Law Review*, *52*(3), pp.813-823
- Lipsey, R.G., 2007. Technological transformation, intellectual property rights and second best theory. *Review of Economic Research on Copyright Issues*, 4(2), pp. 5-28
- Litman. 2011. J., Readers' Copyright. *Journal of the Copyright Society of the USA*, Vol. 58, p. 325
- Loewenheim, U., 2015. Boundaries of use of electronic terminals. *Journal of Intellectual Property Law & Practice*, *10*(5), pp.384-387
- Loren, L.P., 2012. Abandoning the Orphans: An Open Access Approach to Hostage Works. *Berkeley Technology Law Journal*, pp.1431-1470
- Luey, B., 1994. Priorities for publishing research: a multidisciplinary agenda. *Publishing Research Quarterly*, *10*(3), pp. 31-35
- Mattern S., Library as infrastructure – Reading room, social service center, innovation lab. How far can we stretch the public library?, *Places Journal*, June 2014 [online], available at https://placesjournal.org/article/library-as-infrastructure/?gclid=EAIaIQobChMIo5K8_fi41QIVybobCh3lsQ8PEAAYASAAEgLyefD_BwE
- Matulionyte, R., 2015. E-Lending and a Public Lending Right: Is it Really a Time for an Update? *EIPR* 38(3), p.132
- Matulionyte, R., 2016. 10 years for Google Books and Europeana: copyright law lessons that the EU could learn from the USA. *International Journal of Law and Information Technology*, p.18
- Max, S.M., 1984. Tory Reaction to the Public Libraries Bill, 1850. *The Journal of Library History (1974-1987)*, *19*(4), pp.504-524
- McCook, K.D.L.P., 2001. Poverty, democracy and public libraries. *Libraries &Democracy: The Cornerstones of Liberty*, pp.28-46
- McNutt M. My love-hate of Sci-Hub, *Science* 29 Apr 2016: Vol. 352, Issue 6285, pp. 497
- Menell P., 2007. Intellectual Property and the Property Rights Movement, 30 *Regulation* 36
- Merrill, T.W. and Smith, H.E., 2001. What happened to property in law and economics?. *The Yale Law Journal*, *111*(2), pp.357-398
- Müller, H., 2012. Legal aspects of e-books and interlibrary loan. *Interlending & Document Supply*, *40*(3), pp.150-155

- Na, N., 2006. Testing the Boundaries of Copyright Protection: The Google Books Library Project and the Fair Use Doctrine. *Cornell JL & Pub. Pol'y*, *16*, p.417
- Nagaraj, A., 2016. Does Copyright Affect Reuse? Evidence from the Google Books Digitization Project, [online] Available at *SSRN*: https://ssrn.com/abstract=2810761
- Nimmer, M.B., 1968. Project New Technology and the Law of Copyright: Reprography and Computers; Foreword: Two Copyright Crises. 5 *UCLA L. Rev.* 939
- Novos, I.E. and Waldman, M., 1984. The effects of increased copyright protection: An analytic approach. *Journal of political economy*, *92*(2), pp.236-246
- Oberholzer-Gee, F. and Strumpf, K., 2007. The effect of file sharing on record sales: An empirical analysis. *Journal of political economy*, *115*(1), pp.1-42
- O'Brien, D., Gasser, U. and Palfrey, J.G., 2012. E-books in libraries: A briefing document developed in preparation for a workshop on e-lending in libraries. Berkman Center Research Publication No. 2012-15, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2111396##
- Ostrom, E., 2010. Beyond markets and states: polycentric governance of complex economic systems. *Transnational Corporations Review*, 2(2), pp.1-12
- Paepcke, A., Baldonado, M., Chang, C.C.K., Cousins, S. and Garcia-Molina, H., 2000. *Building the Infobus: A review of technical choices in the Stanford digital library project*. Stanford InfoLab
- Palfrey J. 2016, Design Choices for Libraries in the Digital-Plus Era, *Daedalus* 145(1): 79–86
- Peitz, M. and Waelbroeck, P., 2003. Piracy of digital products: A critical review of the economics literature, CESifo Working Paper Series No. 1071. available at *SSRN*: https://ssrn.com/abstract=466063
- Penner, J.E., 1995. The "Bundle of Rights" Picture of Property, 43 *UCLA L. Rev.* 711, at pp. 733-38
- Perzanowski, A. and Schultz, J., 2010. Digital exhaustion. *UCLA L. Rev.*, *58*, p.889
- Perzonowski, A. and Schultz, J., 2014. Legislating digital exhaustion. *Berkeley Tech. LJ*, *29*, p.1535
- Pettegree, A. The Renaissance Library and the Challenge of Print, in Alice Crawford (ed.) 2015. *The meaning of the library: a cultural history*, Princeton University Press
- Picker, R.C., 2009. The Google Book Search Settlement: A New Orphan-Works Monopoly?. Journal of Competition Law and Economics, 5(3), pp.383-409
- Plant, A., 1934. The economic aspects of copyright in books. *Economica*, *1*(2), pp.167-195
- Pollock, Stepan, & Valimakki, Valuing the EU Public Domain, Cambridge Working Papers in Economics 1047 (2010), 1 – 54, p. 4
- Proskine, E.A., 2006. Google's Technicolor Dreamcoat: A Copyright Analysis of the Google Book Search Library Project. *Berkeley Technology Law Journal*, pp.213-239
- R. van der Noll, K. Breemen, V. Breemen, B. Hugenholtz, M. Brom & J. Poort, Online uitlenen van e-books door bibliotheken. Verkenning juridische mogelijkheden en economische effecten,

in opdracht van het Ministerie van Onderwijs, Cultuur en Wetenschap, Amsterdam, 2012, online available at http://dare.uva.nl/search?identifier=743e4ef8-1ae8-45a7-becd-d123354f8ec0

- Reese, R.A., 2002. First Sale Doctrine in the Era of Digital Networks, The. *BCL Rev.*, *44*, p.577
- Rob, R. and Waldfogel, J., 2006. Piracy on the high C's: Music downloading, sales displacement, and social welfare in a sample of college students. *The Journal of Law and Economics*, *49*(1), pp.29-62
- Rogers, D.L., 2007. Increasing Access to Knowledge Through Fair Use-Analyzing the Google Litigation to Unleash Developing Countries. *Tul. J. Tech. & Intell. Prop.*, *10*, p.1
- Rosati E. 2013. Copyright issues facing early stages of digitization projects, Mobile Collections Project [online], available at http://www.digitalhumanities.cam.ac.uk/Copyrightissuesfacingearlystagesofdigitizationprojects.pdf
- Rosati, E. 2015. Online copyright exhaustion in a post-All posters world. *Journal of Intellectual Property Law & Practice*, p.122
- Rose, M., 2009. The Public Sphere and the Emergence of Copyright: Areopagitica, the Stationers' Company, and the Statute of Anne. *Tul. J. Tech. & Intell. Prop.*, *12*, p.123 (also in Deazley, R., Kretschmer, M. and Bently, L., 2010. *Privilege and property: essays on the history of copyright*. Open Book Publishers.)
- Salzberger, E. 2006. Economic Analysis of the Public Domain, *in* Hugenholtz, P.B. and Guibault L. (eds), *The future of the public domain: identifying the commons in information law*. Kluwer Law International, p. 27-57
- Samuelson, P., 1988. Information as property: Do Ruckelshaus and Carpenter signal a changing direction in intellectual property law, *Cath. U. L Rev.*, *38*, p.365
- Samuelson, P. and Giushko, R.J., 1993. Intellectual property rights for digital library and hypertext publishing systems. *Harv. J. L & Tech.*, 6, p.237
- Samuelson, P., 1993. Fair use for computer programs and other copyrightable works in digital form: The implications of Sony, Galoob and Sega. *J. Intell. Prop. L.*, *1*, p.49
- Samuelson, P., 1995. Copyright and digital libraries. *Communications of the ACM*, *38*(4), pp.15-ff
- Samuelson P., Challenges in Mapping the Public domain, in Hugenholtz, P.B. and Guibault L. (eds), 2006. *The future of the public domain: identifying the commons in information law*. Kluwer Law International, pp.7-21
- Samuelson, P., 2009. Google Book Search and the future of Books in Cyberspace. *Minn. L. Rev.*, *94*, p.1308
- Samuelson, P., 2010. Legislative Alternatives to the Google Book Settlement. *Columbia Journal of Law & the Arts*, *34*, pp.697-705
- Samuelson, P., 2015. Possible futures of fair use. *Wash. L. Rev.*, *90*, p.815
- Samuelson P., Justifications for Copyright Limitations and Exceptions, in Okediji, R.L. ed., 2017. *Copyright law in an age of limitations and exceptions*. Cambridge University Press.

- Shavell, S. and Van Ypersele, T., 2001. Rewards versus intellectual property rights. *The Journal of Law and Economics*, *44*(2), pp.525-547
- Shah, M., 2006. Fair Use and the Google Book Search Project: The Case for Creating Digital Libraries. *CommLaw Conspectus*, *15*, p.569
- Shiffrin, S.V., 2001. Lockean arguments for private intellectual property. *New essays in the legal and political theory of property*, *1*, pp.138-167
- Star, S.L. and Ruhleder, K., 1996. Steps toward an ecology of infrastructure: Design and access for large information spaces. *Information systems research*, *7*(1), pp.111-134
- Suber, P. 2011. Open Access for Digitization Projects. In Going Digital: Evolutionary and Revolutionary Aspects of Digitization, ed. Karl Grandin, 70-93. Stockholm: Nobel Foundation, available at https://dash.harvard.edu/bitstream/handle/1/14976387/KVA_Going_Digital_webb%202.pdf?sequence=1
- Towse, R., Handke, C. and Stepan, P., 2008. The economics of copyright law: A stocktake of the literature. *Review of Economic Research on Copyright Issues*, 5(1), pp. 1-22, at p.3
- Travis, H., 2005. Building universal digital libraries: An agenda for copyright reform. *Pepp. L. Rev.*, *33*, p.761
- Travis, H., 2006. Google Book Search and Fair Use: iTunes for Authors, or Napster for Books. *U. Miami L. Rev.*, *61*, p.87
- Urban J.M. 2012. How fair use can help solve the orphan works problem, 27 *Berkeley Tech. L.J.*
- van Houweling, M.S., 2013. Land Recording and Copyright Reform. *Berkeley Technology Law Journal*, *28*(3), pp.1497-1510
- Varian, H.R., 2000. Buying, sharing and renting information goods. *The Journal of Industrial Economics*, *48*(4), pp.473-488
- Varian, H.R., 2005. Copying and copyright. *The Journal of Economic Perspectives*, 19(2), pp.121-138
- von Rimscha, M.B. and Putzig, S., 2013. From Book Culture to Amazon Consumerism: Does the Digitization of the Book Industry Lead to Commercialization?. *Publishing research quarterly*, *29*(4), pp. 318-335
- Wilhelm, T.E., 2006. Google Book Search: Fair Use or Fairly Useful Infringement. *Rutgers Computer & Tech. LJ*, *33*, p.107
- Wilkin J., 2011. Bibliographic Indeterminacy and the Scale of Problems and Opportunities of "Rights" in Digital Collection Building, online available at http://www.clir.org/pubs/ruminations/01wilkin
- Wilkin J., Meanings of the library today, in Alice Crawford (ed.) 2015. *The meaning of the library: a cultural history*, Princeton University Press
- Winter, M.F., 1994. Umberto Eco on Libraries: A Discussion of" De Bibliotheca". *The Library Quarterly*, *64*(2), pp.117-129

- Wu M., 2016. Collaborative Academic Library Digital Collections Post-Cambridge University Press, HathiTrust and Google Decisions on Fair Use, 1 *Journal of Copyright in Education and Librarianship*
- York J., A Preservation Infrastructure Built to Last – Preservation, Community, and HathiTrust, *in Proceedings of the Memory of the World in the Digital Age: Digitization and Preservation*. An international conference on permanent access to digital documentary heritage, 26-28 September 2012, Vancouver, British Columbia, Canada, edited by Luciana Duranti and Elizabeth Shaffer (UNESCO 2013)
- Yu, P.K., 2003. The escalating copyright wars. *Hofstra L. Rev.*, *32*, p.907, concentrating on the battles against online piracy
- Zentner, A., 2005. File-sharing and international sales of copyrighted music: An empirical analysis with a panel of countries. *Topics in Economic Analysis & Policy*, *5*(1), p.21

Press/ Blogosphere/ Web content (alphabetical order)

- Alter A. and Wingfield N., A Trip Through Amazon's First Physical Store, *The New York Times*, March 10, 2016, online available at http://www.nytimes.com/2016/03/12/business/media/a-virtual-trip-through-amazons-physical-store.html?_r=0. [Amazon's announcement at: https://www.amazon.com/b?node=13270229011]
- Amazon press release, *Amazon.com Now Selling More Kindle Books Than Print Books* (May 19, 2011) is available at: http://phx.corporate-ir.net/phoenix.zhtml?c=176060&p=irol-newsArticle&ID=1565581&highlight
- Ashmore, B., The Race to the Shelf Continues: The Open Content Alliance and Amazon.com, 16(1) *Searcher*, January 2008, online available at http://www.infotoday.com/searcher/jan08/Ashmore_Grogg.shtml
- Becker G., Traditional Bookstores and Libraries are Doomed by the Internet, and so too are Many Other Ways to Distribute Content to Consumers, 03/04/2012 blog-post available at: http://www.becker-posnerblog.com/2012/03/traditional-bookstores-and-libraries-are-doomed-by-the-internet-and-so-too-are-many-otherways-to-distribute-content-to-cons.html
- Bellan A., Copyright exhaustion does not apply to digital goods other than software, Hamm Court of Appeal says, *IP Kat Blog*, June 13, 2014, available at http://ipkitten.blogspot.gr/2014/06/breaking-copyright-exhaustion-does-not.html
- Benson, H., A man's vision: world library online - Brewster Kahle hopes to realize his 25-year dream of an international book archive, *San Francisco Gate*, November 22, 2005 online available at http://www.sfgate.com/news/article/A-MAN-S-VISION-WORLD-LIBRARY-ONLINE-Brewster-2593527.php
- Boston Public Library, Access Policy, online available at: http://www.bpl.org/general/policies/accesspolicy.htm and Service Rules, online available at: http://www.bpl.org/general/circulation/borrowers.htm
- Citizens Defending Libraries Blog-post, Love Brooklyn Libraries Files Complaint With Attorney General Eric Schneiderman That Brooklyn Public Library Understates Capital Funds In Order to Sell Heights Library (Sunday, January 31, 2016), online available at http://citizensdefendinglibraries.blogspot.com/2016/01/love-brooklyn-libraries-files-complaint.html
- Cram G., How we expand access to our public domain, *Spotlight to the Public Domain* (NYPL blog-posts), 17 February 2016 [online], available at https://www.nypl.org/blog/2016/02/17/rights-public-domain
- Darnton R., The Library in the New Age, *New York Review of Books*, June 12, 2008, online available at http://www.nybooks.com/articles/2008/06/12/the-library-in-the-new-age/
- Darnton, R., A library without walls, *The New York Review of Books*, October 4, 2010, online available at http://www.nybooks.com/daily/2010/10/04/library-without-walls/
- Darnton, R., Can we create a National Digital Library?, *The New York Review of Books*, October 28, 2010, online available at http://www.nybooks.com/articles/2010/10/28/can-we-create-national-digital-library/

- Darnton, R., Jefferson's Taper: A National Digital Library, *The New York Review of Books*, November 24, 2011, online available at http://www.nybooks.com/articles/2011/11/24/jeffersons-taper-national-digital-library/
- Darnton, R., The National Digital Public Library Is Launched!, *The New York Review of Books*, April 25, 2013, online available at http://www.nybooks.com/articles/2013/04/25/national-digital-public-library-launched/
- Darnton, R., A World Digital Library is Coming True!, *The New York Review of Books*, May 22, 2014, online available at http://www.nybooks.com/articles/2014/05/22/world-digital-library-coming-true/
- Darnton, R. Great New Possibilities for the Library of Congress, *The New York Review of Books*, August 13, 2015, online available at http://www.nybooks.com/articles/2015/08/13/great-new-possibilities-library-congress/
- Digital Public Library of America Inaugural 2013 Fest (Boston MA, October 24-25) and 2016 Fest (Washington DC, April 14-15), details and material available at https://dp.la/info/get-involved/dplafest/
- Digital Public Library of America Strategic Plan 2015 through 2017 available at http://dp.la/info/wp-content/uploads/2015/01/DPLA-StrategicPlan_2015-2017-Jan7.pdf
- Dunlap D.W., A New Queens Building That Can't Be Overlooked: A Library, *The New York Times*, July 20, 2016 [online], available at: https://www.nytimes.com/2016/07/21/nyregion/a-new-queens-building-that-cant-be-overlooked-a-library.html
- Eckersley P. 2013. Farewell to Aaron Swartz, an Extraordinary Hacker and Activist, Electronic Frontier Foundation post, online available at https://www.eff.org/deeplinks/2013/01/farewell-aaron-swartz
- *EnDOW* ("Enhancing access to 20th Century cultural heritage through Distributed Orphan Works clearance") online resources available at http://diligentsearch.eu/resources/
- Flood A., Ebook sales falling for the first time, finds new report, *The Guardian*, February 3, 2016, online available at https://www.theguardian.com/books/2016/feb/03/ebook-sales-falling-for-the-first-time-finds-new-report
- Ghiasi M., Books in the digital revolution, January 1, 2011 blog-post, available at http://ghiasi.org/2011/01/books-at-the-digital-revolution/
- Gillmor, D., The Creator of the Internet Archive Should Be the Next Librarian of Congress, *Slate*, September 10, 2015, online available at http://www.slate.com/blogs/future_tense/2015/09/10/brewster_kahle_creator_of_the_internet_archive_should_be_the_next_librarian.html
- Google's December 2004 announcement of the library project: "Google checks out library books," December 14, 2004, available at http://googlepress.blogspot.gr/2004/12/google-checks-out-library-books.html
- *Google and The World Brain*, 2013, Documentary directed by Ben Lewis, at http://www.worldbrainthefilm.com/

- Google Academics Inc., Google Transparency Project, 14 July 2017, available at http://googletransparencyproject.org/sites/default/files/Google-Academics-Inc.pdf
- Hafner, K., In challenge to Google, Yahoo will scan books, *The New York Times*, October 3, 2015, online available at http://www.nytimes.com/2005/10/03/business/in-challenge-to-google-yahoo-will-scan-books.html
- Hart, M., 1992. The History and Philosophy of Project Gutenberg by Michael Hart, online available at https://www.gutenberg.org/wiki/Gutenberg:The_History_and_Philosophy_of_Project_Gutenberg_by_Michael_Hart
- Hoffelder, N., Milestone: Project Gutenberg Releases eBook #50,000, October 5, 2015 post available at http://the-digital-reader.com/2015/10/05/milestone-project-gutenberg-releases-ebook-50000/
- Ingram M., No, e-book sales are not falling, despite what publishers say, *Fortune*, September 24, 2015, online available at http://fortune.com/2015/09/24/ebook-sales/
- Jean-Noël Jeanneney, Quand Google défie l'Europe*, Le Monde*, January 22, 2005, online available at http://www.lemonde.fr/archives/article/2005/01/22/quand-google-defie-l-europe-par-jean-noel-jeanneney_395266_1819218.html?xtmc=jeanneney&xtcr=38#
- JISC, In from the Cold - An assessment of the scope of 'Orphan Works' and its impact on the delivery of services to the public, April 2009, Report online available at: http://www.jisc.ac.uk/media/documents/publications/infromthecoldv1.pdf
- Kahle B., Announcing the Open Content Alliance, Yahoo Search blog-post, October 2, 2005, available at http://www.ysearchblog.com/2005/10/02/announcing-the-open-content-alliance/
- Kahle, B., How Google Threatens Books, *The Washington Post*, May 19, 2009, online available at: http://www.washingtonpost.com/wp-dyn/content/article/2009/05/18/AR2009051802637.html
- Keller, P., Advocate General Wathelet: Extended Collective Licensing is NOT the answer for mass digitisation!, July 14, 2016 blogpost, available at http://www.communia-association.org/2016/07/14/advocate-general-wathelet-extended-collective-licensing-not-answer-mass-digitisation/
- Kelly K., Scan this Book!, *The New York Times*, May 14, 2006 [online], available at http://www.nytimes.com/2006/05/14/magazine/14publishing.html?_r=0
- Kimball S., Free for All: NYPL Enhances Public Domain Collections For Sharing and Reuse, *NYPL Labs*, *Spotlight to the Public Domain* (NYPL blog-posts), 5 January 2016 [online available], at https://www.nypl.org/blog/2016/01/05/share-public-domain-collections
- La Biblioteca Aperta - Tecniche e Strategie di Conivisione Convegno Biblioteche 16-17 Marzo 2017
- La Stampa – Technologia, Nasce la Digital Library d'Italia: online 101 archivi e 46 biblioteche Il ministro Franceschini annuncia una piattaforma digitale unica per i vari istituti italiani: sarà finanziata con 2 milioni di euro, March 10, 2017, online available at http://www.lastampa.it/2017/03/10/tecnologia/news/nasce-la-digital-library-ditalia-online-archivi-e-biblioteche-FtHZ2lMVzgwErudo5RSMbJ/pagina.html

- Lebert, M. The Project Gutenberg EBook of Project Gutenberg (1971-2009), etext available at http://www.gutenberg.org/files/31632/31632-0.txt
- Lessig L. Aaron's Laws: Law and Justice in a Digital Age. Harvard Law School Appointment Lecture available at http://today.law.harvard.edu/lessig-on-aarons-laws-law-and-justice-in-a-digital-age-video/
- Lewis-Krausdec, The Great A.I. Awakening How Google used artificial intelligence to transform Google Translate, one of its more popular services — and how machine learning is poised to reinvent computing itself, *The New York Times Magazine*, December 14, 2016 [online], available at https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html
- *Libraries enter the digital age*, National Public Radio, Science Friday podcast, May 11, 2011, transcript and sound available at http://www.npr.org/templates/transcript/transcript.php?storyId=10136882
- Library of Congress Announcement, November 29, 2016, Library of Congress, Digital Public Library of America To Form New Collaboration, online available at https://www.loc.gov/item/prn-16-207/
- Malik S., Kindle ebook sales have overtaken Amazon print sales, says book seller, *The Guardian*, August 6, 2012, online available at http://www.theguardian.com/books/2012/aug/06/amazon-kindle-ebook-sales-overtake-print
- Markoff J. & Wyatt E., Google is Adding Major Libraries to its Database, *The New York Times*, December 14, 2004 [online], available at http://www.nytimes.com/2004/12/14/technology/google-is-adding-major-libraries-to-its-database.html
- McLynn F. Review of The Age of Gold: The Story of an Obsession that Swept the World (by HW Brands), *The Guardian*, May 6, 2005 [online], available at https://www.theguardian.com/books/2005/may/07/featuresreviews.guardianreview11
- Miller C. and Bosman J., E-Books Outsell Print Books at Amazon, *The New York Times*, May 19, 2011, online available at http://www.nytimes.com/2011/05/20/technology/20amazon.html
- Milliot J., As E-book Sales Decline, Digital Fatigue Grows, *Publishers Weekly*, June 17, 2016, online available at http://www.publishersweekly.com/pw/by-topic/digital/retailing/article/70696-as-e-book-sales-decline-digital-fatigue-grows.html
- Mullins B. and Nicas J., Paying Professors: Inside Google's Academic Influence Campaign, *The Wall Street Journal*, 14 July 2017 [online], available at https://www.wsj.com/articles/paying-professors-inside-googles-academic-influence-campaign-1499785286
- New York Public Library, Privacy Policy (November 2016), available at https://www.nypl.org/help/about-nypl/legal-notices/privacy-policy/intro
- Pauler, N., and Ettig, D., Germany: Technische Universität Darmstadt wins landmark case with Hogan Lovells – Digital use of copyright protected works in public libraries, *Hogan Lovells Lime Green IP News*, April 20, 2015, blog-post available at http://www.limegreenipnews.com/2015/04/germany-technische-universitat-darmstadt-wins-landmark-case-with-hogan-lovells-digital-use-of-copyright-protected-works-in-public-libraries/

- Pogrebin R., For the Mid-Manchattan Library, a Redesign for the Future, *The New York Times*, November 16, 2016 [online], at: https://www.nytimes.com/2016/11/17/arts/design/for-the-mid-manhattan-library-a-redesign-for-the-future.html
- Posner R., Are Bookstores and Libraries Doomed by Digitization?, 03/04/2012 blog-post available at: http://www.becker-posner-blog.com/2012/03/are-bookstores-and-libraries-doomed-by-digitization-posner.html
- Rosati E., Waiting for a lower court to rein in resale? You'd sooner herd cats, *IP Kat Blog*, May 1, 2013, available at http://ipkitten.blogspot.gr/2013/05/waiting-for-lower-court-to-reign-in.html
- Rosati, E., AG Szpunar says that time-limited e-lending is allowed under EU law and interpretation of copyright norms must evolve with technology, *IP Kat Blog*, June 17, 2016, available at http://ipkitten.blogspot.gr/2016/06/ag-szpunar-says-that-time-limited-e.html
- Rosenwald M.S., This student put 50 million stolen research articles online. And they're free. *The Washington Post*, March 30, 2016, online available at https://www.washingtonpost.com/local/this-student-put-50-million-stolen-research-articles-online-and-theyre-free/2016/03/30/7714ffb4-eaf7-11e5-b0fd-073d5930a7b7_story.html?hpid=hp_hp-more-top-stories_journalwars-1140am%3Ahomepage%2Fstory
- Rosenberg S., How Google Book Search got lost, *Wired* [online], April 11, 2017, available at https://www.wired.com/2017/04/how-google-book-search-got-lost/
- Roush, W., Digitize This! Yahoo hopes to trump Google with its Open Content Alliance publishing venture, *MIT Technology Review*, October 20, 2005, online available at https://www.technologyreview.com/s/404830/digitize-this/
- Schofield, J. *More about Project Gutenberg's Michael Hart, RIP*, September 14, 2011 blogpost available at http://www.zdnet.com/article/more-about-project-gutenbergs-michael-hart-rip/
- Shields M., Tech giants unite against Google, *BBC News*, August 21, 2009, online available at http://news.bbc.co.uk/2/hi/technology/8200624.stm
- Singer P., Whither the dream of the universal library?, *The Guardian*, 19 April 2011 [online], available at https://www.theguardian.com/commentisfree/2011/apr/19/moral-imperative-create-universal-library
- Somers J., Torching the modern-day library of Alexandria, *The Atlantic*, April 20, 2017 [online], available at https://www.theatlantic.com/technology/archive/2017/04/the-tragedy-of-google-books/523320/
- Stavros Niarchos Foundation Cultural Center Announcement, The National Library of Greece, available at https://www.snfcc.org/about/vision/the-national-library-of-greece/
- Stihler C., EU copyright rules holding back cultural institutions, *The Parliament Magazine*, 12 July 2016 [online available], at https://www.theparliamentmagazine.eu/articles/opinion/eu-copyright-rules-holding-back-cultural-institutions
- Suber, P., The Open Content Alliance, SPARC Open Access Newsletter, issue #91, November 2, 2005, online available at: https://dash.harvard.edu/bitstream/handle/1/4552008/suber_oca.htm?sequence=1

- Swartz A. 2008. Guerilla Open Access Manifesto, available at https://archive.org/stream/GuerillaOpenAccessManifesto/Goamjuly2008_djvu.txt
- Taycher L., Books of the world, stand up and be counted! All 129,864,880 of you, August 05, 2010 at: http://booksearch.blogspot.com/2010/08/books-of-world-stand-up-and-be-counted.html
- *The Internet Archive - a chat with Brewster Kahle and a little look inside its book archive*, interview online available at https://www.youtube.com/watch?v=M6rsPgnA87Q
- *The Reports of My Death have Been Greatly Exaggerated: Libraries for the Future*, AIA New York, Center for Architecture, October 29, 2016, program available at http://cfa.aiany.org/index.php?section=calendar&evtid=9882
- Templeton J.W., Trump drops the mother of all bombs on libraries, *The Hill*, May 01, 2017 [online], available at http://thehill.com/blogs/pundits-blog/economy-budget/331362-trump-drops-the-mother-of-all-bombs-on-libraries
- Toobin J., Google's Moon: The quest for the universal library, *The New Yorker*, 5 February 2007 [online], available at http://www.newyorker.com/magazine/2007/02/05/googles-moon-shot
- Waddell K. The Research Pirates of the Dark Web. *The Atlantic*, February 9, 2016, online available at https://www.theatlantic.com/technology/archive/2016/02/the-research-pirates-of-the-dark-web/461829/
- Wyatt E., New Google Service May Strain Old Ties in Bookselling, *The New York Times*, October 20, 2004 [online], available at http://www.nytimes.com/2004/10/08/technology/08book.html

  Wu T., Leggo my Ego – Google Print and the Other Cultural War, *Slate*, October 17, 2005 [online], available at http://www.slate.com/articles/news_and_politics/jurisprudence/2005/10/leggo_my_ego.html