DANIEL REETZ

# The Why in DIY Book Scanning

## I.  INTRODUCTION

We live in interesting times. Digital cameras are now cheaper than textbooks—a profound inversion with roots in Moore's law[1] and Asian manufacturing—with consequences extending from the present moment into a future littered with electronic reading devices. In fact, the number of e-reader prototypes presented at the 2010 Consumer Electronics Show (CES)[2] was so great that it caused some tech reporters to complain.[3] Yet, strangely, none of these prototypes will allow us to read the books already on our shelves.

Though the future is always hazy, it is hardly controversial to predict that future books will be digitally authored and that most books will be digitally distributed. Furthermore, like all existing books, they will rely on the information from previous books. It is apparent, then, that having all extant books in digital form is not just desirable, but critical. The missing link is digitization—the process of converting print media into electronic media—which has historically been prohibitively difficult and expensive, keeping it largely in the hands of big players like Google, Microsoft, and Amazon. In light of this, the fact that the costs of the fundamental units of digitization (i.e., cameras and books) are approaching unity is a matter of great importance. Almost anyone can afford to digitize and, I will argue, no one can afford not to.

A thorough understanding of digitization begins with an explanation of the underlying hardware and relevant legal issues. In this article, I will frame the broader benefits of digitization using specific anecdotes from the Do-It-Yourself (DIY) Book Scanner Community—a growing group of over 700 people[4] developing book scanning technology from scratch and freely sharing their results.[5] I will use the same sorts of examples to show why the responsibility of digitization cannot be left solely in the hands of Google or any other big player, and how personal digitization

---

1.  Moore's law is not a law. It is a description of a long-term trend in which the number of transistors in computer processors and other integrated circuits doubles almost every two years. Peter S. Menell, *Can Our Current Conception of Copyright Law Survive the Internet Age?: Envisioning Copyright Law's Digital Future*, 46 N.Y.L. Sᴄʜ. L. Rᴇᴠ. 63, 72 (2002–03).

2.  The Consumer Electronics Show is an annual, non-public tradeshow where new consumer electronics are exhibited. *See About CES*, Iɴᴛᴇʀɴᴀᴛɪᴏɴᴀʟ CES, http://www.cesweb.org/aboutces.asp (last visited Sept. 12, 2010).

3.  Numerous tech bloggers complained, including such high profile blogs as Gizmodo and ZDnet, not to mention many smaller blogs like Technologizer. *See* Wilson Rothman, *There Are Officially Too Damn Many Ebook Readers*, Gɪᴢᴍᴏᴅᴏ (Jan. 7, 2010, 6:30 PM), http://gizmodo.com/5443102/there-are-officially-too-damn-many-ebook-readers; Joel Evans, *Too Many Ebook Readers and Other CES Thoughts*, ZDNᴇᴛ (Jan. 8, 2010, 11:31 AM), http://blogs.zdnet.com/mobile-gadgeteer/?p=2372; Harry McCracken, *Too Many Ebook Readers*, Tᴇᴄʜɴᴏʟᴏɢɪᴢᴇʀ (Jan. 8, 2010, 9:24 AM), http://technologizer.com/2010/01/08/too-many-ebook-readers.

4.  This reflects the number of members as of the date of this article's publication. The DIY Book Scanner Community gains about three new members a day.

5.  *See infra* note 13 ("Do-It-Yourself Book Scanning is using cheap, compact cameras and Free software to scan books quickly and efficiently.").

is not just feasible and affordable, but desirable and necessary. I will close with some thoughts on the future.

## II.  BACKGROUND INFORMATION ABOUT SCANNING AND COPYRIGHT

Put simply: a physical book enters the digital domain via digitization. Digitization is simply capturing a digital image of a document. Digitization equipment generally falls into three categories.[6] At the low end are consumer flatbed scanners, priced around $40 to $250. In the middle are sheet-fed scanners, for $300 to $500. Perched at the top are face-up and fully automated scanners, which range from $7000 to $150,000. The cost of an individual scanner, however, extends beyond its purchase price; each type exhibits a unique set of tradeoffs. Flatbed scanners are the most widely available and the least suitable for the task of DIY book scanning. While providing excellent image quality, often for remarkably low prices, they are slow. The fastest flatbed scan times are between ten and forty seconds per page, due to the mechanical movement of the imager across the scanning bed.[7] This represents a hard limit to the number of pages that can be digitized in a day. For example, at thirty seconds per page, a 400-page book represents over three hours of nonstop scanning. They also require the book to be pressed flat against the platen[8] to acquire a usable image. Forcing the binding open in this manner can irreparably damage a bound volume. Furthermore, though large format flatbed scanners exist, the vast majority only scan up to letter-size pages, which makes them useless for scanning oversize volumes, maps, and centerfolds.

Sheet-fed scanners are able to ingest and scan a volume of pages automatically, but require the destruction of the book. To prepare for scanning, the binding of the book is sawed off with a table saw or cut off with a powerful hydraulic shear. The speed and automation advantage of sheet-fed scanning is offset by the destruction of the book. In some cases it is possible to re-bind a text after it is cut and scanned in this fashion, but this is expensive and obviously inappropriate for whole classes of books: the old and the rare, those with sentimental value or signatures, and texts with shallow gutters. This process can be particularly difficult for those of us who die a little watching our books hit the bandsaw. After all, one does not scan a book to spite it.

Face-up and automated scanners, on the whole, treat the book more gently. Though designs vary, the salient feature of a face-up scanner is that it uses one or two cameras and a V-shaped platen that descends into the gutter of the open book to press the pages flat. With two cameras, each capable of capturing a page in a fraction of a second, the speed of scanning is limited only by the page-turning mechanism—

---

6.    *See* Douglas Downing et al., Dictionary of Computer and Internet Terms 445–46 (9th ed. 2006).

7.    The imaging element in a flatbed scanner comprises a line of pixels. It must be translated horizontally across the document to capture an entire 2-D image. Resolution is increased by moving this imaging element more slowly and reading it out more often. *See* Downing et al., *supra* note 6, at 226.

8.    The platen is the flat glass surface of the scanner.

which is often a human. Though fully-automated page-turning scanners do exist, most organizations that employ such scanners also employ an individual to watch the mechanism.[9] For this reason, most projects rely on humans to turn the pages, leaving the limit at two to fifteen seconds to lift the platen, turn the page, place the platen, and press the capture button. For archival projects and books for which scanning does not constitute a justification for destruction, face-up scanners are the fastest and most gentle way to scan a book. They can accommodate all sizes of printed matter, including newspapers, though large materials may require custom platen and cradle construction. For those of us who care about the preservation of our books, and who live finite lives, the obvious choice is the face-up scanner.

Face-up scanners, then, are nearly ideal. They are gentle. They are fast. They are flexible. They do not destroy your books. They are, however, accordingly expensive. But pocket digital cameras, as inexpensive and absurdly capable as they have become, disrupt the market. They are an order of magnitude cheaper than the Digital Single Lens Reflex cameras (DSLR) used on face-up scanners.[10] I built a scanner using these cheap cameras and materials scavenged from dumpsters, and shared the plans online.[11] Subsequently, nearly one hundred scanner builders have built scanner frames using materials ranging from IKEA furniture to aircraft aluminum, whisky boxes to Styrofoam. Everything necessary to begin building a scanner, including a complete tutorial, open source software,[12] and moral support is available on the Internet[13] and at your local hardware store. But say you decide to build such a scanner, when is it legal to use?

Perhaps unsurprisingly, such digitization hardware has many potential infringing uses under current copyright law. Each work exists in essentially one of three legal states: (1) public domain, (2) copyright status unknown, or (3) in-copyright. The least restrictive and slowest-growing category is the public domain, which consists of works that may be used without any legal restrictions.[14] These works can be freely copied, modified, and sold with no further action, provided their public domain status can be discovered or inferred. Works that have an unknown copyright status

---

9.  *See Robotic Book Scanner (RBS)*, Qidenus Tech., http://www.roboticbookscan.com/documents/Qidenus%20Technologies_RBS%20Systems.pdf (last visited Sept. 12, 2010).

10. DSLRs are large digital cameras featuring an old-style mirror box and interchangeable lenses. Downing et al., *supra* note 6, at 464.

11. Daniel Reetz, *DIY High-Speed Book Scanner from Trash and Cheap Cameras*, Instructables, http://www.instructables.com/id/DIY-High-Speed-Book-Scanner-from-Trash-and-Cheap-C (last visited Sept. 12, 2010).

12. Open source software is software which permits access to the original source code. Access to the original code permits modification and extension by peers and end users. Downing et al., *supra* note 6, at 355.

13. *See* DIY Bookscanner, http://www.diybookscanner.org (last visited Sept. 12, 2010); *Forum News*, DIY Bookscanning, http://www.diybookscanner.org/forum/ (last visited Sept. 12, 2010) (discussing DIY bookscanning).

14. *See Welcome to the Public Domain*, Copyright and Fair Use, Stanford University Libraries, http://fairuse.stanford.edu/Copyright_and_Fair_Use_Overview/chapter8/8-a.html (last visited Sept. 12, 2010).

are problematic; they might be freely usable or they might require permission from the rightsholder. In-copyright works require the same arduous investigation. Their copyright status and the contact information of the rightsholder must be discovered to obtain permission from him/her before legal copying can occur.[15]

Fortunately, there are key exceptions for the use of copyrighted material permitted under the fair use doctrine. These exceptions are predicated on the purpose and character of the use, the nature of the copyrighted work, the amount of the work used, and the effect of use on the potential market.[16] As noted by many courts, whether or not a specific copying act is a fair use or an infringement is typically determined on a case-by-case basis and there are *no* guarantees.[17] My proposal hinges on the opinion that the fair use doctrine supports the wholesale space shifting[18] of books by their owners, an opinion shared by many courts when discussing the copying of similar forms of copyrighted works.[19] The fair use doctrine is the sole legal mechanism allowing scanner owners and builders to digitize in-copyright works that they have purchased.[20]

It is curious that discovering the copyright status of any work of interest is difficult. All parties producing copyrighted works, including publishers, authors, and artists, assert that their copyrights are of great value. However, historically, there has

---

15.   *See The Basics of Getting Permission*, Copyright and Fair Use, Stanford University Libraries http://fairuse.stanford.edu/Copyright_and_Fair_Use_Overview/chapter1/1-b.html (last visited Sept. 12, 2010).

16.   *Measuring Fair Use: The Four Factors*, Copyright and Fair Use, Stanford University Libraries, http://fairuse.stanford.edu/Copyright_and_Fair_Use_Overview/chapter9/9-b.html (last visited Sept. 12, 2010). For an in-depth explanation, see generally *Fair Use*, Copyright and Fair Use, Stanford University Libraries, http://fairuse.stanford.edu/Copyright_and_Fair_Use_Overview/chapter9/index.html (last visited Sept. 12, 2010).

17.   *See* Mattel, Inc. v. Walking Mountain Prods., 353 F.3d 792, 800 (9th Cir. 2003) ("To determine whether a work constitutes fair use, courts engage in a case-by-case analysis and a flexible balancing of relevant factors." (citing Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 577–78 (1997))).

18.   Wholesale space shifting of books by their owners is perhaps best explained by example. Rob, the www.diybookscanner.org forum moderator, is a connoisseur of science fiction. He has amassed a collection of several thousand science fiction paperbacks. *See* Rob, Comment to *Post Something About Yourself Here (The Hello Thread)*, DIY Book Scanner Forum (June 4, 2009, 1:33PM), http://diybookscanner.org/forum/viewtopic.php?f=5&t=2. Subsequently, he built a scanner to shift these books to electronic formats suitable for reading on his Kindle and Sony Reader. *See id.* For some individuals, purchasing an entire collection of books twice may represent a fair value proposition. But for people with large collections, and for libraries and other book repositories, purchasing their entire collection a second time does not represent a good value or a sane approach, even if it represents the fondest dreams of publishers.

19.   See Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 443–56 (1984), and the 1999 Rio MP3 player decision, Recording Indus. Ass'n of Am. v. Diamond Multimedia Sys., 180 F.3d 1072, 1079 (9th Cir. 1999).

20.   However, even as a non-lawyer, I understand that the fair use doctrine as written is not a paragon of legal clarity or a matter enjoying universal legal concord. This article is about what is possible, pressing, and personal in book digitizing and, as such, where progress runs counter to the present legal environment, distressing as it may be, the suggestion is to reform, re-factor, and revise the law, not to break it.

been no commensurate effort to preserve and make accessible copyright information. Several approaches to solve this problem are underway, including the proposed Book Rights Registry that would be created by the Google Books Settlement.[21] The Book Rights Registry is a not-for-profit entity that will maintain a database of rights information for all books that are covered by the settlement.[22] The Book Rights Registry is not the only entity seeking to make book rights information searchable and accessible. Tulane Law School is actively developing a web-based software tool called the Durationator.[23] Once completed, the Durationator will enable non-experts, such as myself, to reach legally valid conclusions about the copyright duration status of a given creative work.[24] This tool is currently under heavy development. Part of the development cycle is the ingestion of a series of out-of-print, public domain publications from the U.S. Copyright Office. Tulane Law School's archive of these records was destroyed by Hurricane Katrina, but an additional complete set is on loan from another institution.[25] Using the DIY book scanner tutorial as a starting point, Benjamin Varadi, a Tulane graduate student working on the project, has recently constructed a custom scanning system[26] to affordably and non-destructively scan these records. In this way, DIY book scanning technology is engendering and enabling more digitizing and digitizing technology.

## III.  THE BROADER BENEFITS OF DIGITAL BOOKS

The broad benefits of digital books (and, implicitly, digitizing) are plainly visible in the myriad ways people use them. The DIY Book Scanner Community's "Hello Thread" features over one hundred members testifying as to how they have benefited from digitizing their own books.[27] These testimonials include descriptions of accessibility for the print-disabled, ease of transport, reduced space requirements, plasticity and flexibility, and a general imperishability.

---

21. "Google Books Settlement" refers to the proposed settlement agreement submitted to the court by the parties to the *Author's Guild, Inc. v. Google, Inc.* case in October 2008 and amended and re-submitted in November 2009. The court had not approved the agreement as of the publication date of this article. *See* Amended Settlement Agreement at art. VI, Authors Guild, Inc. v. Google, Inc., No. 05-CV-8136DC (S.D.N.Y. Nov. 13, 2009), *available at* http://thepublicindex.org/docs/amended_settlement/Amended-Settlement-Agreement.pdf.

22. *See* Press Release, Google, Authors, Publishers, and Google Reach Landmark Settlement (Oct. 28, 2008), *available at* http://www.google.com/intl/en/press/pressrel/20081027_booksearchagreement.html.

23. *See* Elizabeth Townsend Gard, *About the Durationator*, Durationator.com, http://www.durationator.com/about.php (last visited Sept. 12, 2010).

24. *Id.*

25. *Id.*

26. *See* Benjamin Varadi, *Build Thread: Tulane CIPLC (now with a TON of photos!)*, DIY Book Scanner (Jan. 25, 2010, 6:35 PM), http://diybookscanner.org/forum/viewtopic.php?f=1&t=262.

27. *See generally* Comments to *Post Something About Yourself Here (The Hello Thread)*, DIY Book Scanner, http://diybookscanner.org/forum/viewtopic.php?f=5&t=2 (last visited Sept. 12, 2010).

Accessibility for the print-disabled is an important benefit of digitizing books.[28] Tristin joined the DIY Book Scanner Community in June 2009. A mechanical engineering major, he is responsible for consuming vast amounts of technical literature as a consequence of his coursework. He has some difficulty reading with his eyes—his visual system does not process text as quickly or easily as his ears. An avid reader, both out of necessity and out of a love of learning, he designed and implemented his own DIY book scanning system to enable his computer to read aloud to him. Another forum member is paraplegic and wishes to read an out-of-print science fiction series. After being repeatedly refused a digital edition by the publisher, he joined the forum to design his own book scanner for his family to scan his copies of the physical books. A rightsholder's refusal to provide a digital edition is not an isolated incident. Bestseller J.K. Rowling has famously refused the possibility of electronic versions of her books.[29] Startlingly, though many print-disabled are not only willing, but anxious to purchase accessible electronic texts, texts are not being made available for purchase, or are only available in inaccessible formats.[30] As Chris Danielsen pointed out at the *D Is for Digitize* conference,[31] e-book platforms like the Amazon Kindle are inaccessible to the blind and many e-book formats are useless on Braille reading devices. The distressingly small fraction of all digital books that are accessible to the print-disabled (a class including the blind, dyslexics, paraplegics, amputees, and the enormous and increasing swath of people with age-related vision disorders), is a topic that deserves its own article. For the purposes of this article, it is enough to say that, for better or for worse, one of the best options presently available to the print-disabled is to personally scan and, using Optical Character Recognition (OCR) software,[32] convert the works into usable text, or employ a person or service

---

28. As defined by the Amended Settlement Agreement, "'[p]rint [d]isability' means any condition in which a user is unable to read or use standard printed material due to blindness, visual disability, physical limitations, organic dysfunction, or dyslexia." Amended Settlement Agreement, *supra* note 21, at art. I.

29. Matt Frisch, *Digital Piracy Hits the E-book Industry*, CNN (Jan. 1, 2010, 9:52 AM), http://www.cnn.com/2010/TECH/01/01/ebook.piracy/index.html; *see also* Mike Masnick, *CNN's Take On 'Book Piracy'* (Jan. 4, 2010, 10:59 PM), http://www.techdirt.com/articles/20100104/0429587587.shtml.

30. Digital Rights Management (DRM) is a software mechanism to prevent unauthorized copying of digital files. Downing et al., *supra* note 6, at 156. DRM is closely related to Technical Protection Measures (TPM). DRM-encumbered formats are typically not able to be used with blind reader hardware, which works best with the DAISY format, plain ASCII text, RTF, HTML, etc. Blind reader hardware includes, but is not limited to, text-to-speech converters, which read text aloud in a synthesized voice, and Braille-reading hardware, which converts digital text into physical bumps.

31. The *D Is for Digitize* conference, sponsored by the Institute for Information Law and Policy, was held at New York Law School in New York, New York on October 8–10, 2009. The purpose of the conference was to discuss the forthcoming Google Books Settlement. An archive of the conference is available online. Symposium, *D Is for Digitize*, *available at* http://www.nyls.edu/centers/harlan_scholar_centers/institute_for_information_law_and_policy/events/d_is_for_digitize (last visited Sept. 12, 2010).

32. OCR is the process of converting an image of a character into text. Downing et al., *supra* note 6, at 356.

(such as BookShare) to scan and convert works for them.[33] Here, a DIY book scanner is a rare example of a totally unrestricted, affordable, and customizable solution.

It is unsurprising that many members of the DIY Book Scanner Community have amassed substantial "reference" text collections, considering that the first five people to join the forum and build scanners were two programmers, two mechanical engineers, and an intellectual property lawyer. The subsequent influx of dictionary lovers, religious studies majors, and history buffs (and re-enactors) all brought up the same point: it is absurd to attempt to travel or commute with this kind of reference content, however central to their work and study. Furthermore, a surprising number of masochistic members wish to read their dictionary collections on the beach. And if not the beach, the landfill. The Byzantine laws, regulations, and design guidelines that inform sampling, monitoring, maintenance, and construction of these most permanent book repositories would surely benefit from mobile content.

Another issue is the raw space that a book collection occupies. A dozen books stuff a backpack, but a hundred books fill a bookshelf. Umberto Eco's reported 30,000 volume library is not just practically immovable; it is architecturally as well. [34] These issues are never more glaring than when one is moving, be it across town or across the country. A single hard disk or memory card weighs little and takes up little space. Having all your books scanned means you can take advantage of your first-sale rights[35] and convert your books to bits, heat, and light.

College textbooks are a recurring topic of discussion among scanner builders. Even ignoring for the moment the ever-rising cost[36] of these glossy, unrecyclable hardcovers, the physical burden placed on college students is substantial. Carrying around three to five textbooks in a backpack is hard on the spine, made worse if the student is using a messenger bag, and worse still if the student is on crutches or in a wheelchair. The advantages of digital editions, in this respect, are clear. In addition, there are other less obvious advantages to digital textbooks. As the prices of textbooks rise, they are increasingly targets for theft and resale on college campuses. Digital copies backed up on a flash drive and read on a laptop or e-reader are not as vulnerable or attractive to thieves. In fact, statutory damages would likely make the thief think twice about resale.[37]

---

33.  *See How Bookshare Works*, BOOKSHARE, http://www.bookshare.org/_/aboutUs/howBookshareWorks (last visited Sept. 12, 2010).

34.  Nigel Farndale, *Heavyweight Champion*, THE DAILY TELEGRAPH (May 24, 2005), http://www.telegraph. co.uk/culture/books/3642577/Heavyweight-champion.html.

35.  The first sale doctrine is a limitation on copyright that allows the original purchaser to sell or give away a copyrighted work. *See* 17 U.S.C. § 109 (2006 & Supp. 2010).

36.  *See* Saffron Zomer, *Exposing the Textbook Industry: How Publishers' Pricing Tactics Drive Up the Cost of College Textbooks* (Feb. 2007), http://cdn.publicinterestnetwork.org/assets/awAU_yjlm1I9khl6VOv_mw/ Exposing_the_Textbook_Industry.pdf; *see also* Josh Rohrscheib, *Textbook Example of Price Gouging*, JOSH ROHRSCHEIB (Jan. 21, 2006, 12:13 AM), http://rohrscheib.blogspot.com/2006/01/textbook-example-of-price-gouging.html.

37.  Statutory damages for copyright infringement cases, on the whole, greatly exceed the value of textbooks. *See* 17 U.S.C. § 504(c) (2006 & Supp. 2010).

It is startling that more textbooks are not available in digital editions. No major textbook publisher uses paste-up layout methods; nearly every publisher authors electronically.[38] Converting one electronic document format to another is trivial in comparison to the enormous pipeline necessary to produce and ship a paper book. While publishers claim that new editions are published to keep up with pedagogical practice and correct mistakes, again, digital editions do so more easily and more efficiently. It is far less complicated and resource-intensive to apply a *digital patch*[39] to an electronic textbook than to produce, edit, lay-out, print, ship, stock, and sell a new edition. This patching idea is far from new, even in print books. One particularly prescient example was the 1908 *Nelson's Encyclopedia*,[40] which had a binder-like binding and was refreshed as new facts and new information were made available.[41] Furthermore, this practice need not be limited to non-fiction books. Poet Stéphane Mallarmé envisioned books and staged performances that could be reconfigured as necessary to meet the demands of a democratic society.[42] Flexible textbook models already exist outside the United States and even outside the digital domain. I taught at a university of engineering in Russia where instructors simply collaged together whole textbooks from original works, other texts, and Internet sources. These short, staff-and-faculty produced, frequently updated texts were photocopied, bound, and distributed to the students with remarkable results, including reduced costs, and ensured relevant and consistently up-to-date material. Digital equivalents boast the same advantages but eliminate those last three steps. Unfortunately, the current state of copyright law in the United States forbids this innovative and effective pedagogy.

38. Since the advent of the personal computer in the early 1980s, publishers have moved to digital publishing systems. Rather than painstakingly paste together book layouts, books are now typeset in software on a computer. Authorship has seen a similar shift. Rather than using typewriters, authors now write their books in text editors like Microsoft Word. *See History of Desktop Publishing and Digital Design*, Design: Talk Board, http://www.designtalkboard.com/design-articles/desktoppublishing.php (last visited Sept. 21, 2010); *see also The History of Prepress*, Prepressure, http://www.prepressure.com/library/prepress-history (last visited Sept. 21, 2010).

39. A digital patch is a binary file which can be appended to an existing binary to improve it or correct a problem with it. *See* Downing et al., *supra* note 6, at 372.

40. Nelson's Encyclopedia (Franklin J. Meine et al. eds., 1940). *Nelson's Encyclopedia* was previously published as the *Nelson's Complete Encyclopedia* and copyrighted in 1905. Library of Congress Catalog Record, http://lccn.loc.gov/41004742 (last visited Sept. 12, 2010). *Nelson's* was published through the 1930s. Toward the end of this run, the original books cost $99.50 and $6 per year for updates, which were mailed to the subscriber. Jessamyn West, *"The only perfect reference work" Nelson's Perpetual Loose-Leaf Encyclopedia*, BoingBoing (Jan. 30, 2010, 8:14 PM), http://boingboing.net/2010/01/30/nelsons-loose-leaf-encyclopaedia.html; *see also Popular Mechanic's Advertising Section*, Popular Mechanic's Mag., March 1930 at 2, *available at* http://books.google.com/books?id=cOIDAAAAMBAJ&lpg=PA2&ots=u6xlrPoJEn&dq=%22a%20Reader%27s%20Guide%20to%20Nelson%27s%20Looseleaf%20encyclopaedia.%22&pg=PA2#v=onepage&q=%22a%20Reader%27s%20Guide%20to%20Nelson%27s%20Loose-leaf%20encyclopaedia.%22&f=false.

41. *See Unique Plan For Reference Book*, N.Y. Times, Jan. 4, 1908 at BR2.

42. Anna Sigrídur Arnar explores Mallarmé's obsession with the idea of creating new forms of the book that would actively engage readers by inviting them to rearrange the structure of a text. *See generally* Anna Sigrídur Arnar, The Book as Instrument: Stéphane Mallarmé, the Artist's Book and the Transformation of Print Culture (forthcoming Dec. 2010).

While there were previously fair use exceptions for educators to create "coursepacks" containing photocopied textbook material, these privileges relating to commercial copying were later rescinded in *Basic Books Inc. v. Kinko's Graphics Corp.*[43] Fortunately, there is a nascent movement to produce open source textbooks that may eliminate one aspect of this copyright problem with no change in copyright law.[44] There is a bill in the Senate at the time of this writing that is designed to improve this situation by creating funding for electronic textbooks which will be "made available free of charge to, and may be downloaded, redistributed, changed, revised, or otherwise altered by, any member of the general public."[45] This particular distribution model has the added positive effect of preventing the loss of books by destruction of individual volumes or collections.

> Time and accident are committing daily havoc on the originals deposited in our public offices. The late war has done the work of centuries in this business. The lost cannot be recovered; but let us save what remains: not by vaults and locks which fence them from the public eye and use, in consigning them to the waste of time, but by such a multiplication of copies, as shall place them beyond the reach of accident.[46]

Here, Thomas Jefferson refers to the damage done during the Revolutionary War, foreshadowing the congressional library fire of 1851. The destruction of libraries through burning was not restricted to Washington, D.C.; both Alexandria and Iraq have faced the total destruction of their central archives.[47] The centralized nature of the Google scanning effort is of concern to many.[48] Taking the long view, it makes good sense to scan each book more than once and to have original scans—e.g., not just copies of the images Google makes public, which are not the raw scans[49]—of many books; this is true of the many slight variations and major editions of each book; translations of each book; and even personalized copies. Furthermore, it makes sense to store these scans in many locations. If there is one thing we can still learn from the libraries of Alexandria, Congress, and now Iraq, it is that a distributed library is a robust library. This extends to all works, including in-copyright works,

---

43.  758 F. Supp. 1522, 1531 (S.D.N.Y. 1991); *see also Academic Coursepacks*, Copyright and Fair Use, Stanford University Libraries, http://fairuse.stanford.edu/Copyright_and_Fair_Use_Overview/chapter7/7-a.html#1 (last visited Sept. 12, 2010).

44.  *See* California Open Source Textbook Project, http://www.opensourcetext.org (last visited Sept. 12, 2010); Wikibooks, http://en.wikibooks.org/wiki/Main_Page (last visited Sept. 12, 2010).

45.  Open College Textbook Act of 2009, S. 1714, 111th Cong. § 4(f)(3)(B) (2009).

46.  Robert Darnton, *Burn a Country's Past and You Torch Its Future*, Wash. Post, Apr. 27, 2003, at B1.

47.  *Id.*

48.  *See* Pamela Samuelson, *Cyberspace & The Law: Privacy, Property, and Crime in the Virtual Frontier: Google Book Search and the Future of Books in Cyberspace*, 94 Minn. L. Rev. 1308, 1314 (2010).

49.  Videotape: C is for Culture, D Is for Digitize Conference held by New York Law School, Conversation between Daniel Reetz and Dan Clancy during the panel (Oct. 9, 2009), *available at* http://nyls.mediasite.com/mediasite/SilverlightPlayer/Default.aspx?peid=4f3acab65b424c17b0eeb26ec8be6bc2. Clancy points out that Google scans are saved in JPEG2000 and not perpetually preserved in their rawest format, which would be prohibitively costly in storage terms. *Id.*

and perhaps *especially* to those in-copyright works. After all, the existence of so-called orphan works[50] exposes a fault of the publishing industry's stewardship of copyright information. If they did not store these critical records, it is difficult to imagine that they would have made a store of the printing plates, contact sheets, and other materials that would be needed to reconstruct these books.

Why should this be our responsibility? What does this have to do with DIY book scanning? Why are other people unable to do it? Much of what has been discussed so far could be characterized as a general call for digitization, but for those of us with eyes and limbs intact, when does digitization become personal?

## IV. WHY WE SHOULD DIGITIZE BOOKS

Suryandaru, a village official in Indonesia, was among the first to contact me regarding the scanner tutorial. He described the appalling condition of regional village "holy books," which contained hand-drawn maps of land ownership, birth and death records, and more. Any dispute over land ownership is settled with the information in these books. It is self-evident that these books are essential in disaster conditions, such as those recently faced by Indonesia: earthquakes, fires, and floods. These damp, moldy, and bug-ridden books are dematerializing during a time of great vulnerability. Their fragile condition makes photocopying a poor solution and flatbed scanning takes time better used elsewhere. The DIY Book Scanner Community took up a donation and sent Suryandaru some cameras and, while they were in transit, he and his son Oka built a scanner frame. They are now scanning village holy books to keep the peace in tumultuous circumstances.

This village official's predicament—and effort—exposes something unaddressed by current mass digitization efforts: maintaining government documents which are typically in the public domain and of public interest and import. His radical approach to keeping the peace in and around his village demonstrates the viability of the DIY book scanner approach for underfunded, understaffed government offices of all types. Furthermore, the existence of a variety of free, simple plans and open source software helps ensure that no shipping or financial issue can interfere with the digitization of critical materials, even in less than ideal or disaster conditions. This is particularly true as the world's adoption of mobile computers (in the form of mobile phones, personal navigators, and digital cameras) allows their access and distribution in disaster conditions.[51] Indeed, one can make a DIY book scanner from almost anything, including wreckage.

---

50.  Orphan works are works where the copyright owner is unknown or not enough information is available to contact the copyright holder for permission. Jennifer Suzanne Bresson Bisk, *Book Search is Beautiful?: An Analysis of Whether Google Book Search Violates International Copyright Law*, 17 Alb. L.J. Sci. & Tech. 271, 282 (2007).

51.  It is interesting to note that Canon Powershot cameras used to construct DIY book scanners can function as e-book readers. All digital cameras are effectively small computers with a display. The Canon Powershot series has famously been hacked, and custom open source firmware made available. This firmware incorporates e-book reader functionality. *See* Donald Melanson, *Cannon Firmware Hack*

Rare and precious books are not limited to musty leather-bound first editions in the back of some Parisian shop; they can be handmade; limited edition artist books; sketchbooks or notebooks; instruction manuals for equipment produced in limited numbers; a late grandmother's recipe books; diaries; and even photo albums. Of this class of books, many cannot be removed from their surroundings by archival edict or social contract. In fact, social contract dictates access to many books and bound records; one cannot simply waltz into a small-town city hall and expect access to the records of interest.[52] Furthermore, these books will almost certainly fall through the cracks of mass digitization efforts—after all, Google's book scanning stations will not be delivered to your grandmother's living room in the foreseeable future.[53] In such scenarios it becomes essential to have a portable scanner or to build a scanner on-site. Forum member Ann, who works with the Minnesota Maritime Historical Society digitizing ships' logs, has discovered that particular agencies, such as the National Archives and Records Administration (NARA), have technical restrictions such as no additional lighting.[54] In these circumstances, having control of the construction of the scanner can ensure access.

Where are libraries in all this? Well, the vast majority of libraries have digitized little, if anything, on their own. Of the many contributing factors,[55] three stand out:

1. The cost of digitizing equipment.[56]

2. The cost of labor to run the equipment.[57]

3. The absurdly high statutory damages for copyright infringement.[58]

As discussed above, the DIY book scanner, and related efforts,[59] solves the first problem. The second could be solved as well. Consider that libraries have preserved our books thus far—far enough to grant us the ability to digitize them. To subject these same, strained institutions to the additional pressure of having to purchase

---

*Unlocks Features on DIGIC II–based Cameras*, ENGADGET (Sept. 11, 2007, 1:05 PM), http://www.engadget.com/2007/09/11/canon-firmware-hack-unlocks-features-on-digic-ii-based-cameras.

52. Particularly in a foreign country, as many genealogists might wish.

53. Google has developed custom scanning stations that use proprietary methods to capture books. They are a closely guarded trade secret. *See* Bobbie Johnson, *The Secret Behind Google's Book Scanning Project*, GUARDIAN.CO.UK TECHNOLOGY BLOG (May 4, 2009 9:30 BST), http://www.guardian.co.uk/technology/blog/2009/may/04/google-gadgets.

54. *See Obtaining Copies of Records*, THE NATIONAL ARCHIVES, http://www.archives.gov/research/order/ (last visited Sept. 12, 2010).

55. *See* Samuelson, *supra* note 48, at 1311–14.

56. *Id.* at 1312.

57. *Id.; See* DECAPOD PROJECT (Sept. 12, 2010), http://sites.google.com/site/decapodproject/ (last visited Sept. 12, 2010).

58. *See* 17 U.S.C. § 504(c) (2006 & supp. 2010).

59. *See, e.g.,* DECAPOD PROJECT, *supra* note 57; BOOK LIBERATOR (Sept. 12, 2010), http://bkrpr.org/ (last visited Sept. 12, 2010).

their entire collection again is uncharitable. This situation arises because publishers see digital editions as distinct from physical editions. As such, they feel that it is appropriate to charge for digital editions on a per-volume or subscription basis. Similarly, charging the people tasked with keeping books safe on the shelves with digitizing each and every volume is unfair. While Google and others have been rushing to digitize every possible book to create a competitive advantage, the rest of us do not need to hurry. Digitization projects have been going on for decades, and will continue. This is now the time to distribute the effort, which can move equally quickly when it involves millions of people, their cameras, and their books.

Library patrons could assist in the digitization of the library's volumes. Every library has a copy machine. Next to the copy machine, imagine an inexpensive, simple book scanner. Patrons with late fees could pay their fees by spending a few minutes scanning the book of their choice. Library-loving volunteers could have their names digitally affixed to volumes they digitized, something like a name-etched brick in a donor wall. Book lovers could scan books in exchange for extended check-out times, and out-of-service-area potential patrons could avoid access fees by trading time against the usual fee. Public domain works could be made immediately available and the rest stored. After all, on a long-enough time scale, even infringing acts contribute to the public domain.[60]

Though in-copyright works and orphan works comprise the majority of library collections,[61] whole classes of other materials and special collections exist. By way of example, the North Dakota State University (NDSU) library maintains a sizable collection of government documents, as well as several community-related special collections. Government documents, to the average reader, may seem somewhat uninteresting, but in terms of outreach and education, they have remarkable value. For example, the current organic gardening movement would benefit from scanned copies of the NDSU Agricultural Extension Service's one-pagers of practical advice. Another example is the corpus of U.S. Navy training manuals, which cover everything from the use of basic hand tools to surgery at sea. If you have any reservations about the value of scanning such works, simply search YouTube for "How-To." A practical example: as a part of my first book scanner tutorial, I scanned the *Basic Hand Tools Manual*[62] and made it available for those people who might want to build a scanner but lacked the necessary experience.

Libraries themselves are creators of book-like materials, records, and so on. While the present discussion surrounding Google Books focuses on the desires of publishers and select authors, it is important to note that many institutions—public, private, federal, and state—produce documents that would benefit from fast, inexpensive digitization and distribution.

---

60.  Of course, as has historically been the trend, this is only true if copyright is not extended into infinity.

61.  Brian Lavoie & Lorcan Dempsey, *Beyond 1923: Characteristics of Potentially In-Copyright Print Books in Library Collections*, D-Lib Mag., Nov.–Dec. 2009, *available at* http://www.dlib.org/dlib/november09/lavoie/11lavoie.html.

62.  To find a copy using Open Library, see Open Library, http://openlibrary.org/ (last visited Sept. 12, 2010).

Not all libraries lag. This discussion would be incomplete without a mention of the tireless exemplar, the Internet Archive.[63] Founded by Brewster Kahle, the Internet Archive has cached and made searchable the entire Internet from the days of ARPANET.[64] Furthermore, the Internet Archive has constructed its own book scanners[65] and made many public domain titles available.

In a recent presentation, Professor James Grimmelmann pointed out that one of the first uses of the printing press was to produce many copies of books that existed in other forms, an early form of format shifting.[66] Later, the variety and quality of book copies increased with the advent of pirated books like George Faulkner's *Dublin Editions*, which were smaller, more portable, cheaper copies of existing works.[67] Today, we can not only buy a book in many editions (sometimes back to the first) or in editions with fake ISBNs (such as complementary instructor editions of textbooks),[68] but we also often have books modified after first sale. Ex-libris plates, author signatures, handwritten notes and dedications, and even detritus-like pressed flowers contribute to the individuality of a volume. While the perfection of a digital copy is absolute, there is a personal, human attachment to specific instances of any volume. Being able to produce your own copy is essential. This has even more importance for modified and handmade volumes, as well as volumes culturally allocated for annotation such as yearbooks, scrapbooks, guestbooks, and more. These volumes have a long history: Thomas Jefferson spliced the Greek and Latin *Julius Caesar* together for personal reading purposes.[69] Furthermore, the slight differences between the many editions of books, which are of themselves of academic interest, are perhaps best revealed with digital copies. Differences in textual word content reveal items of interest to lexicographers, literature critics, and historians. Slight differences in

---

63. *See* Internet Archive, http://www.archive.org (last visited Sept. 12, 2010). The Internet Archive, a non-profit digital library dedicated to preserving and providing access to information, has created a searchable archive of nearly the entire Internet as well as some early networks with the cooperation of many individuals, universities, and organizations. *See id.*

64. ARPANET was a military network project which eventually gave rise to the Internet. Downing et al., *supra* note 6, at 27.

65. *See* Jared Benedict, *More Details on Internet Archive's Scribe Book Scanner Project*, The Future Is Yesterday, (Feb. 10, 2006, 9:24 PM), http://redjar.org/jared/blog/archives/2006/02/10/more-details-on-open-archives-scribe-book-scanner-project/; s*ee also* Pictures of the Internet Archive's Scribe Scanning Stations, Scribe Development, http://www.aipengineering.com/scribe/index.html (last visited Sept. 12, 2010).

66. Videotape: Conference on Books, Computers, and the Law, held by Archivists Round Table of Metropolitan New York, Inc. (Jan. 12, 2010), *available at* http://laboratorium.net/archive/2010/01/13/archivists_presentation_archived.

67. *See* George Lillo, The London Merchant ix (William H. McBurney ed., 1965).

68. These books often have a fake ISBN on the outside cover; the ISBN on the inside cover, however, is the same as any other edition. Presumably the publisher's intention is to prevent resale.

69. *Thomas Jefferson's Library: Library Reconstructs, Displays Founding Collection*, Library of Congress, http://www.loc.gov/loc/lcib/0806/detail/jefferson06.html (last visited Sept. 12, 2010).

printing could be used forensically, for example, as a "fingerprint" of a particular press.[70]

The point at which the importance of digitization becomes personal is not necessarily yours to choose. My next example is related, and it is my own. I built my own scanner and scanned all of my textbooks for use on my laptop computer. Over the Christmas holiday, the building I worked in collapsed,[71] with my office at the nadir of a chevron-shaped slump in the façade. The subsequent release of steam boiled my books under the rubble, damaged my computer and monitor, soaked a roll of souvenir Soviet classroom posters, and bled their ink. Yet, I still have copies of my textbooks.
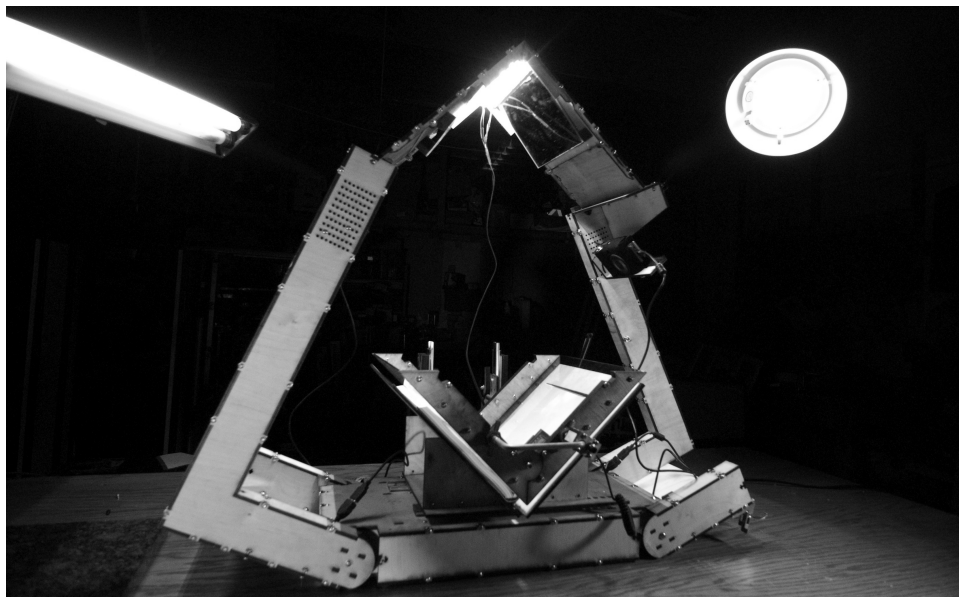


Things become more predictably personal, and more pressing, when we start to consider issues like privacy, which is a necessary condition for free thought, and, by extension, search.[72] Privacy is intimately tied to unpopular, politically unacceptable, and pornographic content, which are, in turn, intertwined with advances in technology. Selecting examples here is like shooting fish in a barrel, from the

---

70.   Historically, such techniques have been used by intelligence agencies to identify the specific typewriter that produced a document.

71.   *Part of Historic NDSU Building Collapses*, Minnesota Public Radio News (Dec. 27, 2009), http://minnesota.publicradio.org/display/web/2009/12/27/building-collapse/.

72.   Robert Darnton, The Case for Books, Past, Present, and Future 46 (2009); Cindy Cohn, *Warrants Required: EFF and Google's Big Disagreement About Google Book Search*, Electronic Frontier Foundation (Aug. 18, 2009), http://www.eff.org/deeplinks/2009/08/warrants-required-big-disagreement-google-book-search; *see generally* James Grimmelmann, *The Google Dilemma*, 53 N.Y.L. Sch. L. Rev. 939, 940 (2008–09).

authority-dodging mobile printing press of the mythical man Martin Marprelate[73] to the DIY book scanners of the Soviet *samizdat*.[74] While others have explained specific privacy issues in depth,[75] the simple act of digitizing one's own collection of unpopular, politically flammable content, or even one's own pornography collection, alleviates the majority of these concerns by the plain virtue that local activity on a local machine is largely undetectable.[76] This stands in stark contrast to activity on a local network, on the Internet, or in the "Cloud,"[77] where each query, including the duration of time spent on any page, image, or passage, and a host of other potentially incriminating or humiliating information is almost certainly being recorded, perhaps under the seemingly harmless auspices of targeted advertising.

Put more bluntly: Are you bold enough to search Google Books for books on Islamic architecture *and* books on airport terminal design? These issues are nothing new. One reason photographic methods were suggested for *samizdat* duplication—or underground duplication—was that the audible clatter of typewriter-copying could alert nosy neighbors through thin walls.[78] The "Cloud" and other service layers on the network offer still more uncertainty. One provision of the proposed Google Books Settlement is that a mere 20% of copyrighted books will be displayed.[79] Taken cynically, this 20% represents an extraordinary opportunity for editing. Imagine such a service being offered in China. Would you expect the 20% displayed for a search of "Tiananmen Square" to be the same as in the Western world?[80] At a minimum, private copies permit comparison and form a kind of distributed ground truth.

---

73. Martin Marprelate was a pseudonym used by the author of heretical tracts. *See* The Martin Marprelate Tracts i (Joseph L. Black ed., 2008).

74. *See* John Dunlop, The New Russian Revolutionaries 52–53 (1976) (discussing K. Glukhov's *Photography as a Method of Multiplying Textual Documentation*, which boasted scanning times comparable to our modern hardware).

75. *See, e.g.*, Cohn, *supra* note 72.

76. Search warrants notwithstanding, of course.

77. The "Cloud" is just another name for an application or service hosted on the Internet in a distributed fashion. The "Cloud" breaks from traditional desktop computing by primarily doing processing and data storage remotely. *See* Walter S. Mossberg, *Learning About Everything Under the 'Cloud*,' Wall St. J., May 6, 2010, http://online.wsj.com/article/SB10001424052748703961104575226194192477512.html.

78. Eliezer L. Ehrmann, Readings in Modern Jewish History, from the American Revolution to the Present 179 (1977).

79. *See* Samuelson, *supra* note 48, at 1310. Google Book Search's "Snippet View" shows information about the book plus some sections of text taken from the book. *Id.* Part of the original settlement agreement is that the total content available from Snippet View would not exceed 20% of the book. *Id. See What you'll see when you search on Google Books*, Google. http://books.google.com/googlebooks/screenshots.html#snippetview (last visited Sept. 12, 2010).

80. To get an idea of the potential contrast, James Grimmelmann's *The Google Dilemma* has screenshots of results for these searches using another Google service—Google Image Search—in both China and the United States. Grimmelmann, *supra* note 72, at 948–49.

At the time of this writing, the MobileRead Wiki lists fifty-seven e-book formats.[81] A substantial portion of these were designed to be read on a particular device, which almost certainly is no longer sold or supported. In other words, when support ceased, or the hardware stopped working, they stopped being accessible. Much like a corrupted iTunes database,[82] a pocket library of Alexandria in any DRM-encumbered format will burn when the authorization servers are switched off—a corporate failure constitutes a technical failure. As Cory Doctorow, Michael Robertson, and others recently noted,[83] for this reason, DRM and TPM[84] destroy the notion of ownership, converting all purchases into rentals. What they said, however, extends to proprietary formats and hardware-specific formats. Though they can occasionally be converted, cracked, or scraped in such a way that they can be used on many hardware platforms, even though such action may be illegal under the Digital Millennium Copyright Act (DMCA),[85] there are no guarantees. No guarantees, that is, if you have not captured the source material yourself with a scanner that has no concept of "digital rights." It is, by definition, easier to move content into a DRM-encumbered format than out of one. Therefore, to keep up with the turnover in consumer electronics, both technically and corporately, you may have to repeatedly re-purchase your books—or simply scan them yourself.

This is only the second time that books have experienced a profound technical shift. While music has shifted from sheet music to wax reels to vinyl to eight-tracks to cassette tape to CD to MP3, and movies from film to tape to beta to VHS to DVD to Blu-Ray to MPEG-4, books have never really been separated from their substrate. Indeed, many of the complaints of book lovers are about this exact shift— the smell, the feel, the look of traditional books causes much gnashing and wailing. Ironically, as these bibliophiles age, they may be forced to rely on the flexible font sizes offered by e-reader hardware. As Mary Murrell so elegantly put it, the e-books offered now are actually "broken books."[86] Strange, atavistic tendencies of publishers and hardware manufacturers, multiplied by what could generously be called a lack of

---

81. *E-Book Formats*, Mobileread Wiki, http://wiki.mobileread.com/wiki/E-book_formats (last visited Sept. 12, 2010).

82. As with many other DRM-protected schemes, it is possible to lose the entire contents of an iTunes library if it becomes corrupted or the underlying hardware changes. *See supra* note 30 and accompanying text.

83. *See* Cory Doctorow, *Digital Rights Management: A Failure in the Developed World, a Danger to the Developing World*, Electronic Frontier Foundation (March 2005), http://www.eff.org/wp/digital-rights-management-failure-developed-world-danger-developing-world.

84. TPM stands for Technological Protection Measures, essentially another, broader, way of saying Digital Rights Management. *See* Gwen Hinze, *Technological Protection Measures in the Draft FTAA,* Electronic Frontier Foundation, http://www.eff.org/pages/technological-protection-measures-draft-ftaa.

85. The DMCA criminalizes the production and dissemination of technology that is intended to circumvent copyright protection. *See* 17 U.S.C. § 101, (2006 & supp. 2010). In certain cases, such as watching DVDs on Linux, this extension of copyright makes it impossible to legally watch a DVD that was legitimately purchased.

86. Videotape: C is for Culture, *supra* note 49.

techno-social imagination, has left us digital bookshelves with a myriad of illegible formats, needlessly restrictive rights management schemes, and absurd limitations.[87] No paper book prevents anyone from writing in it, pressing flowers in it, or passing it on to a friend. No paper book can be revoked or remotely disabled. No paper book needs to be switched off during takeoff and landing. However, no paper book connects to the Internet and, by extension, to social networks; no paper book can be freely and infinitely copied and updated; no paper book can so easily collect and present meta-and-usage data about itself and all the places it has been read from. While the paper book's advantages are largely a result of the physical substrate on which it is printed, the electronic book does not enjoy the full benefit of its ephemeral, networked nature—yet. Furthermore, contrary to the actions of publishers, no pirate would prevent the copying of a book, no pirate would prevent the blind from listening to an audio book, and no pirate would delete your digital book. Scanning your own books ensures that you are not subject to the whims of publishers or the dubious good graces of pirates.

Publishers and manufacturers have invested untold mountains of money and man hours into the restriction of digital content. As a result, pirates have often offered a better consumer experience than corporations. If the same time, money, and energy had been invested in creating the best possible digital book—one which has the irrevocable freedoms of a paper book, plus the incredible advantages of digital—we would be in a much different place. While the spectrum of possible e-book advantages that only a publisher could offer might be limited, a few things come to mind immediately. Publishers could offer books in three (or more) forms: an "Author's Cut," a standard version, and an abridged version. Books could show evidence of the writing process, something previously only shared in interviews.[88] Likewise, they could reveal the hand of the editor, something traditionally invisible to the reader. Footnotes could contain entire other books or huge chunks of them.

While the big players are heavily engaged in legal engineering, developing rights management systems contrary to the inherently copy-able nature of bits, and ignoring the needs of the disabled, individuals can ensure their own access to materials and the access of others simply by building or buying their own digitizing hardware and using it. The very fact that a variety of open source scanner designs are freely available exerts pressure on these big players to improve their product, which benefits all of us.

---

87. An example is the recent decision of the Author's Guild to prevent blind readers from using the text-to-speech features of the Amazon Kindle. The decision is absurd because the Kindle is already effectively inaccessible to the blind due to the screen-based menu system. *See* Michael Kwun, *Does the Authors Guild Want to Sue You for Reading Aloud to Your Kids?*, ELECTRONIC FRONTIER FOUNDATION (Feb. 11, 2009), http://www.eff.org/deeplinks/2009/02/does-authors-guild-want-sue-you-reading-aloud-your.

88. Ben Fry recently produced a remarkable visualization tracing the evolution of Darwin's *Origin of Species* through all six editions. *See* Ben Fry, *On the Origin of Species: The Preservation of Favoured Traces*, http://benfry.com/traces/ (last visited Sept. 12, 2010).

## V.  CONCLUSION

For the second time in history, book technology is undergoing a major change. Movable metal type has given way to malleable digital bits. This sweeping change profoundly affects all of us and so urgently begs for our participation. Leaving the future of books in the hands of a few corporate interests is irresponsible when a clear means of influencing the outcome of this transformation is at hand. To influence the future of digital books, we must take part in creating them ourselves, starting with the books we already have.