Chapter 4

Open Access E-books

E. S. Hellman

Abstract

As e-books emerge into the public consciousness, "open access" (a concept already familiar to scholarly publishers and academic libraries), will play an increasing role for all sorts of publishers and libraries. This chapter of The No Shelf Required Guide to E-book Purchasing *discusses the meaning of open access in the context of e-books, the ways open access e-books can be supported, and the roles that open access e-books will play in libraries and in our society.*

The Open Access "Movement"

Authors write and publish because they want to be read. Many authors also want to earn a living from their writing, but for some, income from publishing is not an important consideration. Some authors, particularly academics, publish because of the status, prestige, and professional advancement that accrue to authors of influential or groundbreaking works of scholarship. Academic publishers have historically taken advantage of these motivations to create journals and monographs consisting largely of works for which they pay minimal royalties, or more commonly, no royalties at all. In return, authors' works receive professional review, editing, and formatting. Works that are accepted get placement in widely circulated journals and monograph catalogs.

In the late 1970s and 1980s, academic libraries became acutely aware that an expansion of research activity had resulted in the growth of both the numbers of journals and the numbers of articles published in the journals. The combination of increased subscription prices and the number of journals needed to support research resulted in a so-called "serials crisis." Libraries were forced to cancel subscriptions. The reduction in circulation forced publishers to raise subscription prices further to make ends meet, and the resulting cycle of cancellations and price increases led to a fear that the whole system would collapse. If few libraries could afford subscriptions, fewer scholars would be able to read the articles, diminishing the attractiveness of publishing.

The advent of Web-based publications in the 1990s led many to believe that the solution to the serials crisis would be a shift of the scholarly publishing industry to so-called "open access" business models. Open access publications are those that can be read at no cost to the reader or the reader's institution. The traditional model of publishing supported by subscription fees was thus styled "toll-access" publishing. It was hoped that the cost reductions from the combination of digital distribution and automation would stop the cycle of rising expenditures.

Perhaps the most successful implementation of open access has been arXiv, a database of digital preprints and reprints ("e-prints") originally focusing on the particle physics community. Started by Paul Ginsparg, a physicist at Los Alamos National Labs, arXiv is now located at Cornell University and hosts more than 700,000 scientific articles in e-print form.[1] Authors deposit articles they've written into the repository, and other scholars are free to search, browse, and download articles without needing any sort of subscription.

arXiv.org e-print archive
http://arxiv.org

One reason for the success of open access archives has been that they have grown up in a parallel

coexistence with the traditional academic journals, which have mostly shifted onto the Web. In the so-called "green" model for open access, many journals allow versions of accepted articles to be made available via repositories. Authors can thus submit their articles to high-prestige subscription-supported journals without worrying about colleagues' access because scholars who need to read their works can always access versions from free sources.

Meanwhile, the shift of traditional journals onto the Web has allowed the rise of secondary distribution channels. Most academic libraries today enjoy access to a much broader range of journals compared to twenty years ago because of the availability of article databases that aggregate content from large numbers of journals.

The past decade has also seen the rise of "gold" open access journals. These journals leverage low-cost Internet distribution to allow articles to be read universally with no subscription charges. These journals cover expenses by charging publication fees to the submitting author (BioMed Central, PLoS) or through private or public funding (SciELO).[2] They build prestige and avoid becoming vanity presses by establishing rigorous review processes.

BioMed Central
www.biomedcentral.com

PLoS
www.plos.org

SciELO
www.scielo.org

What about E-books?

The success of open access journals and articles has for the most part not yet been duplicated in the world of books. There are a number of possible reasons for this. The first is the matter of cost. Publication fees for open access journal articles are in the range of $600 to $3,000;[3] editing and production expenses for a book published by a university press are estimated to be a lot more. While a book that's mostly text might cost as little as $10,000 to produce professionally, a book with figures, photos, equations, and cover art will cost a lot more to edit, design, and produce.[4] Author-funded publication fees this large are unlikely to be practical, even with significant institutional subsidies.

Another factor holding back open access books may be a preference for print books over e-books. Books are much longer than journal articles, and many readers are uncomfortable reading a book on a computer screen. It's only in the past two years that dedicated reader devices such as the Kindle and tablet computers such as the iPad have improved the e-book experience enough to gain wide consumer acceptance.

The business environment for book publishers is another possible factor. The university publisher loses money on much of its catalog but compensates by having one or two titles that cross over to be successful outside the academic environment. Amazon.com has bolstered this pattern by providing wide distribution for small-print-run titles that would never have been available in bookstores before. In contrast, journal articles almost never cross over into nonprofessional markets.

Nonetheless, there have been a few notable attempts to publish open access e-books. I'll cover these later in a section on business models for open access e-books, but it wouldn't be right to omit mention of Project Gutenberg (PG) at this point. PG produced not only the first open access e-books, it produced the first e-books, period. Started by Michael Hart in 1971, PG aimed to take the text of public domain works and make them available via the Internet.[5] To date, PG has put over 36,000 works into its collection, entirely through the efforts of volunteers.[6]

Project Gutenberg
www.gutenberg.org

Distribution of open access e-books can be thought of as an enterprise separate from their production since the costs involved are of a different nature. The scaling laws of Internet distribution favor centralization, and as a result, organizations such as the Internet Archive are able to distribute appropriately licensed e-books on a vast scale; businesses such as Google are able to search and organize them; libraries, blogs, and portal sites are able to select and "curate" them. To some extent, this type of distribution depends on the self-contained nature of the book; it shouldn't require the context of a specific website to retain and accumulate value.

Open access for e-books provides many benefits in addition to allowing people to read for free. Access to the full text of books makes for more complete indexing. The utility of Google Books and the effort Google has put into digitizing books from libraries, even when it is unable to make the books available because of copyright, are testament to the value of indexing the full text. Long-term preservation of our cultural heritage is another public benefit of open access to e-books.

What Does Open Access Mean for E-books?

There are varying definitions for the term *open access,* even for journal articles. For the moment, we'll use

it as a lower-case term broadly to mean any arrangement that allows people to read a book without paying someone for the privilege. At the end of the section, we'll capitalize the term to designate e-books that are freely available in both monetary and practical terms. Although many e-books are available cost-free in violation of copyright laws, they will be excluded from this discussion.

Public Domain

The most important category of open access for books is work that has entered the public domain. In the United States, all works published before 1923 have entered the public domain. Works published in the United States from 1923 to 1963 entered the public domain twenty-eight years after publication unless the copyright registration was renewed. Public domain status depends on national law, and a work may be in the public domain in some countries but not in others. The rules of what is in and out of copyright can be confusing and sometimes almost impossible to determine correctly.

In addition to public domain books that are made available by Project Gutenberg, works digitized by other efforts may be available on an open access basis. It's not true, however, that any digitized public domain book is also open access. That's because the digitizer can use license agreements to restrict access to the works. For example, JSTOR has many digitized public domain works included in its subscription products, but the terms of the subscription prevent republication of its scans. Similarly, Google puts restrictions on the public domain books from partner libraries that it has scanned, digitized, and included in Google Books. While they're available at no cost, there are limits on what you can do with them.

JSTOR
www.jstor.org

The public domain is more than just cost-free; it belongs to everyone. We are free to do with these works what we like. Public domain works can be copied, remixed, altered, or extended. A book publisher can take a public domain text, print up bound volumes, and sell them in bookstores. A movie producer can create a cinematic dramatization of the public domain work; derivative works such as the movie acquire copyrights of their own and are not in the public domain.

Free Copyrighted Content

Laypeople often confuse public domain for "free" (meaning free of cost), and vice versa. Most content available at no cost on the Web is copyrighted, which restricts what people can do with it. Often, the content is made available using an advertising model, trading the opportunity to read and interact with content for the user's attention to ads or links to e-commerce websites. But website users are usually not free to republish content or to e-mail the content to friends beyond the bounds of fair use. They're bound by whatever terms and condition the website chooses to employ; if no explicit terms and conditions are stated, they still can't copy the website's content for other uses.

Even professional publishers are sometimes confused by copyright on the Web. In 2010, the editor of *Cooks Source,* a Massachusetts magazine, got into hot water for republishing a blogger's work without permission. The editor's response to the blogger, on being asked for restitution, made the rounds of the Internet and is striking for the bellicose ignorance it betrays: "But honestly Monica, the web is considered 'public domain' and you should be happy we just didn't 'lift' your whole article and put someone else's name on it!"[7]

Many "free" e-books are available on a similar basis as free websites. They may include advertising or advocacy. Promotional literature and instruction manuals often fall into this category. Many publishers make free e-books available for limited periods of time as a means of marketing them; that doesn't make them free to redistribute, though it happens.

Creative Commons Licensing

Creative Commons (CC) licensing arose to expand the range of creative works available for others to build upon legally and to share. Many authors really want their works to be redistributed for free in venues such as *Cooks Source,* but they want to make sure attribution is given and often want to prevent their work from being altered or chopped into pieces. Others want to make sure that if their work is altered or somehow improved, the altered or improved version will also be available for free. Sometimes, authors are happy to have their works reused noncommercially but want to keep their works from being commercially exploited without permission. CC licenses give authors the tools they need to accomplish these goals.[8]

Creative Commons
http://creativecommons.org

The different licenses available from CC are designated with a special mark, with added code letters that indicate the features invoked by the rights holder. For example, the Attribution-ShareAlike license is denoted by the letters *CC BY-SA* and the mark shown in figure 4.1. This license requires attribution as to the author of the work, and the ShareAlike features bind the

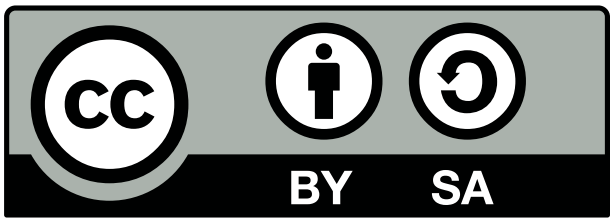

Figure 4.1
The Creative Commons mark indicating an Attribution-ShareAlike license.

licensee to share any modifications or improvements.

It's important to note that in the CC licenses, the owner of the copyright does not give up ownership of the work. Owner may relicense their work under any terms they desire and can still sue people who infringe on the copyrights. The owner licenses the work to the user, who accepts the license as a condition of use. The user can in turn distribute the work along with a copy of the license to other users, who accept the terms of the same license from the copyright owner as a condition of their use.

CC licensing is now widely used for cost-free e-books distributed on the Web. Perhaps the best known e-books using CC are the works of Cory Doctorow, a blogger, science fiction author, and advocate for copyright law reform. It's also used for *Wikipedia* contributions and is supported by Flickr for use in photos.

Copyleft

While CC licenses are the most frequently used for e-books, other licenses can be used to allow for the cost-free reading of books. Noteworthy among these is the GNU Free Documentation License (GFDL), created by the Free Software Foundation to allow software documentation, manuals, and other text to be distributed with strong "copyleft" provisions compatible with the General Public License software they're meant to accompany. The GFDL can easily be applied to e-books; many e-books have been released with this license and with other Free Software Foundation licenses.

Free Software Foundation
www.fsf.org

The idea of copyleft is that licenses can be used to prevent someone from taking from the commons without also giving back. For example, when a book publisher adds commentary and illustrations to the text of a Shakespeare play, the resulting book is covered under copyright, and permission must be given for redistribution even though the underlying work is in the public domain. This would not be allowed by a copyleft license, which requires that any derivative works be made available under the same license. The CC ShareAlike licenses have weak copyleft; the GFDL is stronger, and even forbids the use of digital rights management (DRM). For example, it would not be legal to distribute without permission from the author a GFDL e-book to a Kindle e-reading device in the encrypted format usually used by Amazon.

Open Access versus open access

Consider the book *How Wikipedia Works* by Phoebe Ayers, Charles Matthews, and Ben Yates. Is it an open access e-book? What about *What Matters Now* by Seth Godin? Or how about the author's PhD dissertation, posted on his personal website? If you go by the licensing, you might say the first is "free," the second might be, and the third isn't. *How Wikipedia Works* comes with a GFDL license; it's clear what you can do with it. *What Matters Now* doesn't come with any license except for the statement "feel free to share this"; it's still covered by copyright law. The third has the copyright statement "All rights reserved" and is registered with the copyright office.

The practicalities are quite different. If you search for *How Wikipedia Works* on Google, you'll have to dig quite a bit to get a free e-book. Amazon will sell you the Kindle version for $14.37.[9] You can buy it in three different formats from O'Reilly or from No Starch Press, the publisher, for $23.95.[10] Google Books has the book through its publisher program; it appears to fully available, and Google doesn't try to sell it to you.[11] You can find the e-book in a library through WorldCat, but the libraries that hold it restrict access to their own users. *Wikipedia* itself has a page for it with links to a PDF and the Google Books version. [12] *What Matters Now* and the author's dissertation can both be found on both Google and WorldCat, but the listing for *What Matters Now* in WorldCat covers only a print version apparently published through Lulu.com.[13] The author's dissertation is not only accessible through WorldCat but has been made available as a PDF download by the Stanford University Library.

The lesson of these examples is that for an e-book to be both cost-free and effectively available, there should be an intent by the publisher to make the e-book openly available, expressed with an appropriate license, and there must be effective distribution. This combination is what makes an e-book Open Access with capital *O* and *A*.

Business Models for Creation of Open Access E-books

Any model for e-book publishing must include a business model for recouping the expenses of production:

reviewing, editing, formatting, design, etc. In this section, I'll review methods that can be used to support Open Access e-book publishing.

In 2009 Cory Doctorow put together a collection of short stories called *With a Little Help* and documented the process of publishing it in a series of columns in *Publishers Weekly*.[14] He used a variety of business models to support the project, as detailed below, and the e-book version was released under a CC license.

DIY Publishing Models

One way to meet the costs of e-book production is to keep those costs close to zero. Free blogging sites have made it simple for authors to produce blogs and other sorts of websites; additional tools are available to add keywords, links, and images. Other tools can convert a blog or similar website to the EPUB e-book format; EPUB export is available in Apple's Pages word processor, and it's likely that other programs will soon follow suit.

Given these tools, authors can produce e-books on their own, with no other expense than the value of their time. For *With a Little Help,* Doctorow did most of the production himself; as the title suggests, he got friends to help out with things such as cover and book design.

In the do-it-yourself, or DIY, model, there are essentially no expenses to recoup. If the author wants to earn something, additional money needs to be spent on an ISBN and a bit more to get metadata into a feed for Amazon. But if income is not the object, the e-book can simply be posted on a website and made available to the world. A CC license allows the e-books to be distributed in a wide variety of channels.

In fact, with the consent of the editor, this chapter will be released as a DIY Open Access e-book in EPUB format, with a CC BY-ND license. The author hopes to profit primarily from the experience of doing so.

Freemium Models

Freemium refers to the business model, common on websites, of offering one level of service for free, and then, when the user is solidly hooked on the use of the service, offering a premium level of service for a fee. The difficulty of this model is to have a service that's attractive enough at the free level to drive premium conversions, and at the same time limited enough at the free level that upgrades deliver significant value.

In the e-book space, the traditional premium service is typically either the print version or an updated or otherwise enhanced digital edition. O'Reilly has used this model to great effect, by allowing authors to make free PDF versions available on websites while O'Reilly sells print versions through traditional channels.

In Doctorow's project, he offered print-on-demand versions through Lulu.com for $18 each, along with 250 "super-limited hardcovers" for $275 each: These were hand-bound on acid-free paper and included original paper "ephemera" and came with a memory card with the full text of the book and audiobook. The $275 version turned out to be the big moneymaker.[15]

As e-book readers become preferred over print by users, using print as a revenue engine may run out of steam. Bloomsbury Academic is building a platform that also uses e-book versions as the premium. While CC noncommercial versions are available for reading online, the books will also be issued for purchase in print and on Kindle and Sony readers. It's possible that publishers will look at enhancing e-books with supplementary content or deep semantic markup as their revenue driver; a bare-bones Open Access version would serve as promotional vehicles for the core product.

Bloomsbury Academic
www.bloomsburyacademic.com

Advertising and Promotional Models

Cost-free and Open Access content can promote more than just a premium edition of the same content. E-book formats are much like HTML websites in that they can embed links; even JavaScript functionality is becoming available in e-book content. Publishers can use these types of functionality to generate revenue through advertising. A quick look at iPad or Android app stores reveals a huge selection of free, advertising-supported apps, including many apps that simply wrap e-book content.

In one scenario, an author of a book series might produce an Open Access electronic version of the first in the series. The free e-book could have embedded links or "in-app purchase" buttons for subsequent books in the series. Open Access e-books might also be supported by contextual links or product placement; imagine a story featuring a sports car where the brand and model of the car are chosen based on support from a car company.

Another type of promotion that can be furthered by all types of free e-books is personal brand building. It could be argued that Cory Doctorow's biggest payoff from the *With a Little Help* project was that it increased his fame and thus his ability to make money on appearances, commissions, and the Boing Boing website. (One story in the collection was a $10,000 commission.[16]) Seth Godin, author of the above-mentioned *What Matters Now,* is another personal brand builder. *What Matters Now* is a compilation of inspirational snippets from seventy "big thinkers" that loudly proclaims "share me."

Public Funding

Some books, such as those relating to education, public health, political or social advocacy, or scientific research, fulfill a public purpose. Publication of these

books using a form of Open Access will further their public purpose. The costs of production and release of these books can be financed by foundations, charities, political action committees, private individuals, or governments.

European governments have joined together to fund the digitization and distribution of cultural heritage works through Europeana. Funded by the European Commission and national ministries of culture, Europeana acts as a portal enabling distribution of large numbers of Open Access e-books. In the United States, books created by the federal government belong by law to the public domain, but there's no centralized funding of Open Access e-books or their distribution.

Europeana
www.europeana.eu/portal

In developing countries, governments seeking to provide textbooks to large numbers of students will eventually find that producing e-textbooks, released for free, is the only scalable method of providing for their national educational needs. Many states in India, for example, already release their state-published textbooks on an Open Access basis.

A variation on public funding for Open Access e-books in the context of academic monograph publishing has been proposed by Frances Pinter. Her idea is for libraries to join together in a cooperative, diverting a fraction of their acquisition budgets to fund the fixed costs of producing new monographs by university and commercial scholarly presses; the monographs would then be made Open Access. She estimates that individual libraries could save over 75 percent, depending on the participation rate.[17]

Another sort of public funding model with a long history of use is the "tip-jar", or more profitably, the pay-what-you want model. Here, the creator urges his audience to leave some money as a "thank you" in return for value received. Doctorow has reported receiving over $2,700 using a PayPal-powered donation box, much better than his print-on-demand offering.[18]

Crowdsourcing

Wikipedia and the more specialized wiki sites it has spawned are excellent examples of Internet resources created by large numbers of individuals working together virtually. These volunteer collaborations have replaced printed encyclopedias for most people and might be considered to be the largest, most dynamic Open Access e-books in existence. Most users wouldn't consider these websites to be books, even though the printed equivalents certainly were.

An organization called Distributed Proofreaders is an aggregation of volunteer effort clearly focused on e-books. Many of the digital texts in Project Gutenberg have been produced by Distributed Proofreaders volunteers, who check and correct OCR (optical character recognition) transcriptions of scanned books. While OCR can be very accurate for modern books, books and magazines printed in the nineteenth century and earlier present a variety of challenges. The resulting digitized works are dedicated to the public domain.

Distributed Proofreaders
www.pgdp.net/c/

Crowdfunding

The model that the author is working on at Gluejar, Inc. is crowdfunding. It's analogous to the method by which public radio and public television stations are funded in the United States, except that every book that's to be released with a CC license has a fund drive of its own. Once the producer's price has been matched by reader pledges, an Open Access e-book is released. The pledge drives are managed by a website.

Gluejar
www.gluejar.com

Authors have used crowdfunding websites such as kickstarter.com to cover the expenses of completing a new book. For example, Mur Lafferty raised over $19,000 from more than 250 backers to fund book design, cover design, and e-book conversion for a fantasy audio series.[19] In a few cases, the projects use CC licenses. Stephen Duncombe, a professor at NYU, raised $4,350, more than his goal of $3,500, from 111 backers to fund the further production of an open-source version of Sir Thomas More's *Utopia,* which is distributed with a CC BY-SA license.[20] (Of course the underlying work is in the public domain, but the new translations, annotations, and commentary are subject to copyright.)

Kickstarter
www.kickstarter.com

The Open Utopia
http://theopenutopia.org

To get a better idea of how crowdfunding might scale to large numbers of books, consider the author of a romance series. Rights for the earliest books in

the series have reverted to her, but there's no cash to convert the book to e-book formats. She contacts the pledge-drive website and enters an offer to release the first book under a CC license in exchange for a lump sum payment that she considers to be fair and that covers the conversion to e-book. Fans of the series can then go to the site and pledge support. If the author's offer price is met, supporters get billed, and the author gets the payment. The resulting e-book file is sent to all the people who have pledged and put on a feed for the rest of the world to pick up. Since the e-book is now CC licensed, it can be redistributed for free.

In another scenario, a reader, perhaps someone who has found the book in a library, launches the pledge campaign. The library metadata is pushed to the pledge-drive site, and other fans can pledge their support. Eventually, the pledge amount gets big enough to attract notice from rights holders, who can then show up, deliver the e-book, and take the cash off the table and divide it among themselves.

Libraries and Open Access E-books

One of the missions of libraries is to provide access to all sorts of information, including e-books. So if an e-book is already open access, what role should libraries play?

Here's a thought experiment for libraries: imagine that the library's entire collection is digital. Should it include Shakespeare? Should it include *Moby-Dick*? These works are available as public domain works from Project Gutenberg; providing them in a library collection might seem to be superfluous. Many librarians have been trying to convince their patrons that "free stuff on the Internet" is often inferior in quality to the information available through libraries. There are certainly e-book editions of these works available for purchase with better illustrations, editing, annotations, etc. Should libraries try to steer patrons to using these resources instead of using the free stuff?

For the most part, libraries have not done a good job of incorporating resources such as those available from Project Gutenberg into their digital collections. OverDrive, the leading provider of e-books to public libraries, now offers Project Gutenberg titles for no extra charge, but they are offered as a separate collection. At present, if a user searches for *Moby-Dick* in a library collection, a result will be returned only if the library has a purchased edition of *Moby-Dick,* which may be in use by another patron. A separate search must be done to retrieve the free edition.

As we saw in the section on types of open access, for an e-book to really be Open Access, there must be an appropriate license (or public domain status) and effective access. There are a number of ways that libraries can work, both individually and through cooperative effort, to make that access effective. Similarly, Open Access e-books can play an important role in supporting the mission of libraries. This section will consider libraries' roles in access, selection, archiving, community, and production of Open Access e-books.

Access and Storage

Most libraries can avoid worrying about access and storage of Open Access e-books, thanks to services such as the Internet Archive's Open Library project and HathiTrust, a "partnership of major research institutions and libraries working to ensure that the cultural record is preserved and accessible long into the future."[21] These services provide reliable low-cost file storage and bandwidth. Adding effective access to cost-free e-books at other sites may need a bit more work; figuring out and tracking stable, persistent URLs at multiple locations can create a logistics burden for libraries that could help manage access. Library-oriented "knowledge base" services from vendors such as OCLC, ProQuest, and Ex Libris may prove to be useful in this regard.

Open Library
http://openlibrary.org

HathiTrust
www.hathitrust.org

As users shift towards reader devices and tablet computers, libraries will find themselves spending a lot of time helping users figure out how to move Open Access e-books onto their devices. In principle, Open Access e-books shouldn't require DRM and should thus be compatible with most devices. In practice, getting content free content onto a device can be nonintuitive and often "side-loading" or other indirect procedures are required; most e-reader devices have book shopping functionality, and the vendors are not motivated to push users to content that doesn't generate revenue.

Selection and Description

Metadata-based discovery and browsing have been a strength of libraries; without the motivation to sell copies, many cost-free e-books lack even basic metadata, let alone good quality catalog records. This is clearly an area where libraries can make significant contributions, especially when they work cooperatively.

With a flood of free content already available and much more on the way, there is a continuing need to highlight the material most suited to the needs of the user. Multiple editions can exist of public domain works; it makes sense for libraries to help patrons find the best ones.

Perhaps the best example to date of work in a library on selection and description of Open Access e-books is the Online Books Page at the University of Pennsylvania. Edited by John Mark Ockerbloom, it indexes over a million online books, all of them available for free to users.

The Online Books Page
http://onlinebooks.library.upenn.edu

Archiving and Preservation

One of the biggest uncertainties presented by e-book licensing is whether today's e-book acquisitions will meet the needs of future readers. As we celebrate the fortieth anniversary of the first e-books, it's hard to ignore the fact that most libraries have print collections that reach back a hundred years and more. We don't know what parts of today's written culture will be in demand one hundred (or even forty) years from now or how readers will expect to approach them. For that reason, texts must be in a form that can evolve with reading technology, and the evolution must not depend on the permission and continued existence of publishing companies, platform vendors, rights management software, proprietary software, or hardware. Formats must adhere strictly to standards. The forty-year-old texts from Project Gutenberg can still be read today because they used very simple formats; these are being converted to newer, more capable formats such as EPUB for easy consumption on e-book readers. Going forward, there will be continuing challenges in the evolution of photos, graphics, mathematics, scripting, and linking of e-books.

Another concern for the future is the potential loss of material. The threats facing e-book survival include media degradation, bit rot, cosmic rays, natural disasters, wars, bankruptcies, legal actions, societal breakdown, and human stupidity. Many preservation specialists believe that making many dispersed copies of material is the most robust way to ensure long term; the more important the material, the more copies should be made. This is the philosophy behind LOCKSS, a peer-to-peer preservation system in which libraries are taking the lead in preserving e-journals and other websites. LOCKSS has been working to extend its digital preservation efforts to e-books; about 45,000 e-books are "in-process," and it's expected that another 30,000 will be added in 2012.[22]

LOCKSS
www.lockss.org

Community and Context

Open Access e-books give libraries new ways to reach out to the communities they serve. The social aspects of reading are well known to libraries; the story times and book clubs nurtured by public libraries are excellent examples. Although an e-book isn't tied to location the way a print book is, people and their social circles are tied to places. There are two types of advantages for the use of Open Access e-books in a library's outreach efforts. Cost is an obvious factor; public libraries have an obligation to support reading by community segments that might not be able to afford the books they need. A second advantage is that of context building. The sort of annotation, commenting, and discussion around books that can take place in a group of friends and neighbors is quite different from that which occurs anonymously in a global forum. At the same time, the availability of free, untethered e-books from libraries, free from DRM or Internet monitoring, allows individuals to obtain and read books with real privacy.

Participation

As technology lowers the barriers to e-book production, more and more people will be able to produce and distribute e-books. Just as the combination of YouTube, cheap video cameras, and editing software allows Rebecca Black to become a viral sensation, the corresponding e-book technologies are already starting to nurture grassroots authorship. Libraries may play an important role in enabling and promoting community-created content. Books that may not be commercially viable may still be important to a community, and libraries can play a role in connecting local authors to communities both near and far.

Libraries can also fill the need for educating grassroots authors about the meaning and importance of public licenses. Some authors will of course need to use traditional licensing strategies, but most will be unfamiliar with CC and other types of licenses. The social benefit of the use of these licenses is aligned with the library's mission of promoting access to information, and libraries should not be hesitant to promote their use.

Changing the World

As applied to the scholarly journal, the goals of the Open Access movement have been diverse. The success of the movement must be judged against those goals. There's no doubt that Open Access has been successful at its core goal of increasing access to many types of information. But some other hopes pinned on the movement have been unrealized. Serials budgets at libraries have continued a seemingly inexorable rise.

It's been estimated that 4 billion books are printed

each year.[23] That seems like a big number until you remember that the world's population is almost 7 billion. A large fraction of the world's population has minimal access to books. Yet the number of cell phones in the world has been estimated at 4.6 billion.[24] As more and more cell phones become capable of delivering e-books, the fraction of the world's population with access to e-books may soon exceed the fraction of the world's population with access to physical libraries.

The majority of the people in the world will not be able to pay $9.99 for an e-book. Even in wealthy countries, the cost of food, clothing, shelter, transportation, and medical care limit many people's ability to buy content licenses. Yet the thirst for literature, learning, and culture is not confined to the wealthy of the world. Open Access e-books can help to slake this thirst and help to create a global community of understanding and knowledge. Through shared access to culture and ideas, Open Access e-books can erase some of what separates the nations of the world, rich and poor.

For Open Access e-books to have this sort of impact, their production and distribution must be effective. Production can occur through a variety of business models, including models that reward authors and creators for their efforts. New distribution channels must be created and supported. Libraries have a clear and vital role in this process and must work cooperatively to meet the needs of their diverse communities. Venues for such cooperation already exist (OCLC, Open Library, HathiTrust, Europeana, and various national libraries) or are being planned (the Digital Public Library of America), but new ones will also be needed.

Together, we must strive to make sure that the best and most thoughtful of the world's e-books are not lost in a deluge of free drivel, free come-ons, and free polemics. If people are to govern themselves in peace, they should have easy access to good ideas and honest information.

Digital Public Library of America planning Wiki
http://cyber.law.harvard.edu/dpla/Main_Page

Notes

1. arXiv.org e-print archive, http://arxiv.org (accessed Oct. 13, 2011).
2. BioMed Central is part of Springer Science + Business Media, a leading global publisher in the STM sector; PLoS (Public Library of Science) is a nonprofit organization that publishes open access journals including *PLoS Medicine, PLoS Biology,* and *PLoS ONE*; SciELO (Scientific Electronic Library Online) is a multinational endeavor to provide open access to scientific research produced in Latin America and Spain.
3. Several journal submission sites were surveyed for this article.
4. The $10,000 figure was obtained by analyzing the operating budgets for independent university presses as discussed in this blog post: Eric Hellman, "A Library Monopsony for Monographic eBook Acquisition?" *Go to Hellman* (blog), Aug. 10, 2010, http://go-to-hellman.blogspot.com/2010/08/library-monopsony-for-monographic-ebook.html. Numbers were confirmed privately.
5. Project Gutenberg, "Obituary for Michael Stern Hart," last updated Oct. 8, 2011, www.gutenberg.org/wiki/Michael_S._Hart.
6. Project Gutenberg home page, www.gutenberg.org (accessed Oct. 13, 2011).
7. Mary Elizabeth Williams, "*Cooks Source:* The Internet Roasts a Plagiarist," *Internet Culture* (blog), Salon.com, Nov. 5, 2010, http://life.salon.com/2010/11/05/cooks_source_internet_revenge.
8. Descriptions of all of the Creative Commons licenses are available at http://creativecommons.org/licenses.
9. Amazon.com, "How Wikipedia Works [Kindle Edition]," www.amazon.com/How-Wikipedia-Works-ebook/dp/B002MZAR68/ref=sr_1_1?s=digital-text&ie=UTF8&qid=1318339355&sr=1-1 (accessed Oct. 11, 2011).
10. O'Reilly Media, "How Wikipedia Works," http://shop.oreilly.com/product/9781593271763.do (accessed Oct. 11, 2011); No Starch Press, "How Wikipedia Works," http://nostarch.com/wikipedia.htm (accessed Oct. 11, 2011).
11. Google Books, "How Wikipedia Works," http://books.google.com/books?id=lHdi1CEPLb4C&lpg=PP1&ots=Ffjr0nFODq&dq=How%20Wikipedia%20Works%20ayers&pg=PP1#v=onepage&q&f=true (accessed Oct. 11, 2011).
12. "How Wikipedia Works," *Wikipedia,* http://en.wikipedia.org/wiki/How_Wikipedia_Works (accessed Oct. 11, 2011); *How Wikipedia Works* is available (GFDL license) as a PDF file at http://dl.dropbox.com/u/10330141/How%20Wikipedia%20Works.pdf, and it's listed on a GNU webpage at www.gnu.org/doc/other-free-books.html.
13. Seth Godin's *What Matters Now* is available for download at "What Matters Now: Get the Free Ebook," *Seth Godin's Blog,* Dec. 14, 2009, http://sethgodin.typepad.com/seths_blog/2009/12/what-matters-now-get-the-free-ebook.html.
14. Cory Doctorow's *With a Little Help* project is documented on the *Publishers Weekly* website. The first column is Cory Doctorow, "Doctorow's Project: With A Little Help," Oct. 19, 2009, *Publishers Weekly* website, www.publishersweekly.com/pw/by-topic/columns-and-blogs/cory-doctorow/article/15883-doctorow-s-project-with-a-little-help.html.
15. Cory Doctorow, "With A Little Help: The Early Returns," Feb. 14, 2011, *Publishers Weekly* website, www.publishersweekly.com/pw/by-topic/columns-and-blogs/cory-doctorow/article/46105-with-a-little-help-the-early-returns.html.
16. Doctorow, "Doctorow's Project."
17. Frances Pinter, "Rethinking the Role and Funding of Academic Book Publishing" (keynote presentation, O'Reilly Tools of Change for Publishing conference, New York, Feb. 24, 2010), www.toccon.com/toc2010/public/schedule/detail/11833.

18. Cory Doctorow, "Financial Report, September 28, 2011," With a Little Help website, http://craphound.com/walh/2011/09/financial-report-september-28-2011.
19. Mur Lafferty, "The Afterlife Series: Heaven, Hell, Earth, Wasteland, War," Kickstarter, www.kickstarter.com/projects/869477073/the-afterlife-series-heaven-hell-earth-wasteland-w (accessed Oct. 11, 2011).
20. Stephen Duncombe, "The Open Utopia: A New Kind of Old Book," Kickstarter, www.kickstarter.com/projects/1713881779/the-open-utopia-a-new-kind-of-old-book (accessed Oct. 11, 2011).
21. HathiTrust Digital Library, "Welcome to the Shared Digital Future," www.hathitrust.org/about (accessed Oct. 11, 2011).
22. Victoria Reich, private communication, April 19, 2011.
23. Peter Lyman and Hal R. Varian, How Much Information? 2003 website, www.sims.berkeley.edu/how-much-info-2003 (accessed Oct. 13. 2011).
24. CBS News, "Number of Cell Phones Worldwide Hits 4.6B," Feb. 18, 2010, www.cbsnews.com/stories/2010/02/15/business/main6209772.shtml.