

E-Books Are Not Books

Mark T J Carden
Ingram Digital
300 Poole Road, Poole,
Dorset, BH12 1AZ, United Kingdom
+44 1202 753 232
mark.carden@ingramdigital.com

ABSTRACT

Currently, in the early days of their development, e-books are essentially following the evolutionary path of physical books, a path that started thousands of years ago. Yet physical books are containers for a wide variety of information types, and are accessed in a wide variety of ways, which offers the possibility of differing electronic manifestations. The evolutionary approach will continue to be reasonably successful in meeting current needs, but the real growth in the adoption of e-books will happen when the traditional book is deconstructed and reconstructed (textually, behaviorally and commercially) in order to create new paradigms for storing and delivering content in electronic forms.

ACM Categories & Subject Descriptors:

A.0 General Literature

General Terms:

Theory

Keywords

Electronic Book; E-Book; Ebook; Digital Content.

1. THE EVOLUTION OF E-BOOKS

Electronic books commonly in use today are primarily computerized representations of physical books. They may be scanned page images (viewable as PDFs), or re-flowable text-streams that are reconstructed by a software application to resemble pages on a reading device.

Large repositories of this e-book content are being built, both through digitization of older books and through 're-publishing' of modern book content in e-book form. Modern content is mainly available through publisher or aggregator repositories (such as OCLC's NetLibrary and Ingram's MyiLibrary), while older books have mainly become available through harvesting of library content, whether commercially by Google or as part of institutional archiving.

Creating 'electronic pictures of books' is a reasonable and understandable phase at this point in the evolution of the e-book, comparable to the way that early automobile designs strongly

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

BooksOnline '08, October 30, 2008, Napa Valley, California, USA.
Copyright 2008 ACM 978-1-60558-249-8/08/10...\$5.00.

resembled the 'horseless carriages' that they indeed were. Yet it is not necessarily a paradigm that will, or should, be adopted in the longer term, as new ideas and technologies start to allow us to escape from the limitations of a collection of bound pages.

The physical book is a storage and presentation tool that can be used to support a wide variety of information forms (data, stories, arguments), access methodologies and purposes. The fundamental structure of the book, and the way it is used, is broadly unchanged since the invention of the codex in the 3rd century BC. Refinements to the codex have undoubtedly been helpful. The movement from papyrus (or wax tablets) to paper, the introduction of moveable-type printing, and the adoption of the index and the table of contents have all improved accessibility and usability to a significant, but not transformational degree. The latest stages of this evolutionary refinement in the electronic era have included making the pages visible on an electronic screen, making the text searchable and embedding multimedia.

Yet within the pages of conventional books are contained a wide variety of information types, which are accessed in profoundly different ways. A parent and child reading together through a picture-book are doing something very different from a citizen looking up a telephone number, a student working through a textbook or a commuter reading a novel. Yet we currently think of all of these activities as in some sense similar – they are all 'reading a book'.

Some books are to be tasted, others to be swallowed, and some few to be chewed and digested: that is, some books are to be read only in parts, others to be read, but not curiously, and some few to be read wholly, and with diligence and attention.

Sir Francis Bacon, English author & philosopher (1561-1626)

If we treat all of these experiences as 'reading a book', and transfer this traditional model into the world of the e-book, it seems likely that we will misunderstand both 'reading' and 'book' in the electronic age. We will be in danger of not only creating misconceived or dead-end electronic solutions, based on this single model, but also of missing opportunities to enhance each of these separate experiences through appropriate (and distinct) new technologies. There is a need for a more in-depth analysis of the many different ways in which we use books at present, and for the development of some creative approaches to how we might take this content+use combination and re-implement it in the electronic world.

2. A NEW ELECTRONIC PARADIGM

There have of course been many attempts by academics to define or classify e-books, which could help to tease apart the 'what is a

book?' question; yet for all their positive intentions, many definitions have been constraining rather than illuminating, and many classifications have tended to be focused on unhelpful dimensions or characteristics, such as the reading device, the content format or the source of the book. (Space limitations preclude a review of these definitions and classifications in detail here.)

Meanwhile the marketplace is taking its course and we are starting to see some polarization of the different natures of 'the book' as we are drawn, apparently reluctantly, into the electronic world. These developments are to some extent being forced upon the producers and consumers of these books, by market demands or technology opportunities, in the absence of a coherent anticipation of what the new models should look like. While some of these 'accidental' developments are encouraging, others are worrying, and may result in unanticipated disruptions in both business models and the flow of knowledge.

2.1 E-Books As Databases

An example of a significant migration away from the traditional 'book' model, that is already taking place, is the way in which dictionaries, encyclopedias and other similar reference materials are rapidly becoming searchable databases. (As a young library patron is reputed to have said, "Why would anybody print out a whole encyclopedia?") This change has created considerable disruption in the economics of the publishers and distributors of such 'books', yet many of these publishers are adapting well to the electronic world and finding new revenue streams through electronic licensing of these databases or pay-per-view models. A reference book in the physical world is not turning into an e-book at all, but into an electronic searchable (and browsable) database of discrete factual components.

2.2 E-Books As Learning Objects

Meanwhile, many textbook publishers have embraced (whether enthusiastically or resignedly) the opportunity to take their content online in an enhanced or extended format. At its simplest level this can mean embedding the textbook content in an application that permits searching, highlighting, annotation and similar useful features. Beyond this, we are also seeing integration of textbook content with assessment tools, homework management systems and learning environments. There is starting to be some acceptance that traditional physical textbooks are, in reality, collections (or perhaps sequences) of learning components bound together in the pages of a book, and although there may be commercial reluctance to deconstruct these into their individual components, or to allow these to be 're-purposed', there is clearly an opportunity to use new technology to provide a richer learning experience than can be provided by pages in a book.

But while we may be starting to look at textbooks as collections of learning objects, we need to remember that they are also often used in a way analogous to the reference books described previously. Students (and others) will often use a textbook to look up a single piece of information (and this is now being made significantly easier by the search capabilities of e-textbook delivery applications). But this implies that we need to start thinking of textbook content serving two separate needs simultaneously. This is something that was previously achieved by a single physical book being used in two distinct ways (either by locating data via the index or by working sequentially through

the text, guided by the table of contents), but in the electronic world there is potentially both the need and the opportunity for two completely separate delivery environments (and commercial models). The e-textbook content needs to be accessible as both learning objects and a searchable database; the 'nature' of the book being determined neither by the traditional book-type classification, nor by the content of the book itself, but by the intent of the content consumer at the moment the book is opened.

2.3 E-Books As Viewable Resources

Perhaps the most successful use so far of 'traditional e-books' (ie e-books based on the existing printed book model) is electronic access to academic monographs via publisher or aggregator repositories. Libraries have invested millions of dollars in licensing this type of content, and the modest amount of research to date suggests that content available in online repositories is very actively used by both students and academics. This content is usually delivered in a page-by-page format via a web browser (with limited copy, download or print capabilities), and users seem to be accessing this content in two main ways: firstly as a reference tool to access not just an item of data (as in an online dictionary), but a larger element of textual description (a multi-page account, proposition or argument, perhaps); and secondly as a very quick and convenient way to scan publications to identify items of potential interest for further in-depth study.

This 'enhanced mimicry' of the existing physical book model has been relatively successful, and offers additional functionality, such as full-text searching, hyper-linking of the table of contents, and bookmarking. This model also has the potential for further functional enhancements, such as hyper-linking of references and of the text itself.

The limitations here seem to be partly about format portability, with users wanting to 'take away' portions of text after viewing, and commercial models, which are still evolving and thus have some undesirable complexity and variability. Both of these issues are not really problems of technology but of copyright, licensing and commercial terms. The tensions between e-book providers and librarians (and library users) are to some extent based on the misunderstandings of the similarities and differences between books and e-books. Users (librarians, students or faculty) complain that it is unacceptably limiting if they are restricted from embedding e-book content in learning materials, printing whole e-books, or lending e-books to other institutions, yet content owners are legitimately concerned about the implications of a 'buy-once, duplicate-many-times' philosophy adopted by many of those who obtain access to e-content.

2.4 E-Books As Narratives

A common reaction by people when asked about e-books is that they could never read a 'whole book' on a computer screen or a hand-held device. Firstly, of course, such comments fail to recognize the many ways in which a book can be used without 'reading the whole book', as described above; but consumers of, for example, trade fiction are legitimately highlighting how unsuitable computers and small-screen devices (like PDAs and Smartphones) are for long sessions of reading text. So where the book is a container of a narrative, a story (or perhaps an academic account or argument), which requires access through sustained continuous reading, a better user experience is needed. Despite

this significant and legitimate concern, millions of fiction titles are downloaded every year and consumed on these unsuitable devices (or printed out), but if suitable reading devices were widely and affordably available it seems likely that consumption of this sort of material would take off rapidly.

Recently we have started to see the emergence of dedicated reading devices using electronic paper technology, such as the Kindle, Sony, Iliad, Cybook and others, which with next-generation improvements in usability and, most importantly, price are set to trigger this upswing.

Today, narrative content is typically available in three main forms: page images (typically PDFs, often too large for e-book readers or small-screen devices), structured re-flowable text (typically specific to a particular e-book viewing application and digital rights management protocol) and unstructured re-flowable text (such as the free TXT and DOC files found in many web repositories).

The key to the delivery of the narrative e-book is the ability of readers to access it on a suitable device; this is why re-flowable text is so important, and why open or proprietary standards, DRM, and vertical lock-in models like Amazon's Kindle, become contentious issues. Encouragingly, publishers are starting to provide content into the supply chain in the re-flowable EPUB standard, which can then be delivered directly to devices or transformed into one of the common (secure) formats.

Other approaches to delivering narrative are also being pursued; easily the most successful of these being audio-books, which must be omitted from the scope of this paper. Other approaches include text that is 'read aloud' by software, and flashing one word at a time on small-screen devices like cellphones (using WordUP and ICUE, for example).

2.5 E-Books As Imagery

Finally, returning to the scene of the parent and child reading a picture-book together, we can easily imagine that this experience, when electronic, could be enhanced by animation and interactivity. Parents already complain that children prefer the liveliness of television and computer games to the apparently passive printed book, but it might be that books that are presenting imagery or narrative to younger children could be transformed by new technologies. It is not clear, however, whether such a development would help to lead children on a path toward more traditional reading experiences, or simply hasten a profound preference for highly animated media.

3. RE-INVENTING THE E-BOOK

Having reviewed the varied types of content that are stored within traditional books, and the varied ways in which this content is accessed, we might conclude that we need to reappraise what an e-book could potentially become, and indeed consider whether the e-book is simply a soon-to-be-obsolete phase, rather like the 'horseless carriage' in the evolution from the brougham to the modern automobile.

Consequently, it might be useful to have a methodology to deconstruct the insufficiently differentiated clump of content and purposes that is 'a book', in order to create a range of conceptual models and technology solutions that deliver this content in appropriate electronic ways (with appropriate commercial

arrangements). Perhaps these can still be collectively called e-books, but they may well turn out to be too diverse to be held together by any single term.

To do this, we need to look closely at what the book content really is, and at how it is going to be accessed by the user. This paper is not going to attempt this process in a systematic way, but some possibilities are outlined below.

3.1 Content Types and Access Behaviors

The content itself could helpfully be classified into some groupings along the lines of "data", "explanation", "instruction", "account", "argument", and "narrative". User behaviors also need to be identified, and these will map in a many-to-many way to the type of content. These behaviors might turn out to be things like: "look up", "skim", "view", "consume", "immerse", and "enjoy".

There are also related categorizations that may be helpful. We probably would benefit from differentiating between "pure text" "text plus pictures", "pure pictures", "multimedia" and "audio".

3.2 Commercial Models

It would also be useful to examine models of acquisition and possession. As we know, it is today possible to "skim" or "view" some "instructional" material though the relatively simple and inexpensive library-based licensing and browser-based viewing model. But to "consume" the content in an "immersive" way probably implies an ownership model, downloaded content and powerful software application package that can be delivered by a sophisticated textbook platform. Both of these models exist and function well today, but there is confusion amongst publishers, aggregators, librarians and consumers regarding how this all 'should' be working, and regarding what content, commercial model and application is appropriate for the various traditional books being brought into this electronic world. This may get worse, not better, without some new paradigms.

3.3 Page Layouts

In content production and presentation we need to give more attention to the page layouts in which we structure text for e-book consumption, whether we are aiming to produce static pages or 'intelligent' re-flowable text. The constraints and opportunities presented by electronic screens, and by the ability of devices to reformat text (and images) on-the-fly, mean that what works for a printed page may well be unsuitable for electronic devices. This article, for example, is formatted to fit a 180x235mm window on a 216x279mm page, in a portrait format, with a two-column layout – just about ideal for a printed journal, yet hopeless for a landscape-screen laptop, an e-book reading device or a small-screen device. To read this article, as it is currently formatted, on a Blackberry would be an impossibility. Yet it could be easily reformatted (if the underlying text streams and formatting tags are preserved) to be perfectly readable on such a device.

3.4 Reading Devices and Formats

Devices are also important, not just because they have a significant impact on consumers of information. Someone will say "I'm not interested in e-books – I could never read a novel on my computer screen." Yet they would probably be willing to look up a recipe on a computer screen, or read a novel on a dedicated

e-book reading device. Effective devices are available today and we can expect hardware improvements and significant price reductions in 2009, which should take e-book reading devices into the mainstream.

A significant challenge will be the inevitable 'format wars', where some players will attempt to lock consumers into a single vertical model, while others will wish to participate in an open approach which would allow users to make independent choices about device, format, retailer and content. Linked to this will be digital rights management (DRM) issues, where some will want to use DRM technology defensively (or offensively) to enforce the lock-in, others will use it in a manner sufficient to legitimately protect against piracy, and others will seek to follow the music industry in allowing open access to content (if they can find a business model that will support this).

4. CONCLUSIONS

The e-book, in its current early stage of development, is already a success, with widespread adoption of e-content in the academic and educational world, and new excitement in the trade book

market. This trend will continue, and we can be confident that there will be a further growth in the content available, an increasing sophistication of the academic and educational offerings, and a rapid growth of consumer purchase and use of e-books as suitable devices become mainstream.

Yet we are only at the start of an evolution from codex to e-book. While we continue vigorously to support this evolutionary development, we need simultaneously to start to analyze, model and construct the revolution that will take us into a world of electronic content, electronic repositories and electronic access that is only tangentially connected to the traditional book.

5. ACKNOWLEDGMENTS

This paper has been informed by: the insights of colleagues at Ingram Digital and others in the book business, by the JISC E-Book Observatory project and the research done by Professor David Nicholas and his team at CIBER, by the Random House/Zogby Poll of reader behaviors, by the 2000 EBONI paper "The Problem of Defining Electronic Books", and by many other unattributable yet valuable sources and conversations.