FuturePress

Fred Chasen
Jake Hartnell
AJ Renold

Advisor: Robert Glushko

May 2014

Abstract

This paper offers a vision of digital book distribution based on open web standards, shows solutions implementing these standards, and discusses the possibilities such technological stacks can enable. The World Wide Web Consortium (W3C) has championed open standards to create a content publishing system that is extensible, configurable, and compatible across browsers, applications and devices. We argue digital books should be built with and distributed by these same open standards. We discuss how “books as web resources” will bring a number of efficiencies, new affordances, and benefits for multiple stakeholders in the information economy. The benefits discussed include open annotation, linking inside of books, interoperability, decentralized distribution, and syndication. We present a technology stack built on these standards, and offer a use case showing how this architecture can be used to support new ways of creating, distributing, and monetizing content.

Table of Contents

Introduction

A book used to be a physical object, a static and relatively unchanging container for knowledge, but the "book" has changed remarkably in the age of the Internet. When we think of books what often comes to mind is the idea of the novel, but books serve many informational purposes. The same information that was once printed in books has been largely transformed into digital information systems we don't often call "books."[1] Far from the romantic view of the book as the ideal format for all knowledge, the physical book is not always the best way of organizing and presenting information, as Peter Brantley points out in his essay *The Curation of Obscurity*:

> It is true that we are surrounded by books that, by their nature, should always have been digital and never books at all. We froze them into a physical form because that was all we could do with tabular data and rich information. Atlases made rigid as oversized picture books, compendia of various facts and speculations printed as so many beautifully designed encyclopedias. Cookbooks and phonebooks presented as a series of manually navigable facets: soups, vegetables, meats, and desserts on one hand; an alpha-sort order listing by name, and type-of-business on the other.[2]

The choices that technologists, content creators, consumers, and distributors made in the creation of the current systems have had important implications in the realms of intellectual property, economics, and culture. However the focus of this paper is not to give a complete picture of the publishing industry as it currently operates, but to offer a vision of digital book distribution based on open web standards, solutions implementing these standards, and the possibilities they can enable.

Open Standards

The World Wide Web Consortium (W3C) has championed open standards to create a content publishing system that is extensible, configurable, and compatible across different browser applications and devices. They have an open and inclusive process to develop new standards and refine existing ones through topic based working groups. Digital books should be based on, and distributed by, these same open standards and thus remain a vital part of the web platform.

The keystone of these standards is hypertext markup language (HTML) now in it's fifth revision. It defines markup elements that give structure to content. These included headers, paragraphs and tables. It also specifies how to include images, audio, and video that can be displayed in any implementing web browser. HTML creates the mechanisms for linking together pages on the internet, creating a web of pages that have come to define what the internet means to most users.

---

[1] Savikas, A. (2012). Distribution Everywhere. In H. McGuire & B. O'Leary (Eds.), Book: A Futurist's Manifesto. O'Reilly Media. Retrieved from http://book.pressbooks.com/chapter/distribution-everywhere-andrew-savikas

[2] Brantley, P. (2012). The Curation of Obscurity. In H. McGuire & B. O'Leary (Eds.), Book: A Futurist's Manifesto. O'Reilly Media. Retrieved from http://book.pressbooks.com/chapter/curation-of-obscurity-peter-brantley

Cascading Style Sheets (CSS) define how those markup elements will be displayed, separately from the elements content. It allows detailed descriptions of how text will look—similar to the tools a print editor would layout a physical book with. The text capabilities include defining color, font families, weight, size, justification, letter spacing and line height. CSS also defines advanced layout capabilities, such as grids, columns and flowing text around images. Finally it can adjust how elements are styled to different media types, such as smaller screen or printed pages.

User interaction with these elements are programmed through Javascript (or its standards name ECMAScript). HTML elements have a programmatic interface that allows their styles, events and content to be controlled through methods exposed by a browser to its Javascript interpreter. Javascript can also request content from other sources, such as a database server and load it into a webpage.

Epub 3

The Epub specification, maintained by the International Digital Publishing Forum (IDPF) is a popular open digital book format based on web platform technologies.

> "EPUB is the distribution and interchange format standard for digital publications and documents based on Web Standards. EPUB defines a means of representing, packaging and encoding structured and semantically enhanced Web content — including XHTML, CSS, SVG, images, and other resources — for distribution in a single-file format."[3]

In the third version of the Epub specification, content is created in HTML5, styled with CSS and interacted with through Javascript. In addition to these, it defines the Open Packaging Format (OPF), which adds a Extensible Markup Language (XML) manifest of all the files needed for the book, book specific metadata, and a spine which gives the linked HTML chapters an ordering.

Additionally, the Epub format standardizes linking in ebooks through the use of Canonical Fragment Identifiers (CFI). These allow linking down to the character level, as well as linking to ranges of text and fuzzy matching to text that has changed. Epub also defines a Page list, which contains a reference to the location of physical printed pages to their counterparts in a reflowable digital text. Finally, Epub defines a specific way to package all this information, using Zip compression, for single-file distribution.

OPDS Catalogs

The Open Publication Distribution System (OPDS) is an open specification for the syndication of digital content based on Atom, an XML format for publishing a feed of content on the web which anyone can subscribe to for updates. "OPDS Catalogs enable the aggregation, distribution, and discovery of books, journals, and other digital content by any user, from any source, in any

[3] http://idpf.org/epub

electronic format, on any device."[4] Publishing an OPDS catalog allows a bookstore, library, publisher, or author to maintain an online list of all the book titles they have available, with descriptions, metadata, and links directly to the content in different formats or to a place to purchase or borrow them.

The Internet Archive's Open Library and Project Gutenberg list huge libraries of free content through the format, and publishers such as Feedbooks, Smashwords, and O'Reilly offer newer content for purchase.

Open Annotation

With longer web documents comes the need for more advanced forms of linking. What if you want to talk about a specific sentence? Or a specific comma that's wrong? Or a specific part of an image? The Open Annotation data model and the many projects implementing it have provided this.[5][6]

Open Annotation for note taking, would mean that notes would no longer be locked in a platform or trapped within the file format (as they currently are with PDF). Notes themselves would be web resources that can be used in new contexts such as writing papers. The annotations themselves contain all the appropriate metadata for robust citation, and provide a direct link to the resource itself so that a reader would be taken to the exact context in which something was said or written.

Open Annotation could also allow for better collaboration. Be it researchers reading scientific papers together, or an office that is reading a book on marketing together and having a discussion related to their specific clients on top of it. Annotation provides many useful tools for collaboration when part of an open platform.

Why Use Open Standards?

Interoperability

Open standards promote interoperability of books through standardizing the set of features an implementer must support through an open and community driven process. Using the Epub format is that purchased ebook content will work with any implementing reader, be it an e-ink device, tablet, phone, desktop app, or browser.

However several large ebook distributors have chosen to use their own proprietary formats instead (as seen in Table 1), inhibiting content purchased from them being used on devices or applications they choose not to support. Each of these formats have different capabilities and limitations, causing issues with consistency for cross-platform ebooks. This has lead publishers

[4] http://opds-spec.org/about/
[5] http://www.openannotation.org/spec/core/
[6] http://annotatorjs.org/showcase.html

to create ebooks that cater to the lowest common functionality between the systems, supporting only the most basic text markup, styling, navigation, and linking.

Table 1: Comparison of Supported Ebook Formats[7]

Bookstore	Formats	DRM
Amazon Kindle	KF8 and Mobi	Amazon-specific
Barnes & Noble NOOK	Epub 2	Adobe ADEPT
Apple iBookstore	iBooks Author	Apple Fairplay
Kobo	Epub 3	Adobe ADEPT
Google Play Books	Epub 3	Adobe ADEPT

Lack of interoperability also creates issues for consumers. Their purchased content is locked to a particular platform, making them unable to read that content on other devices they own or with applications. Consumers with disabilities, such as those who suffer from visual impairment, are not able to use reading applications specifically designed to overcome those difficulties.

Access Controls

Many proprietary ebook formats are based on open standards, but add features that are thought to be necessary for a company's legal and business practices. An iBook from Apple is an Epub 3 document secured with a Apple's FairPlay DRM (Digital Rights Management). Kindle's .mobi file format extends Epub's predecessor Open Ebook (OEB) with custom HTML elements and wrapped with DRM tied to the device it was purchased on. This protection limits the number of devices a book can be read on by capping the number of IDs it can be associated with. Amazon's newer format, KF8, is based on Epub 3 but "still carries forward some HTML constructions that have been de-standardized for years and left over from MOBI's use."[8]

DRM systems replicate the affordances of physical goods, such as scarcity, and difficulty to copy while reducing the affordances of digital books. While the intention of DRM is to limit unauthorized access, there are adverse effects for authorized usage such as removing the ability for users to share, loan or resell the digital files they own. These important interactions with books have more limitations in their protected digital versions than they do in their physical versions. DRM systems limit the portability of digital books between reading systems or applications. This can be detrimental to the reading activity, as different reading systems can

[7] Data adapted from "eBook Retailers" eBook Architects https://ebookarchitects.com/learn-about-ebooks/retailers/

[8] http://wiki.mobileread.com/wiki/KF8#The_Format

offer a range of features, but digital books are rarely portable between the systems.[9] Finally, DRM can easily be stripped from ebooks and there are many instructional sites and tools available online for this purpose. Additionally, once stripped of DRM, the book can be listed on torrent websites for the easy downloading and reading of those who may be less technically inclined.

As Lawrence Lessig argues in his book *Remix: Making Art and Commerce Thrive in the Hybrid Economy*, "Access is the mantra of the YouTube generation. Not necessarily free access. Access."[10] People more than ever have an expectation of content that works across devices, that is linkable, that is *on the web*. One of the most successful strategies in combating piracy has been making content easy to access and purchase digitally.

Open standards on the Internet can better suit the needs of publishers then current DRM systems. Rather than controlling *usage*, publishers should focus on controlling *access*. Web servers have developed sophisticated systems at determining whether or not someone has access to a certain web resource. OAuth is openly developed standard for token based authorization, it enables applications to request fine-grained permission to access to a web service. For instance Facebook uses OAuth 2.0 with their Graph API to grant access to a third party to view, post or edit items on a Facebooks users activity feed depending on the permission granted to the party by a user. When books are web resources, rather than a single downloadable file, permission to view specific chapters of them can be granted and revoked using a similar mechanism.

Economics

Open standards will allow for new economic models for content publishers and improve competition, hopefully leading the information economy to a state that looks more like a Pareto-optimal distribution system: a system in which more people have a higher payoff. There is much work being done on digital payment systems and standards, which are beyond the scope of this paper, but, coupled with open standards for distribution and documents could lead to a great amount of experimentation as micropayments become a reality.[11] In addition, there are also benefits and savings to be gained from interoperability and simplified distribution.

Open standards would allow content creators to be the distributors of their electronic content. Currently, distributors take a cut for delivering content from their servers to consumers devices. (Table 2 shows what the current distribution fees and market shares of the different distributors are). Not only that, but they collect extensive analytics on the consumer's reading experience which are not made available to content makers. Open standards could allow content makers to own the methods of distribution, notably bypassing the commission that would have to be paid by a traditional distributor. This would also create more competition in the market. Rather than

[9] This also creates large barriers to entry for anyone wanting to make a better reading application.
[10] Lessig, L. (2008). *Remix*. Blackwell Publishing.
[11] http://www.w3.org/community/webpayments/

having the market controlled by a small number of distributors, anyone with a website and content could be in competition with other content makers.

Table 2: Comparison of Distribution Fees[12]

Bookstore	Market Share	Distribution Fee
Amazon Kindle	60%-70%	30% for book priced $2.99 - $9.99 plus delivery charge of $0.15/MB. 65% for other prices.
Barnes & Noble NOOK	15%-25%	35% for books priced $2.99 - $9.99. 40% for other prices.
Apple iBookstore	10%-20%	30%
Kobo	<5%	35% for books priced $1.99 - $12.99, 55% for other prices.
Google Play Books	<5%	48%

Books as Web Resources (Books unbound)

We propose that books are treated as web resources, which are by their nature cross-platform and interoperable. We think a web-based solution—separating content from functionality and following the lead of W3C, which is pushing standards in this direction—is superior to solutions that are intrinsically closed and don't support the collaboration that we believe is essential for the next generation of publishing. The Internet offers a low-barrier distribution mechanism for information, as can be seen by the over 900 million websites that are online today.[13]

Linking

As a web resource an Epub file has the potential to be more than a container. Because the Epub format is based on established web standards, book content can be addressable through universal resource identifiers (URLs) creating numerous advantages over a physical book.

[12] Data adapted from "eBook Retailers" eBook Architects https://ebookarchitects.com/learn-about-ebooks/retailers/

[13] 958,919,789 reported by Netcraft in April 2014 - http://news.netcraft.com/archives/2014/04/02/april-2014-web-server-survey.html

> "The Web has proven that the concept of hyperlinking is tremendously powerful, but EPUB Publications have been denied much of the benefit that hyperlinking makes possible because of the lack of a standardized scheme to link into them. Although proprietary schemes have been developed and implemented for individual Reading Systems, without a commonly-understood syntax there has been no way to achieve cross-platform interoperability. The functionality that can see significant benefit from breaking down this barrier, however, is varied: from reading location maintenance to annotation attachment to navigation, the ability to point into any Publication opens a whole new dimension not previously available to developers and Authors."[14]

Sharing books become much easier, when a link takes you directly a point in the text. For classroom use, a linked syllabus can be transformed into a course reader, allowing students to refer to the content they are discussing rather than page numbers.

Decentralized Distribution

Until recently, finding and purchasing a book was an activity that exclusively took place at a physical location, a bookstore. Bookstores had familiar sections for genres as well as new releases, specials, or best sellers. Many readers found books by browsing the store, others actively discussed books in clubs, received book recommendations from friends, or read book reviews in a magazine. Yet there are now more books in existence than there ever were, and this number continues to grow. Digital bookstores are unbound by physical restrictions in the quantity of books they keep in stock.

Currently, different distributors require roughly the same metadata, but support their own proprietary file formats. These proprietary systems result in many silos of content and inefficiencies due to non-interoperability. This creates a large amount of work for publishers who have to submit similar files and metadata to many different distributors.[15] Additionally, when submitting a file to a single distributor the content will be limited to a listing by only this distributor.

By publishing with the open standards Epub and OPDS, authors can take part in decentralized distribution where a book listing on their own server can support an unlimited number of listings in other locations or with larger distributors. For example, a book review on a personal blog might embed the author's book into the review, allowing the review reader to sample different sections of the text, as well as acting as an additional distribution point. In another example, ebook retailers, such as Feedbooks, can act as aggregators of content while the book listing itself is sourced from the author's own server. Historically, this is similar to RSS feeds, where content is published into an XML feed for syndication in other locations.

---

[14] http://www.idpf.org/epub/linking/cfi/epub-cfi.html#sec-overview-purpose-and-scope

[15] While there are meta-distribution services such as Smashwords that offer to simplify this process, these often take extra commission off the top.

Discovery

Linking and decentralized distribution can create new ways of discovering books and could also provide new ways to monetize and incentivize content creation, evaluation, and sales. For example, stores could be much smaller and more hand curated, and they could be everywhere, the store being embedded with a link. This profusion of curated storefronts could help more content to be discovered through social networks or interest groups, rather than through top seller lists and algorithms. In *The Long Tail,* Chris Anderson notes:

> "For too long we've been suffering the tyranny of lowest-common-denominator fare, subjected to brain-dead summer blockbusters and manufactured pop. Why? Economics. Many of our assumptions about popular taste are actually artifacts of poor supply-and-demand matching - a market response to inefficient distribution."[16]

Unfortunately, we still have inefficient distribution in terms of ebooks. Currently, unlike other forms of media that can be easily embedded into HTML, ebooks cannot be viewed in web browsers.[17] As a web resource, this is no longer a problem. Linking is a powerful form of sharing and thus distribution. Any website that links to the book becomes a storefront. Links to locations in web-based books can be shared in social mediums such as Facebook or Twitter.

Moreover, decentralized distribution could create new economic models that reward people that contribute to new parts of the information economy. In current e-commerce systems, consumers are rarely rewarded for their participation in the success of products. Websites like Amazon and Yelp have popularized product reviews and star ratings and consumers have come to rely on these mechanisms for purchasing decisions.[18] For example, Harvard Business School research has shown that a one star difference in a Yelp rating can cause a 5% to 9% increase or decrease in volume for a business.[19] Consumers who generate this information are almost always unrewarded for their content, time, and effort, yet their information contribution generates significant value for the service who hosts, or owns, their review.

Some businesses, including Amazon, have experimented with rewarding behaviors that might increase the sales of a product. Affiliate marketing for example, rewards website owners for linking to an Amazon product and the resulting purchases from web traffic. Other e-commerce sites offer cash or store credit rewards, typically around $20, to users who refer a new user that makes a purchase. More recently, Kickstarter has popularized a model for buying into a new or

---

[16] Anderson, Chris, The Long Tail, *Wired* 12.10, Oct 2004.

[17] At least without extensions.

[18] Kosur, J. (2013, September 2). *Yelp Affects 93% Of Researched Purchases, Increases Revenue For Small Businesses, Study Finds*. Social News Daily. Retrieved May 7, 2014, from http://socialnewsdaily.com/16367/yelp-effects-purchases-and-business-revenue-study-finds/

[19] Luca, Michael. "Reviews, Reputation, and Revenue: The Case of Yelp.com." Harvard Business School Working Paper, No. 12-016, September 2011. (Revise and Resubmit at the *American Economic Journal - Applied Economics*.)

potential product, much like a presale event, to both give producers a signal for the business potential and to generate buzz for the product itself.

We believe that these attempts to incentivize information generating behavior can be improved and further experimented with. Previous work by others hints at the potential new economic models could provide for incentivizing discovery. Popcuts a former Berkeley's School of Information project, centered around the idea of rewarding trendsetters and encouraging music buyers to be ahead of the curve.[20] Even more relevant were the results of the 2009 DARPA Network Challenge. MIT's very effective winning strategy involved recursively incentivizing the spread of information.[21] In the next section we will describe a technological stack that will give flexibility for more experimentation with regards to these ideas.

Solutions

EPUB.js - A Javascript Epub Renderer

To implement books on the web, we created Epub.js, a Javascript library that allows any webpage to render Epub documents. Epub.js contains a flexible rendering engine and provides a simple interface for common ebook functions such as styling, persistence and pagination.

An unzipped Epub 3 is a collection of HTML5 files, CSS, images and other media – the same as any other website. However, the Epub specification standardizes the table of contents, provides a manifest that enables the caching of the entire book, and separates the storage of the content from how it's displayed in ways that the W3C does not. Epub.js provides an interface for the core functionality needed to display a digital text. This includes loading, unpacking, rendering page content, pagination, linking, offline storage, annotation, and smoothing out cross-browser quirks.

More advanced ebook functionality can be built on these primitives. The ability to mark your place and pick up where you left off on any device, to automatically load pages in advance as the reader progresses through a book, search for a string of text, and to support annotations all depend on simple but robust linking.

By handling the primitives with Epub.js, the display of the content of the book is separated from the user interface of a reading system. This makes it possible to quickly customize the reading interface for different platforms and uses, such as for mobile devices or a class reader.

Rendering a book is as simple as including the Epub.js library

---

[20] Lim, K. M., Hesse, H., & Liu, Y. (2008). Buy Music, Make Money:An Incentive-based System for Digital Music Distribution, 1–13.

[21] Defense Advanced Research Projects Agency. (2010). DARPA NETWORK CHALLENGE, 1–17. Retrieved from http://www.eecs.harvard.edu/cs286r/courses/fall10/papers/ProjectReport.pdf

```
<script src="../build/epub.js"></script>
```

then setting the path to the book and selecting a HTML element to render to.

```
<script>
    var Book = ePub("url/to/book/");
    Book.renderTo("area");
</script>
```

An active community is beginning to develop around the Epub.js project. This can be attributed to following the lead of other successful open source projects, such as PDF.js, a PDF rendering library create by Mozilla. Following their model, development takes place on GitHub (a social network for code, based on the git version control system) and major decisions or changes are discussed on a mailing list that anyone can join.[22] Additionally, we try to maintain a presence on IRC (a chat room) to answer questions users may have. Finally, we have created documentation and tutorials to get new users quickly setup and rendering their first book.

As of May, there are 123 forks, which are user maintained versions, of the project on Github. 666 users have “starred” the library, as bookmark that alerts them of updates. Most importantly we have had 10 users contribute code to the project and 134 issues have been raised to suggest areas in which the project code fix bugs or improve, 84 of which have been addressed.

FuturePress - OPDS Catalog Server

FuturePress is a python server that enables distribution of Epubs as web resources through publishing a feed of available books using OPDS catalogs. There are two types of OPDS catalogs, both of which our server supports. The first are navigational feeds with links to other parts of the catalog, for instance new additions, discounted or best seller lists. Acquisition collections, the second type, list resource entries with metadata, such as title, description, author, and ISBN and provide links to acquire the content in various formats.

Links to content contain a “href” URL, a content MIME “type” and a link “rel”ationship.

```
<link rel="http://opds-spec.org/acquisition/open-access"
    href="/content/4561.epub"
    type="application/epub+zip" />
```

These links can lead directly to the resource (as shown above) or indirectly to it (for instance through a purchase page shown below).

[22] The mailing list and archive can be found here: https://groups.google.com/forum/#!forum/epubjs

```
<link rel="http://opds-spec.org/acquisition/buy"
    href="/product/1"
    type="text/html">
  <opds:price currencycode="USD">1.99</opds:price>
  <opds:indirectAcquisition type="application/epub+zip" />
</link>
```

OPDS defines six methods by which a resource can be acquired, expressed as a relationship. They are shown in Table 3 below.

Table 3: OPDS Relationship Definitions

Relationship	Definition
Acquisition	Generic retrieval link
Open Access	Retrieved without payment or registration at link
Borrow	Retrieved through after lending transaction
Buy	Retrieved through after purchase transaction
Sample	Subset of the Resource is available at link
Subscribe	Retrieved as part of a subscription after registration

Though the resource can be of any type, in all cases it is expected that a link (whether directly or indirectly) will eventually lead to the downloading of that packaged resource. In the case of an Epub (with the MIME type “application/epub+zip”) this is the Zip compressed package.

We propose streaming as a seventh relationship to enable using OPDS with Epub.js and our vision of books as web resource. A “Stream” relationship represents that the resource is available for immediate viewing without the need to download the entire file first.

This would require different behaviors for different file formats, for instance the mechanisms for streaming video file are well established but very different than streaming a html page. However, FuturePress focuses on what is needed to implement the streaming relationship for the Epub format.

For Epubs, a direct streaming relationship would link directly to the package document of an uncompressed Epub container. This follows the precedent set by the IDPF in their specification intra-publication linking schema.[23] The package contains the location of all the other resources that are needed to display the book content. Unlike in a contained Epub, these could be hosted

[23] http://www.idpf.org/epub/linking/cfi/epub-cfi.html#sec-intra-cfis

anywhere on the Internet and are only sent to a reading application as they are requested. Therefore a reading application that handles a streaming Epub must be able to handle standard HTTP error codes sent in place of the content, such as file not found or not authorized.

```
<link rel="http://futurepress.org/acquisition/stream"
    href="/content/4561/ops/package.opf"
    type="application/oebps-package+xml" />
```

An indirect link would instead direct a reader to a html page in which the book would be displayed. In the case of our FuturePress server, the rendering of content on this page would be accomplished using Epub.js, though it is by no means limited to use with our library. Several other similar methods for displaying ebook content online exist and method of display is of no importance in the catalog listing beyond declaring its type as “text/html”.

```
<link rel="http://futurepress.org/acquisition/stream"
    href="/read/1234"
    type="text/html">
  <opds:indirectAcquisition type="text/html" />
</link>
```

Once an Epub has been uploaded to the FuturePress server an entry will be created for the book and added to the appropriate acquisition feeds according to the permission set. Depending on permissions, the entry will contain links to the hosted or streaming content along with creation of all the necessary indirect acquisition pages.

We believe that content providers should openly list their on their catalog of books through catalog feeds. It is in the content makers interest to have their content listed in as many places/stores as possible. Feeds can be published by each publisher, library, or aggregated into large collections by content platforms.

Reader - Embedded Online Reading Application

To demonstrate the capabilities of a browser based reading application we created a reference implementation of a Epub.js based reading application. It uses the primitives provided by our library in combination to enable the functionality that is expected in a digital reading environment. It implements robust linking, annotations, endnotes, footnotes, search, reading position, bookmarks, nested table of contents, transclusions, embedded interactive content and more (shown in Table 4).

Table 4: Epub.js Reader Functionality

Functionality	**v0.2** - As of May 2014	**v1.0** - Planned for January 2015

Linking	Handles character offset and ranges. Can be passed through the URL request.	Fuzzy matching locations in text.
Search	Provides links to full string matches of search queries in a book. Basic highlighting of search terms.	Advanced faceted search and tokenized search queries.
Responsive User Interface	Adjusts controls to fit varying screen sizes. Supports touch events.	Phone vs Tablet specific controls.
Pagination	Creates and parses Epub page lists to match reflowable text with its print equivalent. Save reading position.	Improved generation of page counts when styling is adjusted.
Settings	Sets single or double page spreads, font size, adjustable width viewing.	Adjust background color, themes, offline storage.
Annotation	Integrates with Hypothes.is for inline annotating of text. Footnote and Endnote popups.	Group annotation functionality and support for longer form notes.
Rich Media	Supports SVG, video, audio and MathML.	Supports EPUB media overlay
Interactive Widgets	Support injecting arbitrary iframe widgets.	Implement the forthcoming EDUPUB widget specification.
Text Rendering	Supports rendering in all left-to-right rendered languages. Loads custom fonts.	Support right-to-left text rendering. Localized user interface. Better handling of custom font quirks.

It can be embedded in any webpage by using the HTML iframe element, just as a YouTube video is embedded.

```
<iframe src="/reader?bookPath=moby-dick.epub" width="400" height="600"></iframe>
```

In newer versions of web browsers it can also be embed as a custom web element.

```
<epub-reader src="/moby-dick.epub" width="400" height="600"></epub-reader>
```

Custom elements enable a pages author to add an ebook with the same ease that they can currently add a video or audio file. Unlike the iframe method, it adds semantic meaning to the content, indicating that the tag “epub-reader” contains Epub content.

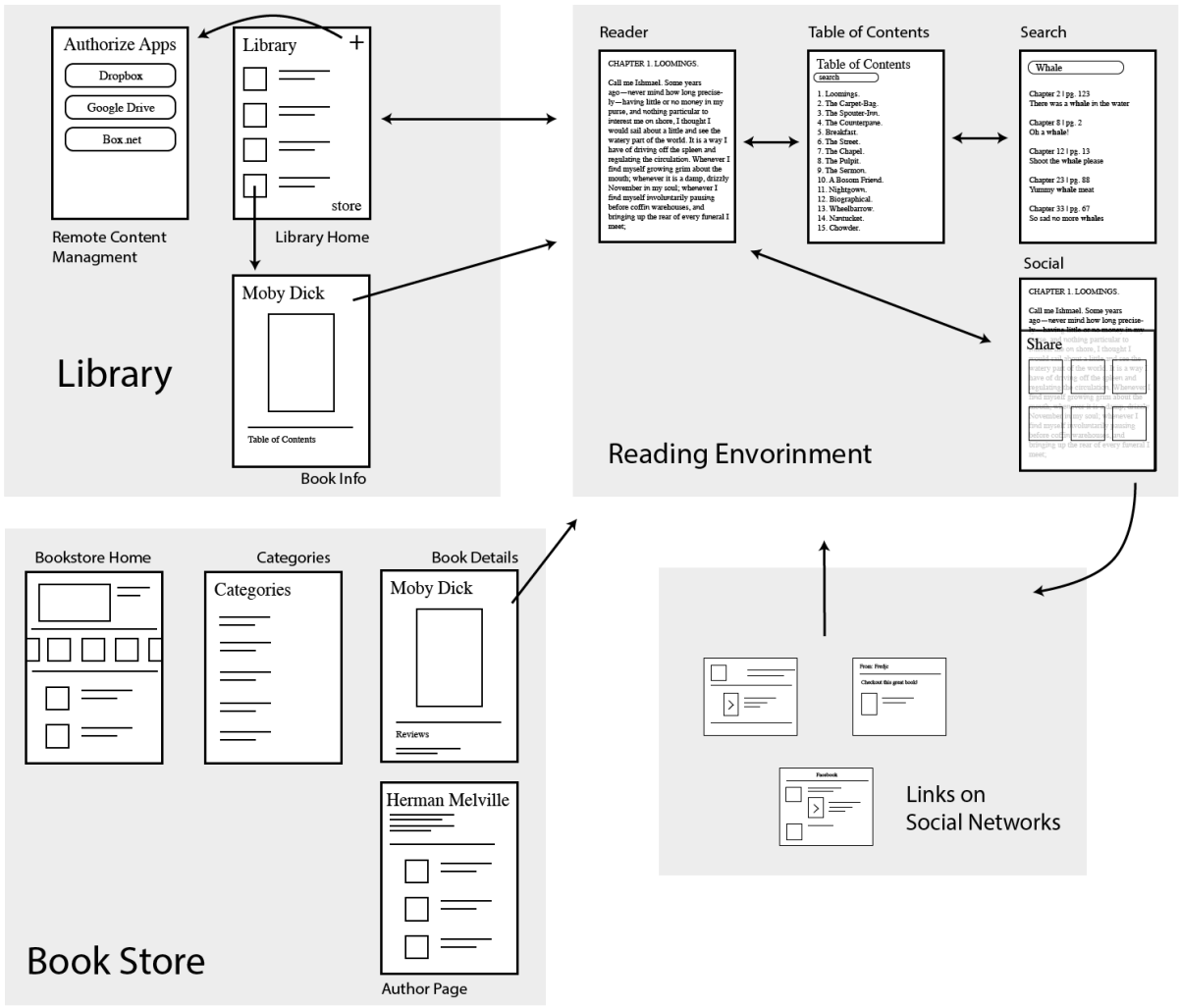
Shelf - iOS Application

Treating ebooks as web resources does not mean that they cannot be used in native applications. To demonstrate this we create a IOS app called Shelf that wraps the functionality of Epub.js in a native user interface. It allows readers to borrow and loan books through their

friends on facebook, as well as borrow books from OPDS catalogs. These catalogs can be pulled from any distributor, so the Shelf app manages books from many different sources. We seamlessly transition from clicking on a shared link to reading from that position in the book.

Shelf enables users to open Epub files on their mobile device in a unique but familiar interface, responsive to device's size. It respects an author's creation of epub styles without compromising the usability of the reading experience.

Figure 1: App Map



Our App Map in Figure 1. is broken down into three areas, Reading, Library, and Bookstore. These are the three core functionalities of our application.

The Reading portion also includes social screens which are an important part of the onboarding experience for a new user. Interactions within the Reading portion support a user's reading activity as well as giving the user actions to take with a book, such as bookmarking, annotation, sharing, and navigation with search or table of contents.

The Library portion of our App Map is where a user will interact at a higher level with the books that they purchase or import. Here a user will be supported to organize the books they own and access information about the book such as reading progress, notes, bookmarks, or social activities.

The Bookstore portion of our App Map includes a high level map of our OPDS functionality. This includes adding new Catalog feeds, browsing those feeds, book overview and downloading.

Possibilities: *Towards A Revolution in Publishing*

The technological stack we have described can enable many exciting new possibilities. This section will illustrate how these technologies would be combined to create new experiences and business models. To illustrate these possibilities we will use a hypothetical book called *Towards A Revolution in Publishing*. We'll go through the life cycle of the book and talk about the novel possibilities.

The book consists of a collection of essays on the future of publishing from a variety of perspectives. Some of the authors are well known, others less so. The editor of this book has decided to use it to experiment with a new type of book, something she calls a "Conference Book;" which is more accurately a cross between a MOOC,[24] social network, forum, and book.

The technological stack we have described for FuturePress could enable many exciting new economic and business models not possible in today's centralized environment.

Open Annotation: *the book as a place*

Open Annotation can allow for many exciting new interactions to take place inside books. In current reading systems, community and the content are separate, but with annotation these could be integrated. The website for the book not only provides the digital text but also holds the community around the digital text.

Months before the digital conference takes place, marketing begins. Emails are sent out to publishers and people who might be interested in taking part in a unique one week co-read in which the authors of the essays will engage with and respond to questions. People pay to register for when the book comes out, and on the day it does, they are free to read, annotate, ask the author's questions, and engage in discussion on publishing problems in forums. All of this happens in the browser, and powered by open annotation.

Data-Mining

New web-based distribution could allow for content makers to get access to much more information, allowing them to better understand their readers and content. Since *Towards a*

[24] Massive Open Online Course

Revolution in Publishing is hosted on the publisher's own servers, they are able to collect very detailed analytics about each of the essays in the collection.

The publisher can feed this information into what will be the final version of the book. Perhaps a certain essay in the book draws little readership and engagement; editors then make the decision to remove it from the collection.

Editors use information contributed and collected from readers to further refine and add to the content. They comb through the contributions and contact the owners of the best ones so that their entries are included in the books supplemental content. Anyone whose contribution is included in the book is given a small percentage of digital sales. They are given a certain percentage of shares, in other words a stake in the book, and are thus incentivized to talk about the book to their social circles. If a book has 100 authors, there is a good chance that those authors will be sharing that book to their Twitter followers.

Distribution

The book is then released for the general public on FuturePress and published on OPDS. Content aggregators and stores use the OPDS listing to syndicate the book. It's listed on bookstores and blogs across the Internet. Every sale results in seamless royalty payments. Since the original author and publisher get a cut of each payment, this produces a win win situation.

Towards A Future in Publishing is on the web and can also be indexed by search engines. A search query for "Essays on publishing's future," returns results from the pages within the book. A reader can click on the listing and read for a few pages before being asked to purchase the book. Moreover, the book is unbundled in that rather than purchasing the whole book, a reader can buy individual chapters.

Incentivized Sharing

Anyone who buys the book is given a special link to the book, which they can use to share. People can read a sample and purchase the book just by clicking on the link. They can post the link on social media sites, their blogs, or through email. Anytime their activity results in a purchase, they get a small percentage, thus people have ownership in the success of the book. The many authors, contributors, and readers of *Towards a Revolution in Publishing* have an incentive to spread knowledge of the book and talk about it within their social networks.

Advantages

While simple in concept, this is innovative in a number of regards, and manages to monetize content and network value in a number of places. First, by having exclusive limited windows of time to ask the author(s) questions, publishers are able to capitalize on the fact that authors are interesting people and people will pay to engage with them. Second, this round of feedback and discussion can produce useful extra content. Third, by rewarding the makers of this secondary content, publishers can effectively and ethically crowdsource certain aspects of market

research, writing, editing, and marketing. Forth, all of this produces content that can then be sold in more traditional book markets.

This hypothetical example shows how new information architectures can be used to create new types of value. We believe there is much more experimentation to do in this area, but this work hinges on open standards for distributing content and digital payment systems that lack transaction fees.

Conclusion

Almost twenty years ago, John Perry Barlow wrote *The Economy of Ideas* for Wired Magazine. The central problem put forth in his essay has still not been resolved:

> If our property can be infinitely reproduced and instantaneously distributed all over the planet without cost, without our knowledge, without its even leaving our possession, how can we protect it? How are we going to get paid for the work we do with our minds? And, if we can't get paid, what will assure the continued creation and distribution of such work? [25]

This gets to the heart of what we are trying to address; how do we fairly spread the wealth of the information economy among more participants? There is a great tension that exists between building open, accessible, and powerful information architectures and creating systems that allow intellectual property to be fairly compensated. While controversial, many such as Jaron Lanier have argued, “As our economy turns into an information economy, it will only grow if more information is monetized.”[26] We argue that the best path towards monetizing more information will be to create content distribution systems that are based on standards, flexible, and allow for innovation.

Clark et al. in their visionary essay *Tussle in Cyberspace: Defining Tomorrow’s Internet* introduced the term “tussle,” which they defined as an ‘ongoing contention among parties with conflicting interests.’[27] Indeed, there are many conflicting interests in the digital publishing world. A very clear example is Google, which was accused of copyright infringement by the Authors Guild while digitizing millions of books. Ultimately, Clark et al saw tussle as beneficial to the future internet, but they stressed ‘the need for network designers to think explicitly about tussle and the design requirements it implies.’[28] In the current paradigm of proprietary centralized distribution systems, the tussle has been fairly lopsided; tech companies have largely designed the technical systems to benefit themselves. Ultimately, content makers and others who have an interest in digital publishing would benefit from participating in standards bodies, and having their

---

[25] Barlow, J. P. (n.d.). The Economy of Ideas. Wired. Retrieved from http://archive.wired.com/wired/archive/2.03/economy.ideas_pr.html

[26] Lanier, Jaron. *Who Owns the Future?*, Simon & Schuster, 2013. ISBN 978-1-846145223.

[27] Clark, D. D., Wroclawski, J., Sollins, K. R., & Braden, R. (n.d.). Tussle in cyberspace: defining tomorrow's Internet. IEEE/ACM Transactions on Networking, 13(3), 462–475. doi:10.1109/TNET.2005.850224

[28] Clark, D. D., Wroclawski, J., Sollins, K. R., & Braden, R. (n.d.). Tussle in cyberspace: defining tomorrow's Internet. IEEE/ACM Transactions on Networking, 13(3), 462–475. doi:10.1109/TNET.2005.850224

say in the designs of future systems. FuturePress is our prototype of an architecture that is advantageous to all participants, and we hope that our open platform will be used to implement new possibilities for creating, interacting with, and monetizing content.

Bibliography

Abeel, T., Van Parys, T., Saeys, Y., Galagan, J., & Van de Peer, Y. (2012). GenomeView: a next-generation genome browser. Nucleic Acids Research, 40(2), e12–e12. doi:10.1093/nar/gkr995

Abel, S. (2012, May 15). Hey, Book Publishers! Evolve or Get Out of the Way | The Content Wrangler. thecontentwrangler.com. Retrieved September 11, 2012, from http://thecontentwrangler.com/2012/05/15/hey-book-publishers-evolve-or-get-out-of-the-way/

Anderson, Chris, The Long Tail, *Wired* 12.10, Oct 2004.

Barlow, J. P. (n.d.). The Economy of Ideas. Wired. Retrieved from http://archive.wired.com/wired/archive/2.03/economy.ideas_pr.html

Bastable, M. (2013). How I overcame snobbery to self-publish an e-book. The Telegraph.

Brantley, P. (2012). The Curation of Obscurity. In H. McGuire & B. O'Leary (Eds.), Book: A Futurist's Manifesto. O'Reilly Media. Retrieved from http://book.pressbooks.com/chapter/curation-of-obscurity-peter-brantley

Braunschweig. (n.d.). Textbooks round the world: It ain't necessarily so. The Economist. Retrieved October 3, 2012, from http://www.economist.com/node/21564554

Cheng, J. (2012, April 13). Apple: we broke "Amazon's monopolistic grip" on e-book industry. Ars Technica. Retrieved April 13, 2012, from http://arstechnica.com/apple/2012/04/apple-we-broke-amazons-monopolistic-grip-on-e-book-industry/

Chiang, C.-N. (2010). A MULTI-DIMENSIONAL APPROACH TO THE STUDY OF ONLINE ANNOTATION, 1–329.

Chuang, J. and Sirbu, M., Optimal Bundling Strategy for Digital Information Goods: Network Delivery of Articles and Subscriptions. *Information Economics and Policy* 11(2):147-176, 1999.

Clark, D. D., Wroclawski, J., Sollins, K. R., & Braden, R. (n.d.). Tussle in cyberspace: defining tomorrow's Internet. IEEE/ACM Transactions on Networking, 13(3), 462–475. doi:10.1109/TNET.2005.850224

Clark, L. (2013, April 28). Raspberry Pis hide narrative in the city to complete offline book. Ars Technica. Retrieved April 28, 2013, from http://arstechnica.com/business/2013/04/raspberry-pis-hide-narrative-in-the-city-to-complete-offline-book/

Dannen, C. (2013, May 10). This Is What Happens When Publishers Invest In Long Stories. Fast Company. Retrieved May 11, 2013, from http://www.fastcolabs.com/3009577/open-company/this-is-what-happens-when-publishers-invest-in-long-stories#comments?utm_source=buffer&utm_medium=twitter&utm_campaign=Buffer&utm_content=buffer65a7f

Defense Advanced Research Projects Agency. (2010). DARPA NETWORK CHALLENGE, 1–17. Retrieved from http://www.eecs.harvard.edu/cs286r/courses/fall10/papers/ProjectReport.pdf

Davies, S. (2011). Still building the memex. Communications of the ACM, 54(2), 80. doi:10.1145/1897816.1897840

DeMarque. (2013). Power to propel you, 1–12.

Digital Content. (2013). Digital Content. American Libraries.

Downey, A. B. (2010, January 6). The Textbook Manifesto. greenteapress.com. Retrieved October 3, 2012, from http://greenteapress.com/manifesto.html

eBook Architects. Data adapted from “eBook Retailers” Retrieved May 1, 2014, from https://ebookarchitects.com/learn-about-ebooks/retailers/

Economist, T. (n.d.). E-book publishing: A too-cosy world? | The Economist. The Economist. Retrieved September 5, 2012, from http://www.economist.com/node/21552599

Economist, T. (n.d.). J.K. Rowling and Pottermore: The Hogwarts express | The Economist. The Economist. Retrieved September 5, 2012, from http://www.economist.com/blogs/prospero/2012/04/jk-rowling-and-pottermore

Economist, T. (n.d.). Enhanced e-books: Truly moving literature | The Economist. The Economist. Retrieved September 5, 2012, from http://www.economist.com/blogs/prospero/2012/02/enhanced-e-books

Hsiao, I.-H., & Brusilovsky, P. (2011). The role of community feedback in the student example authoring process: An evaluation of AnnotEx. British Journal of Educational Technology, 42(3), 482–499. doi:10.1111/j.1467-8535.2009.01030.x

International Digital Publishing Forum (2011, October 11) EPUB 3 Specification. Retrieved September, 2012, from http://idpf.org/epub/30

Kosur, J. (2013, September 2). Yelp Affects 93% Of Researched Purchases, Increases Revenue For Small Businesses, Study Finds. Social News Daily. Retrieved May 7, 2014, from http://socialnewsdaily.com/16367/yelp-effects-purchases-and-business-revenue-study-finds/

Lanier, Jaron. *Who Owns the Future?*, Simon & Schuster, 2013. ISBN 978-1-846145223.

Lessig, L. (2008). *Remix*. Blackwell Publishing.

Lim, K. M., Hesse, H., & Liu, Y. (2008). *Buy Music, Make Money:An Incentive-based System for Digital Music Distribution*, 1–13. http://www.ischool.berkeley.edu/files/Buy%20Music%20-%20Make%20Money.pdf

Lopez, S. (2013, April 27). iPads in school: a toy or a tool? - latimes.com. latimes.com. Retrieved April 29, 2013, from http://www.latimes.com/news/local/la-me-0428-lopez-ipads-20130428,0,197667.column

Marshall, C. C. (2010). Reading and Writing the Electronic Book. (G. Marchionini, Ed.)Morgan & Claypool, 1–201.

Marshall, C. C., & Brush, A. J. B. (2002). From Personal to Shared Annotations. ACM, 1–2.

McGuire, H., & O'Leary, B. (Eds.). (2012, January 10). Book: A Futurist's Manifesto | Essays from the bleeding edge of publishing. O'Reilly Media. Retrieved April 15, 2013, from http://book.pressbooks.com/

Nelson, Les, Convertino, G., Pirolli, P., Hong, L., & Chi, E. H. (2009). Impact on Performance and Process by aSocial Annotation System: A Social Reading Experiment. Augmented Cognition, 270–278.

Open Publication Distribution System (2011, June 27) OPDS Catalog 1.1 Specification Retrieved March 1, 2014, from http://opds-spec.org/specs/opds-catalog-1-1-20110627/

Pearson, J., Buchanan, G., & Thimbleby, H. (2011). The Reading Desk:Applying Physical Interactions to Digital Documents. ACM, 1–4.

Pearson, J., Thimbleby, H., & Buchanan, G. (2012). Co-Reading: Investigating Collaborative Group Reading. ACM, 1–10.

Rags Srinivasan, M. P. (2012). Freemium has run its course.

Roscorla, T. (2011, September 26). Indiana U. Helps Shape Economic Terms of eText Transition. centerdigitaled.com. Retrieved November 12, 2012, from http://www.centerdigitaled.com/classtech/Indiana-U-Helps-Shape-Economic-Terms-of-eText-Transition.html?utm_medium=direct&utm_source=embedded&utm_campaign=University-Leaders-eTexts

Savikas, A. (2012). Distribution Everywhere. In H. McGuire & B. O'Leary (Eds.), Book: A Futurist's Manifesto. O'Reilly Media. Retrieved from http://book.pressbooks.com/chapter/distribution-everywhere-andrew-savikas

Shiller and Waldfogel, Music for a Song: An Empirical Look at Uniform Song Pricing and its Alternatives, *Journal of Industrial Economics*

Siegel, L. (2013, May 3). Seeking Out Peer Pressure. nytimes.com. Retrieved May 4, 2013, from https://www.nytimes.com/2013/05/05/fashion/seeking-out-peer-pressure.html?smid=fb-share&_r=0

Signer, B., & Norrie, M. C. (2011). A model and architecture for open cross-media annotation and link services. Information Systems, 36(3), 538–550. doi:10.1016/j.is.2010.08.002

Streitfeld, D. (2012, November 4). Shunning Amazon, Booksellers Resist a Transformation. nytimes.com. Retrieved November 5, 2012, from https://www.nytimes.com/2012/11/05/technology/shunning-amazon-booksellers-resist-a-transformation.html?_r=0

Trujillo, S. (2013, February 28). 5 Despicable Things People Do for Good Online Reviews. cracked.com. Retrieved March 1, 2013, from http://www.cracked.com/article_20266_5-despicable-things-people-do-good-online-reviews.html

Varian, H. How to Build an Economic Model in Your Spare Time. In *Passion and Craft: Economists at Work*, edited by Michael Szenberg, University of Michigan Press, 1997.

Varian, H., Section 6 "Switching costs and lock-in" in Economics of Information Technology. Mattioli Lecture at Bocconi University, Milan, Italy, 2001.

Victor, B. (2012, September 5). Learnable Programming. worrydream.com. Retrieved October 24, 2012, from http://worrydream.com/LearnableProgramming/

Weiseth, P. E., Munkvold, B. E., Tvedte, B., & Larsen, S. (2006). The Wheel of Collaboration Tools, 1–10.

Wikert, J. (2012, February 9). It's time for a unified ebook format and the end of DRM - Tools of Change for Publishing. toc.oreilly.com. Retrieved September 16, 2012, from http://toc.oreilly.com/2012/02/unified-ebook-format-end-drm.html

Wischenbart, R. (2011). The Global eBook Market:, 1–43.