Journal of Librarianship and Information Science

http://lis.sagepub.com

Books in a virtual world: The evolution of the e-book and its lexicon

Chris Armstrong

Journal of Librarianship and Information Science 2008; 40; 193
DOI: 10.1177/0961000608092554

The online version of this article can be found at:
http://lis.sagepub.com/cgi/content/abstract/40/3/193

Published by:
SAGE
http://www.sagepublications.com

Books in a virtual world:
The evolution of the e-book and its lexicon

CHRIS ARMSTRONG

Over recent years there has been considerable confusion over the use of the term 'e-book', and this article examines the variety of definitions used to date while proposing a definitive construct. Beginning by examining the definitions of 'book', the paper moves on to consider the essential element of a book – the content, and to examine publishing and structural aspects of e-books, as well as their place in libraries, before arriving at a final definition. The definition and its derivation embrace all of the issues that affect the way in which e-books are understood and used today. In conclusion, the article looks at both the genesis of e-books, and the stage of acceptance and adoption that they have reached, with brief reference to 3rd-generation e-book readers available at the time of writing.

KEYWORDS: e-book definitions; e-book readers; e-books; e-resources; libraries; library collections

Chris Armstrong is Managing Director of Information Automation Limited (IAL), a consultancy, research and training company in the library and information management sector which was established in 1987. His work focuses on electronic information resources, and their effective use in libraries and information centres; in recent years, e-books have been the subject of much of the company's research. Chris and colleague, Ray Lonsdale, jointly run a series of regular training courses on the management of e-book collections in libraries. Chris publishes, and speaks at conferences regularly. He is a Fellow of the Institute of Analysts and Programmers (FIAP), as well as of CILIP (FCLIP), and is currently a National Councillor of the latter and an active member of UKeiG: the UK eInformation Group – the CILIP Special Interest Group which focuses on the accessibility, use and management of electronic information. He also publishes 'info NeoGnostic', a blog (http://i-a-l.blogspot.com/) on which current e-book issues are regularly aired.

Address: Information Automation Limited
Penbryn
Bronant
Aberystwyth
SY23 4TJ
Tel: +44(0) 1974 251302
E-mail: lisqual@cix.co.uk
Website: www.i-a-l.co.uk
Blog: http://i-a-l.blogspot.com/

INTRODUCTION

About a quarter of the articles and papers written about e-books contain a statement such as this: 'The term "e-book" includes the hardware, a suitable device to read electronic media, perhaps better called "e-book readers"' (Whalley, 2006). It is partly in an attempt to dispel this misleading impression of the inseparability of content and hardware that I wrote this article.

THE BOOK

The word 'book' is so much a part of everyday conversation that there is no doubt but that every reader already has a mental image and a complete understanding of what is being discussed. No two readers will have the same picture, but all will have a clear idea of the essential features: what it is for, how it is used, and so on. Even some 10 years after we began to see e-books as a real presence, the same cannot be said of the term 'e-book'.

Partly, this is because the idea of an electronic book, or an e-book, is still relatively new. It is also due to the variety of forms, formats and features given to e-books. Even the physical form of this virtual entity gives rise to debate!

What follows will discuss the emergence both of electronic texts and of the concept of an e-book; it will also look at how the term has been used in the literature; and will present a definition of the term which is fit for 21st century-purpose – a definition that I believe covers adequately its place today for writers, publishers, librarians and readers.

Vol 40 (3): 193–206
[DOI: 10.1177/0961000608092554]

Some of the understanding and recognition of how a print-on-paper book may be used have a considerable influence on our understanding of, and expectations for, an e-book. Consequently, if mundanely, a short discussion about books – the physical, print and paper-based reading material – is a necessary starting point.

The legacy: Definition(s) of a book

While it is true that we all readily and immediately understand what is meant by the term 'book', this does not mean that there has never been a need to define the word formally. Dictionaries, including the *Oxford English Dictionary (OED),* have comprehensively explained the word and how it is (and has been) used. The 1989 second edition of the *OED* seems to have covered all angles: a book is a 'written or printed treatise or series of treatises, occupying several sheets of paper or other substance fastened together so as to compose a material whole'. It goes on to explain that,

> [i]n this wide sense, referring to all ages and countries, a *book* comprehends a treatise written on any material … put together in any portable form, e.g. that of a long roll, or of separate leaves, hinged, strung, stitched, or pasted together.

This is not enough for the definition, which goes on to confirm modern usage:

> a treatise occupying numerous sheets or leaves fastened together at one edge called the *back*, so as to be opened at any particular place, the whole being protected by binding or covers of some kind.

OED goes on to explain that as

> either the form of the book or its subject may be mainly or exclusively the object of attention, this passes on either side into [either] [t]he material article so made up, without regard to the nature of its contents, even though its pages are occupied otherwise than with writing or printing, or are entirely blank [or a] literary composition such as would occupy one or more volumes, without regard to the material form or forms in which it actually exists.

And almost as a footnote there is the aside that, in one sense, every volume is a book while in the other, one volume may contain several books. Not unreasonably, the *OED* makes no reference to physical size, number of pages, authorial or editorial intent, or purpose. That being the case, it is forced to conclude that, 'No absolute definition of a "book" in this sense can be given'.

Others, too, have attempted a definition. The United Nations Educational, Social and Cultural Organization defines a book as a 'a non-periodical printed publication of at least 49 pages, exclusive of the cover pages, published in the country and made available to the public' (UNESCO, 1964) and this stricture on length has found its way into encyclopaedias such as *Britannica* (Publishing, history of, n.d.) and glossaries, such as that of the African American Literature Book Club in Harlem, New York (http://writers.aalbc.com/publishing_glossary.htm) and a commercial design and printing company in Tucson, which uses a similar, very precise, although different, definition:

> A general classification to describe papers used to print books; its standard size is 25x38 inches. A printed work which contains more than 64 pages. (AlphaGraphics, nd)

The emphasis is on the physical entity. Noam Chomsky (2000: 20) wrote:

> Books are concrete objects. We can refer to them as such ('the book weighs five pounds'), or from an abstract perspective ('who wrote the book?'; 'he wrote the book in his head, but then forgot about it'); or from both perspectives simultaneously ('the book he wrote weighed five pounds', 'the book he is writing will weigh at least five pounds if it is ever published').

In 2001, a virtual conference was hosted by Bibliothèque publique d'information: Text-e functioned through a number of online debates, several on the nature of an e-book (text-e, 2001). Robert Casati began one debate with the question: 'The philosopher may wonder: what is a book? This is an ontological problem: the book is at once a physical and a mental entity', and Umberto Eco responded to a question on what distinguishes a book from any other form of information by suggesting that it is the reader's 'psychological mechanisms of attention' which vary between a book and other read forms. An idea perhaps first suggested by Patrick Bazin, who wrote of the book as a 'stable, reliable and public interface' between the author and the reader (Bazin, 1996: 158).

The concentration on the physical attributes as the essence of a book is strange, and may be one reason why, as we shall see, there has been so much difficulty in accepting the term e-book. The *OED* speaks also of the content '. . . literary composition such as would occupy one or more volumes, without regard to the material form or forms in which it actually exists'. And this binary approach is surely important. A book without its content would be useless except for developing new content. A book is designed, serves no purpose other than to convey content. So when we talk of a book surely we refer implicitly to the content. If we speak of David Crystal's book entitled *Linguistics* or Anthony Burgess' book, *Language Made Plain,* we are not referring to, have no real interest in, the 267 pieces of paper that make up the former or the 206 of the latter. The purely physical aspect is only important in so far as it allows an author's wit, wisdom or whimsy to be communicated easily, to be 'transported' from author/publisher by way of bookshop or library to readers. An interface. The book is the sum total of its contents made visible and thus also, perhaps, tangible.

If we accept this, then the corollary is that the actual format or construction of the book is unimportant. The texts that Crystal or Burgess developed certainly need a medium for their transport and for convenience, up until quite recently, necessarily this would have been a physical stack of paper hinged at one side.

This argument establishes that one significant attribute of a book is its content. Wilson (2000), in a discussion on information needs, noted that the word 'information' may be variously understood as the physical document, as channels of communication, or as the *subject data* contained in a document (or transmitted orally). Wilson's term 'subject data' is interesting, bringing into play as it does the primacy of the content as – in terms of logic – the thing about which a judgement is made or which has attributes. We may say that the content (subject data) has physical attributes. The content is the reason that a physical document (in this case, a book) was constructed as a channel of communication. The subject or the defining essence of a book is the content, which has attributes that describe its physical presence.

If a book is to be defined by its content, we must ask what aspects of the content are important in defining it as a book? Content may be pictorial, graphic, photographic or textual, or any combination of these, and historically all have been found in books; similarly it can be seen that content can be generated by one or many authors or contributors. Clearly the nature and source(s) of the content are not significant. The amount of content, as has already been shown, seems important. While an arbitrary page length seems inappropriate, our understanding needs to exclude tracts, pamphlets and other lesser publications. Journals and magazines must be excluded by nature of their serial publication. Books are not always linear – that is, to be read beginning at page one and on through successive pages until the end, the form also embraces directories and dictionaries, for example. With all these complex aspects, it is easy to see why past definitions have concentrated on the physical!

Perhaps we may suggest here that a book is 'a substantive amount of content published in some physical form – normally paper sheets bound at one edge – and which is not distributed serially'.

Now we are in a networked age, content can be given substance and imparted in ways other than the physical – certainly a lesser substance, a transient substance – but a substance nonetheless that serves the exact purpose of a codex, a manuscript or our old friend, the paper book. All that was needed was a name for this nouveau-book! Probably because of the near universal acceptance of the term ejournal or e-journal, e-book slipped into the language. There has been some rear-guard action for digital book – as in digital library and the digitizing of paper content to create many e-books – but this seems destined to fail.

The conceit of an e-book seems entirely appropriate and yet 'e-book' remains a term of which people are unsure, which is defined variously, and which is still, after some years, struggling for acceptance. Given that codices – manuscript volumes which were the prototype of the modern paper book – were in common use by the 6th century and we still have variations in understanding of the term 'book' in the 21st century, it is hardly surprising that e-book has no universally accepted definition after so few years.

The Text-e discussion referred to above became more animated when the concept of an e-book was called to question. If there was no consensus reached, several of the contributors did agree that an e-book could not be defined by its particular physical standard (for example, those of Microsoft or Adobe), and that any ultimate or standard physical form would be governed by commercial considerations. The 'physical' was again subordinate to the 'mental'.

THE E-BOOK

Existing definitions

Just as dictionaries have defined 'book', so too has at least one dictionary produced a definition of 'e-book'. The *OED* published a draft definition in 2001 which did little to clear up one of the main debates: is the e-book only an e-book in its physical entity, that is, when it is given substance by an e-book reader, or is it simply the content?

> A hand-held electronic device on which the text of a book can be read. Also: a book whose text is available in an electronic format for reading on such a device or on a computer screen; (occas.) a book whose text is available only or primarily on the Internet.

It will have become clear from the discussion of 'book' that the first *OED* definition is not acceptable, although it can be seen that some writers have taken this approach. Curiously, the *OED* reference in support of this part of the definition was taken from a 1988 prediction on 'technicological gadgets [that] will soon be developed' for *American Libraries* by futurist Robert Olson (1988: 391). The prophesy was given some credence more than a decade later by Morgan (1999: 36) who wrote, 'E-books are not the same as e-texts. E-books connote a hardware/software combination used to read electronic data on a portable electronic device specifically designed for such purposes'. Christine Borgman (2000: 90) brought Olson's prophesy up to date when she wrote, 'Reading a novel on an electronic book while traveling by air seems probable'. All of these despite an authoritative definition to the contrary published in 1997, stating that an e-book is a 'term used to describe a text analogous to a book that is in digital form to be displayed on a computer screen' (Feather and Sturges, 1997: 130)

Lynch (1999) begins his two-part article, 'Electrifying the Book', 'Many people have now seen, or at least heard about, the new consumer electronics appliances popularly called "e-books" or "electronic books" or (more accurately)

'electronic book readers"'. This ambiguity continued to haunt both the industry and Lynch, and in June 2001, he wrote, 'Imprecise and inconsistent terminology has been a major source of confusion in the hype over e-books, and an obstacle to disentangling the issues involved'. The over-emphasis on the technology was still evident some months later when Hillesund (2001) wrote that a 'narrow definition treats an e-book as a digital object designed to be read on a handheld reading device or to be listened to from a speech-generating tool'. In 2002, Sawyer's review of the literature to establish a definition of electronic books suggested that it 'is currently related more to hardware and software (viewing technology) than content' (Sawyer, 2002). That some confusion over meaning still existed in 2002 was highlighted in a paper by Garrod and Weller (2004), which describes how respondents to research undertaken by Loughborough University (Dearnley et al., 2002: 18) understood the term to mean CD-ROMs. Garrod and Weller also suggest that the naming of the Rocket ebook 'may have contributed to the notion of ebook as device'. Esposito (2003) also attempted to clarify the situation but had to acknowledge that his suggestion was unlikely to be adopted.

> As we begin to publish some books in electronic form, the print package gets tossed out and only the content remains. The content of such a package, however, is also called a book, and that is the kind of book I wish to discuss … Some interested parties now use the term etext to distinguish the content from its package. This would be more helpful if enough people subscribed to the convention.

Unfortunately, at the time of writing in 2006–7 it remained the case that many, including many librarians, understand the word in terms of a particular technology. A report commissioned in the academic sector stated: 'confusion caused by the lack of an acceptable definition of what an e-book is has become abundantly clear during the course of the research undertaken for this report, and this confusion in itself also constitutes a barrier to uptake' (Gold Leaf, 2003).

By 2004, pundits were using phrases such as 'content is king' and suggesting that 'content consumers are format-agnostic' (OCLC, 2004: 2). Perhaps unconsciously echoing the Hodgkin (2003: 13) or Lynch (2001) view that the e-library or personal digital library (and not individual books) are what we should be thinking about, the Corporate Vice President of netLibrary (an e-book aggregator or supplier) said (in a personal conversation) that 'in 10 years, there will be no e-books, only e-content'. All of this is to suggest that users 'do not care much what sort of container – such as a book, journal, blog or a Web page – the content comes from' (OCLC, 2004: 2) rather than, in our narrower context, that the content rather than delivery mechanism is the essence of an e-book. However, the general emphasis on content supports this view, and Pace (2004: 74), writing on e-books, notes that 'technology without good content is form without substance. The future success of the e-book marketers depends on the ability to offer valuable content'. In the few years since the turn of the century, opinion had begun to focus on content. Cox and Mohammed (2001) supported the 1997 definition that an 'e-book is a term used to describe a text analogous to a book that is in digital form to be displayed on a computer screen' while McKnight and Dearnley (2003: 235) suggested that an e-book is the 'content of a published book made available to the reader in electronic form'. In the same year, Ohene-Djan and Fernandes (2003) wrote that 'This paper defines the content of an e-book to be a network of digital information units which may comprise text, graphics, video, animation and/or sound', and Rao (2003: 86–7) used

> text in digital form, or a book converted into digital form, or digital reading material, or a book in a computer file format, or an electronic file of words and images displayed on a desktop, note-book computer, or portable device, or formatted for display on dedicated e-book readers.

Also in 2003, the *International Encyclopedia of Information and Library Science* updated its definition, continuing to use the book-analogy approach, although this definition does not seem to be widely known:

> The result of integrating classical book structure, or rather the familiar concept of a book, with features that can be provided within an electronic environment is referred to as an electronic book (or e-book), which is intended as an interactive document that can be composed and read on a computer. (Landoni, 2003)

Four years later in 2007 the *Wikipedia* definition remains focused on the need to clarify the confusion over content and technology:

> An ebook is an electronic (or digital) version of a book. The term is used ambiguously to refer to either an individual work in a digital format, or a hardware device used to read books in digital format. Some users deprecate the second meaning in favour of the more precise 'ebook device. (*Wikipedia, 2006*)

As Garrod and Weller (2004) say, 'It is important to distinguish between digital content and the technology which enables a reader to access this content'. Their own definition – 'the term ebook tends to mean actual content – i.e. books that are available in electronic form, and which can be downloaded from the Internet and read on a variety of hardware platforms with the aid of reading software' – like that of both the *Wikipedia* and McKnight and Dearnley appears to exclude books written *for* the medium, and focuses only on print books that have been digitized or replicated electronically. Most writers would not subscribe to this narrower interpretation. Landoni's definition, on the other hand, seems to exclude books which are *not* composed on a computer.

Without in any way subscribing to the OCLC suggestion that consumers are format-agnostic (OCLC, 2004: 2),

it seems clear that no *definition* should be media-dependent. It also begins to be apparent that there are other issues that have to be considered. What is less apparent is the degree to which these issues affect any understanding of what is meant by an e-book.

Digitized vs born-digital

The first e-books – those in archives such as the Gutenberg Text Archive – were typed manually by volunteers. As technology has moved forwards, this approach has almost completely been replaced by digitization, or scanning followed by optical character recognition (OCR) to produce a text that can be manipulated and read on screen. In some cases, such as the Culturenet Cymru's *Books from the Past* (http://www.booksfromthepast.org/), or *The Complete Works of Charles Darwin Online* (http://darwin-online.org.uk/), this is saved as a text version in parallel with the page images of the original book; in others, the text is used to produce an e-book in one of the standard formats.

A second starting point is also possible. A book may be first created for use only in a digital form, having no previous version existing on paper – the term 'born-digital' has been coined – and this may be typified by a much greater use of the design features available. If, from its inception, a book has been planned and designed for reading electronically the author and publisher are both able to take advantage of the medium to add value to plain text, and it may seem, at least initially, less like a conventional book. However, if it has the essential qualities of a book – a substantive amount of related content, or content with inherent continuity, that is not published serially – is 'book-like' and is accessible to be read on computers or on e-book readers, the origin does not affect its classification.

Linear text vs hypertext

With the exception of reference works, the majority of print-on-paper books are designed to be read linearly. In most cases this means beginning at the first page and continuing until everything has been read and absorbed. There are also cases – collections of stories or poems, conference proceedings, edited works containing chapters by different authors, textbooks, etc. – where the linear read may be confined to one or more sections or chapters. Essentially, paper books, however, are read linearly. The very essence of computer use (computer-mediated reading and writing) in the 21st century is to have open multiple windows and to jump from one to another in order to follow a train of thought. While writing this article I have frequently moved to my web browser, my e-mail client and other software products, often in mid-sentence to follow up a chain of thought or to work on an unrelated task. Some e-books, particularly those written for the medium, may take advantage of this multi-faceted delivery. Additional windows may open with further information, the reader may be able to move around the e-book to follow a theme or idea that cuts across several chapters, it may be possible to annotate or post a message to the author, film clips may be interpolated with the text, but – just as a dictionary, in which users may chase their way through a chain of definitions, remains a book – the way in which an e-book communicates its message, and may be read, does not affect its 'book-ness'. A particularly rich example of a born-digital book is *City Sites: Multimedia Essays on New York and Chicago, 1870s–1930s* (http://artsweb.bham.ac.uk/citysites/).

Plain text vs value-added text

In the same way that the fashion in which an e-book may be read cannot change its categorization as a book, the integration of a large amount of non-textual material – audio and visual clips, moving images, still images, tables and graphs – or extensive added functionality does not detract from the book-ness. Just as the inclusion of engravings in books did not create a new form, and the addition of a thumb index does not turn a book into a 'reference machine', so the addition of illustrative material and other user aids appropriate to the new medium does not, of itself, distinguish an e-book from a book, or an e-book from some other form of resource. The 2003 *International Encyclopedia of Information and Library Science* definition quoted earlier, continues '[f]rom the conceptual side, it is an attempt to overcome the limitations of paper books by adding a series of useful features that are made possible through the nature of an electronic environment' (Landoni, 2003: 168) – a definition that would be endorsed by one of the earliest conceptualisers of e-books, Alan Kay (2005). Whether the concept of an e-book was born of the need to overcome the limitations of paper, or not, it is certainly true that most e-books include a number of features that would have not been possible in a bound, paper-based book.

Research undertaken in 1998 identified 12 specific types of added value, including: resource links; links to reviews; author biographies; and links to curricula, professors' or other educational sites (Armstrong and Lonsdale, 1998: 35). e-Book products from 68 publishers were examined during the survey and the relative popularity of these mechanisms is shown in Table 1.

If the same survey were to be repeated now, a number of additional features, such as the ability to manipulate data, might be included. Recent research in the academic sector has emphasized the potential of e-books to offer educational institutions added value such as additional exercises, demonstrations, assessment materials (Education for Change Ltd et al., 2003: 48), and the great value of these texts as instruments for research and reference.

> One of the great things about it is you can actually search across entire books. I gave a class a few weeks ago to some PhD students, one of them was searching caricatures of the medical profession in Dickens. She was absolutely

Table 1. Use of value-added features

Type of added value	Publishers (%)
Resource links	47.22
Review links	25.00
Editorial/Feedback	20.83
Author links	18.06
Links to journals	15.28
Author biographies	15.28
Companion websites	13.89
Simulation/Animation	12.50
Exercises/Questions	11.11
Links to employment sites	8.33
Bookshop links	6.94
Links to professor/curricula	2.78

> delighted that she could actually search thousands of books for Dickens . . . doctor, nurse medic, midwife and so on – to search thousands of books [in the past] would have taken her years to do. In that way, that's transforming how she can do her research. (Gold Leaf, 2003: 148)

e-Books differ from books in some quite fundamental ways, but it is suggested here that while some e-books may add more 'around the edges' than is possible in a paper equivalent, they remain books because of their essential objective: the making public of a large amount of content.

Delivery mechanisms

The definition that we are moving towards encompasses any electronic/digital text that is book-like in nature and that can be read on some form of computer – whether desktop, laptop, handheld or a specialist form known as an e-book reader or e-book device. It has been suggested that it is not limited in content or form or structure; nor by virtue of its origin. In addition, it seems reasonable to suggest that the means of delivery does not – in most cases – affect whether something is, or is not, an e-book. With the use of the Internet and the World Wide Web increasing annually, it is unsurprising that most e-books are delivered in this way; however, e-books can also be delivered on diskette, CD-ROM and memory card as well as over mobile telephone networks.

One form of delivery, however, must be set aside. Some writers include print-on-demand books within the e-book stable. Rao (2004: 263), for example, writes 'textbooks in e-book format can be made modular and offered as "print-on-demand" depending on the copyright situation'. Print-on-demand may be based at a commercial publisher as is the case with McGraw-Hill's PRIMIS system or it can be localized (for example, at book shops as was suggested by Peurell (1999: 179–86) for the Swedish company, Podium). This could also be a mechanism used by university presses, as in the USA by Rice University, or extended to individual lecturers who could create their own hybrid textbooks to be printed on demand. In these cases, delivery to the local printer is almost certainly electronic but the end-product is very much print-on-paper, and this author would subscribe to the view put forward by Crawford (2004) in his review of the previously cited article by Andrew Pace (2004), in which he suggests that the cross-classification has been made as a part of

> the ongoing efforts of ebook advocates to pad the sales figures for ebooks. The BookMachine can't be the 'savior of the e-book format' for a simple reason: It has *nothing* to do with ebooks. It's a print-on-demand system. PoD is a great development, but it *strengthens* print books. A stack of paper sheets with ink or toner on them bound into a heavier cover is called a book. Not an ebook: A book.

A similar division has arisen over audio books. Audio books – on tape and CD-ROM – have been around for some time and no one has suggested that they be called e-books until recently. E-books require software to 'translate' the computer files into something that can be understood by human readers, and some software products incorporate the facility to read aloud the text on the screen. They thus become audio books! But to create a definition of e-books that specifically includes audio books seems to be as perverse as one that is accepting of print-on-demand. It would mean that all those cassettes and CD-ROMs of famous actors reading everything from Agatha Christie's *Death on the Nile* to Zola's *Germinal* become e-books! While an e-book may have audio-book capabilities, it would be false logic to extend this statement to deduce that all audio-books are e-books!

Sectoral issues

Is an academic librarian's perception of an e-book different from that of a school, special or public library? Would the essential definition vary? It is impossible to imagine such a debate taking place about a book so it may be wrong to entertain it for e-books! A view expressed in the recent past by a public librarian was that in his sector, books had to be physically loanable – so a web-based product was not, for him, an e-book. For him, e-books were only possible when downloaded onto an e-book reader. As will already be apparent, this is not a criterion for the definition being developed here. In exactly the same way that delivery, content, style and origin do not define an e-book, the medium on which it is read is no more and no less than the medium on which the e-book is read.

It will have become evident that it is the view of this author that none of the above matters of detail (except in the two extreme usages of print-on-demand and cassette-based audio books) have bearing on the definition for an e-book presented here.

Perceptions and semantics

Perhaps the one area that offers the most difficulty in our attempt to characterize an e-book is that of the semantics. In our discussions about the added-value elements, some readers may, for example, have wondered when, or how, an e-book differed from a website. Equally, there may seem little to distinguish some e-book reference titles from the older concept of a full-text database; they can, for example, be interrogated in much the same way with extracts of the whole delivered as answers.

In some cases, e-books are presented on the screen as facsimiles of paper books, so that the reader comes close to the same reading experience. The layout suggests text on a page, there are virtual pages to turn, bookmarks may be placed, and the general metaphor used is that of a book. This is not always the case. Some e-books display text and images on the computer screen with little attempt to ape a book and surprisingly little attempt to render the product very readable – a necessary functionality whether or not the product looks or feels like a book. Landoni et al. (2001: 190) demonstrated that a better reading experience produced improved cognition and comprehension, and a later study on fiction e-books stated 'that concentrating on the appearance of text, rather than the technology itself, can lead to better quality publications to rival the print versions' (Malama et al., 2005). Some user studies have been undertaken in the US – all have been limited to a single institution – and Columbia University concluded that for e-books to be a success, it is necessary to 'provide books that will have the greatest utility in the online format to selected user groups with interfaces and functionalities that are both easy to use and valuable to the scholar' (Summerfield, 2001). As Landoni et al. (2000: 407) noted with respect to their first study,

> the process of reading and the tasks readers are attempting to complete have a central role in judging whether such a translation [from paper to an electronic environment] is suitable. The cognitive overhead associated with a computer-based environment is an important reason for carefully considering the appropriateness and the method of realizing this conversion

although Van Dam (2001) has suggested that the 'fundamental reading experience (viewing the written word as a means of receiving and digesting ideas and creative expression)' has not altered.

While these studies hint at the desirability of a product that mimics the experience of using a book, they do not suggest that it is a determinant in the semantics of electronic information resources. As I have suggested that a book is a substantive amount of content published in some physical form, it would be logical to suggest that an e-book is book-like content made available electronically or digitally. Perhaps what distinguishes an e-book from a website, what makes it book-like, is a continuity of purpose which is often reflected in a book-like structure of chapters and pages.

Finally, there is an issue that – for librarians at least – has a bearing on how e-books are viewed, and it is important to give this cognizance. There was a time when the purchase of a book was relatively straightforward, leaving librarians with a physical copy, with which – under first sales doctrine – they were free to do as they saw fit. The publisher exercised – and *could* exercise – no more control. Among other corollaries, this meant that the book could be loaned or sold; and that, even if the publisher subsequently vanished from the face of the earth, the library would still have a copy of the book to meet the needs of its users. As McCracken (2004: 123) wrote,

> Licensing rights is different from buying a physical object in a very important and fundamental way. The first sale doctrine does not apply . . . This means, for example, that if a library buys a copy of a book or subscribes to a print magazine or learned journal then, under the first sale doctrine, it may retain old copies as part of its archive or grant access to readers who are not members of the library. This applies even if the subscription lapses . . . However, if the library instead subscribes . . . it does not buy a physical copy of the work and the doctrine does not apply. Depending on the terms of the subscription licence, access to archived editions . . . may evaporate along with a lapsed subscription and casual readers may be excluded.

Not all e-books are licensed in this way, but for libraries, most are. As we arrive at a definition of 'e-book', it is important to understand this final aspect of the nature of an e-book: e-books may be purchased, licensed, given and loaned; many may be used together as one resource; they may also be taken out of stock, replaced and probably read in several places at once – their very form makes for a far greater spectrum of use.

So we arrive at a definition. An e-book is:

> any content that is recognisably 'book-like', regardless of size, origin or composition, but excluding journal publications, made available electronically for reference or reading on any device (handheld or desk-bound) that includes a screen.

This definition and its derivation have covered issues that affect the way in which e-books are understood. At the time of writing, in 2007, it is still true to say that many people who hear the term, do not readily and easily picture an e-book in their mind. Some reader perceptions and some of the issues that have been discussed – perhaps most importantly, that of licensing – may well affect their ability to reach a critical mass of readers.

At the turn of the century, Borgman wrote: 'New technologies such as . . . "electronic books" try to replicate the features of paper' (2000: 183), although Landoni sees e-books more as an attempt to overcome the *limitations* of paper books (2003: 168); it is suggested here that e-books have arrived at an adoption stage and that, in different ways and in different usages, both these statements may

be true. Instances exist of e-books which far exceed the features of paper, while others do not even attempt to replicate it. The technology – memory, storage and display – has advanced to a point where some hand-held e-book readers offer a reading experience that is virtually equivalent to the accepted comfort of a paper book; although, of course, without the indefinable pleasure to be gained from handling a physical book. At the same time a lack of standards militates against their easy acceptance. It seems clear that a phenomenon similar to the electronic journal has arrived; what is less clear is what libraries and readers are to make of it in an age where even the established e-journal is fighting a turf-war with the open archives initiative.

Relatively few e-books have been written as born-digital books, while many thousands of existing print works have been digitized, and a smaller number are published in parallel with the printed book. Fiction, poetry, drama, textbooks, monographs and reference works all exist as e-books. Some publishers have not taken to the medium (in this context, Rod Bristow, President of Pearson Education UK was quoted in 2002 as saying: 'It's very hard to deliver what you get in a textbook more effectively in other ways. Textbooks are extremely convenient' (Midgley, 2002)); others self-evidently have embraced it. Some authors, Stephen King with *Riding the Bullet* for example, have experimented with the medium while others, such as Umberto Eco, have declaimed against it for works which are intended to be read virtual-cover to virtual-cover – novels. Eco suggested:

> There are two ways to publish a novel in the Internet. One is to publish it online; in this sense I don't believe too much. I think that you cannot read a book on the screen – after three pages your eyes become two tennis balls: it's insane. Then there is another way to mean a novel by Internet – there's hypertext novels in which the reader can 'mix them up'. Then I believe that the beauty of a great novel, and so even of the small ones, is that they give you the image of destiny. With a sort of Internet hypertextual maneuver you can arrange, let's say Tolstoy's *War and Peace,* in order that Prince Andrew [Bolkonski] doesn't die because you love him too much, you don't want he dies. The beauty of the story is that he dies – and he cannot do anything to stop it; so a novel is a sort of a story about destiny, about the ineluctability of human affairs – so to make it so open to the personal initiative, well I don't think this is interesting to me. (*Front Row*, 2005)

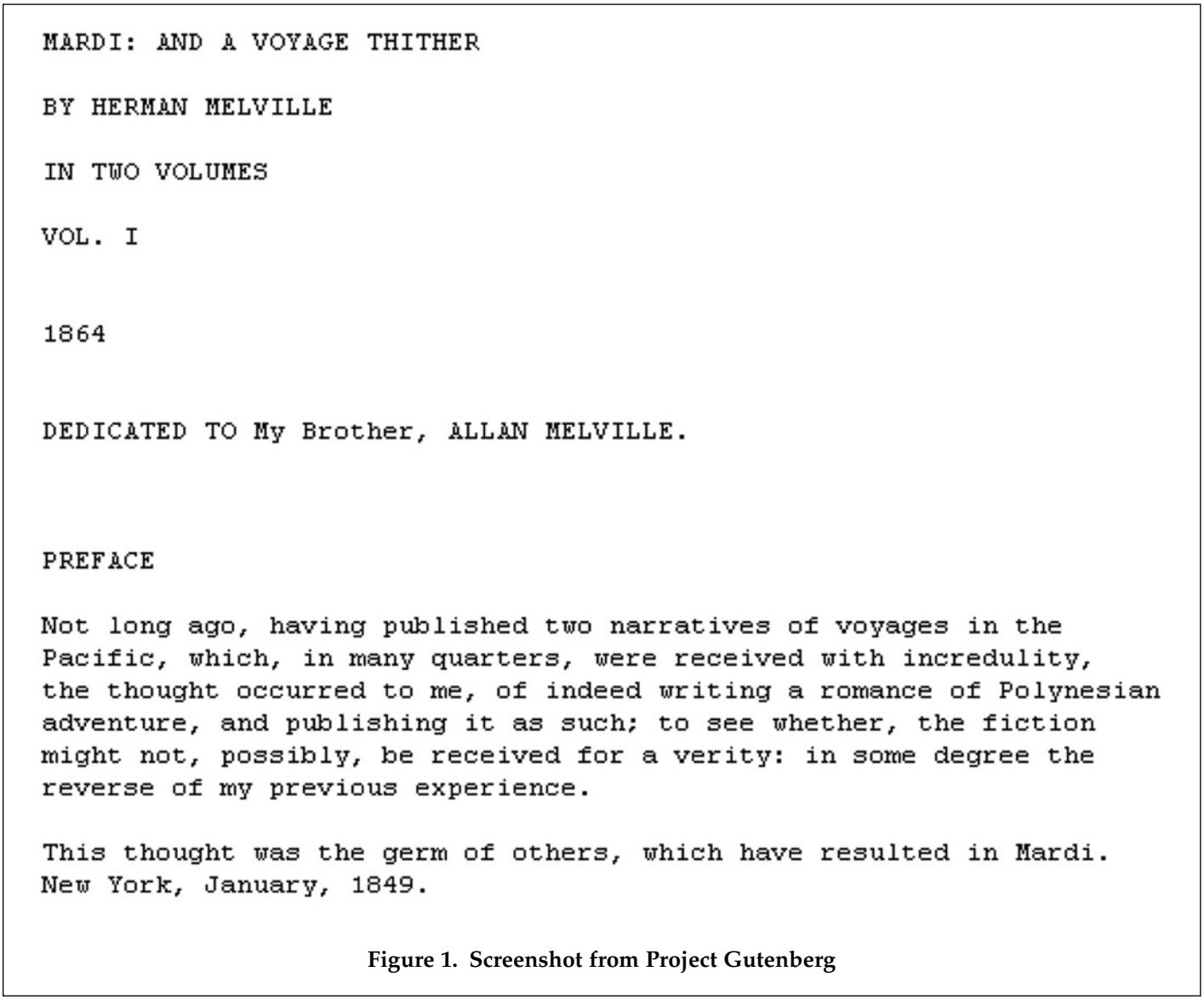
Meanwhile, pundits such as Clifford Lynch (2001) have envisioned a future of 'portable personal digital libraries, not portable electronic books'; a future in which every reference book you are likely to need, as well as the textbooks, novels and monographs you are currently reading can be carried around in something no larger than an Apple iPod. And we are nearly there. The two things which stood in the way of this particular Utopia: the poor reading experience and the variety of software packages needed if a library is not to be severely limited in its choice of titles either have been, or are being, addressed. Ashcroft and Watts wrote, 'Access to e-books continues to develop, with numerous platforms available, and lack of standardisation an ongoing problem' (2004: 284). On websites such as Manybooks.Net (http://manybooks.net/) or FictionWise (http://www.fictionwise.com/home.html), e-books are available in a range of formats: Adobe PDF, hiebook, Microsoft Reader, Palm, RocketBook and MobiPocket. With the advent of the Sony Librie in 2004 – the first e-book reader with the much-praised near-paper-like e-ink display, which could have revolutionized e-book use – yet another proprietary software (Broadband Electronic Books) was introduced, thus limiting its use to yet another very small selection of titles. With its latest Reader (available only in the States) Sony remain determined to lock in readers with books available from their own Connect® eBookstore. However, other e-ink e-book readers – the iRex Technologies' iLiad (http://www.irextechnologies.com/) which sets out to accommodate different formats and, with a capacity of around 200 books the CyBook from Bookeen (http://www.bookeen.com/) which accepts the MobiPocket format supported by Amazon with over 50,000 titles available – offer libraries many possibilities.

GENESIS OF E-BOOKS

Although Vannevar Bush, President Roosevelt's Science Advisor, wrote in his seminal *Atlantic Monthly* article of a 'memex . . . a device in which an individual stores all his [or her] books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility', it was nearly three decades before the first electronic books began to appear (Bush, 1945: 106). In about 1968 Alan Kay developed the Dynabook concept: 'Just as the book was an extension of the oral medium, so is the computer an extension of the print medium. There are many things that books can do, but computers have an extra dimension that seemed to me incredibly important, and this is key to the Dynabook idea' (Kay, 2005). At that time, the technology was not in place but Bush's idea of a hand-held reader was affirmed.

However, it is arguable that their eventual appearance *was* presaged by other technology. Andries Van Dam (the term eBook is normally attributed to him) has suggested that

> the eBook in its present incarnation – namely, the ability to store text in electronic form and subsequently retrieve it through the use of a computerized device – can trace its lineage back to the early methods of automated textual storage and retrieval developed as long as a century ago, as well as to the development of electronic document creation, storage, retrieval and output mechanisms in the 1950s and 1960s. (Van Dam, 2001)



MARDI: AND A VOYAGE THITHER

BY HERMAN MELVILLE

IN TWO VOLUMES

VOL. I

1864

DEDICATED TO My Brother, ALLAN MELVILLE.

PREFACE

Not long ago, having published two narratives of voyages in the Pacific, which, in many quarters, were received with incredulity, the thought occurred to me, of indeed writing a romance of Polynesian adventure, and publishing it as such; to see whether, the fiction might not, possibly, be received for a verity: in some degree the reverse of my previous experience.

This thought was the germ of others, which have resulted in Mardi.
New York, January, 1849.

Figure 1. Screenshot from Project Gutenberg

Although Karen Wiesner suggests that the first e-books were developed as ASCII files of his own work by Bob Gunner in the 1970s (Wiesner, 2003: 18), the more obvious birth of the genre lies in the Gutenberg archive of electronic texts, which was started in 1971 by Michael Hart. The archive began with the US Declaration of Independence, and then added the Bill of Rights, quickly followed by the whole US Constitution. The Bible was added, as individual books of the Bible were not that large, and then Shakespeare (one play at a time). Project Gutenberg (http://www.gutenberg.org/) continues today with the same philosophy to encourage the creation and easy distribution and use of e-books. The home page states, 'Most of the Project Gutenberg eBooks are older literary works that are in the public domain in the United States. All may be freely downloaded and read, and redistributed for non-commercial use.' The project has avoided pressures to create 'authoritative editions', seeing its target audience as the general reader who does not care whether a certain phrase in Shakespeare has a ':' or a ';' between its clauses. Lebert (2004) has summarized the history of the Gutenberg Project. An example of a 'plain Vanilla ASCII' Gutenberg text is shown in Figure 1.

Other similar text archives (or collections of free e-books) were developed later, for example, the Oxford Text Archive (http://ota.ahds.ac.uk/), which opened in 1976, and the Etext Center at the University of Virginia Library (http://etext.lib.virginia.edu/), set up in 1992,

which states on its home page that it aims to 'build and maintain an internet-accessible collection of SGML and XML-encoded texts and images' and to 'build and maintain user communities adept at the creation and use of these materials'.

Although the term e-book was not used at the time, many reference books found their way onto CD-ROM in the 1980s, either marketed as databases or simply as '[title] on CD-ROM'. Examples include the *Oxford English Dictionary* on CD-ROM, Microsoft's Bookshelf (including *Roget's Thesaurus, Bartlett's Familiar Quotations* and the *World Almanac and Book of Facts* from the CIA), and the Bible Library on CD-ROM.

Also in the 1980s, university presses in the USA began to engage with the concept of electronic publishing, at least partly 'as a result of fears about the future of the [print] scholarly monograph' (Armstrong and Lonsdale, 1998: 18). The 1990s saw an international rise of both electronic reference and electronic fiction publishing, as well, in the middle of the decade, as the emergence of aggregators such as netLibrary, Questia and ebrary. Similar to e-journal aggregators, these companies acquired the rights to titles from many publishers and brought them together in subject collections which could be directly licensed from the aggregator. From about this time onwards, commercial publishers (Penguin, Routledge, Cambridge University Press, etc.) began making titles available electronically, and large numbers of individual e-books began to appear on the Web. Figure 2 shows a screenshot from a commercial publisher

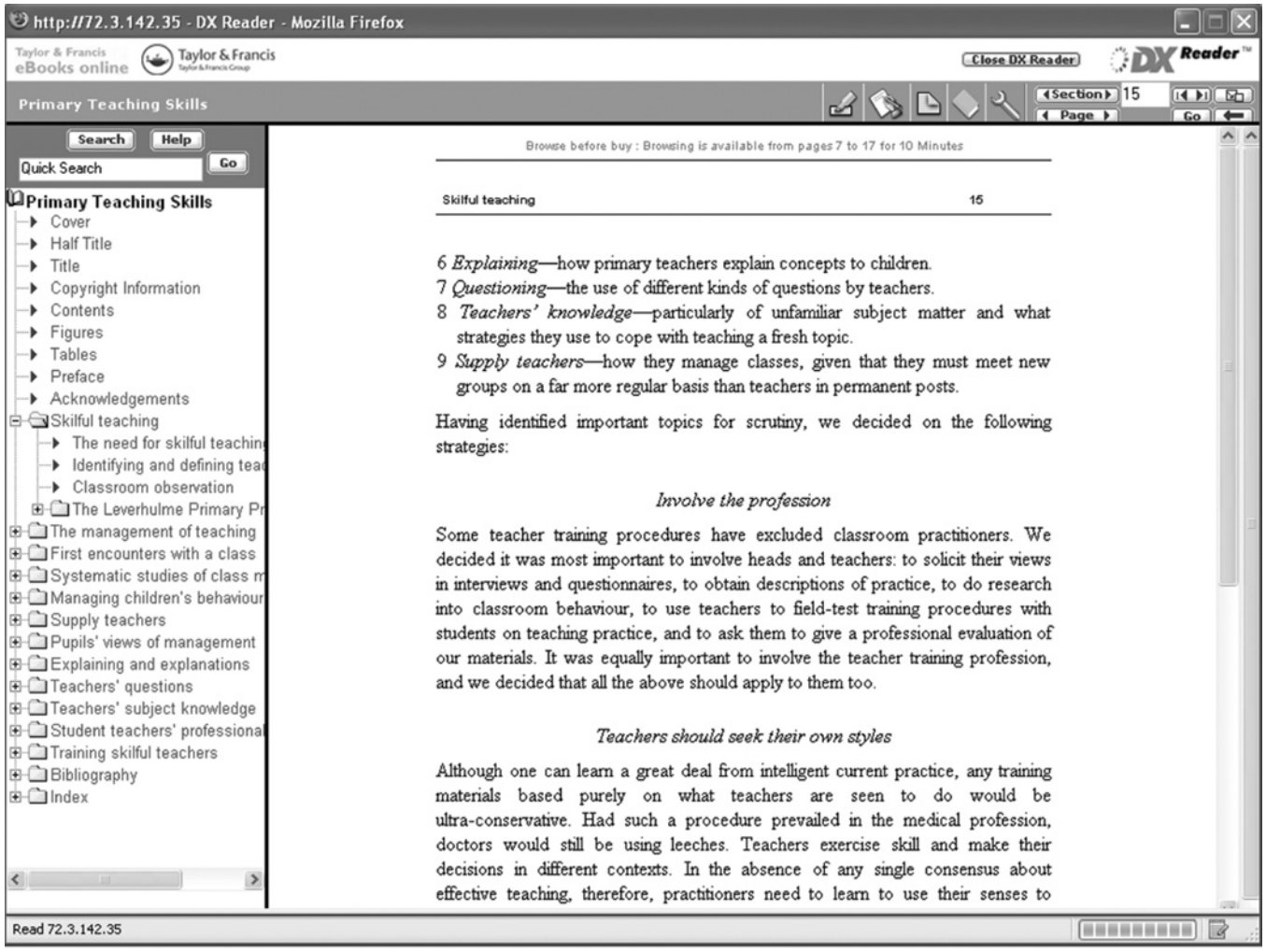


Figure 2. Screenshot from Taylor & Francis e-book interface

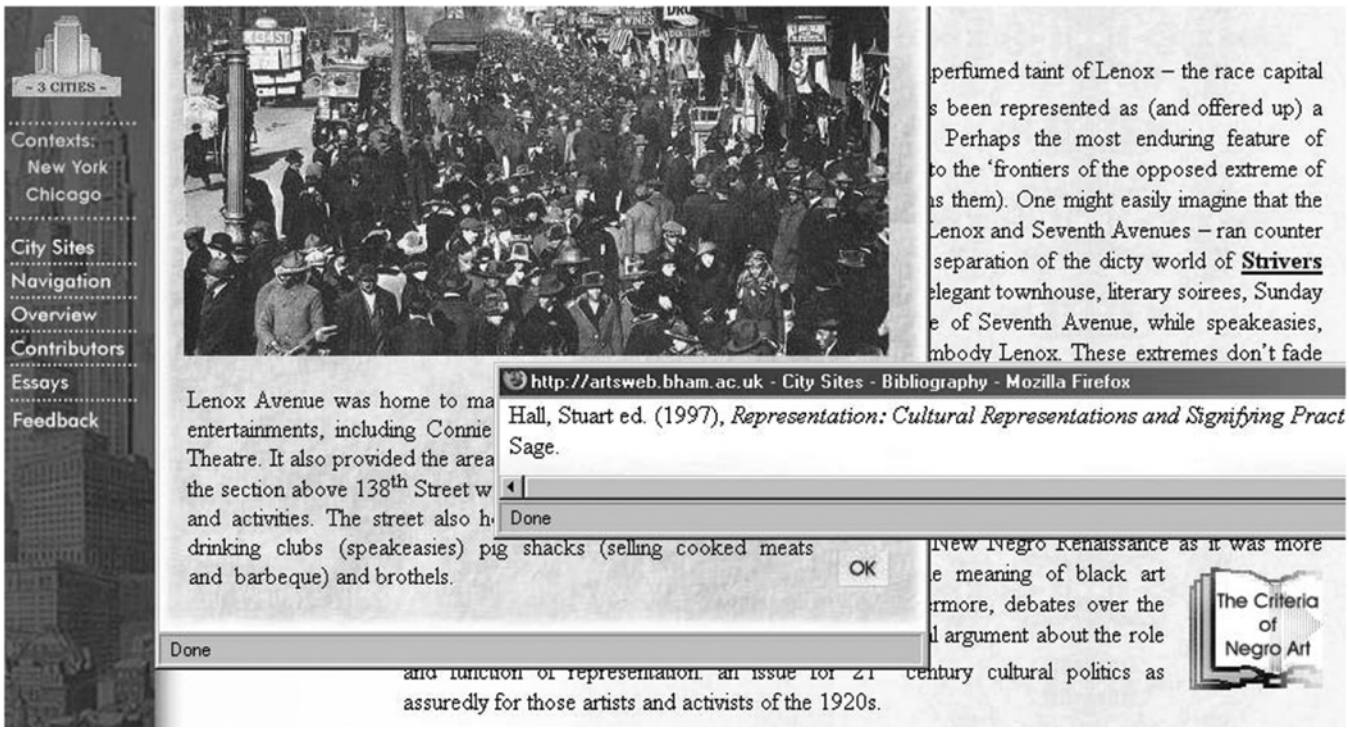


Figure 3. Screenshot from *City Sites*. Born-digital e-book with two overlay windows called up, and links within text to different sections of text ('The Criteria of Negro Art').

(Taylor & Francis) using an interface design typical of the early to mid-2000s, and Figure 3 shows *City Sites*, the stand-alone e-book mentioned above, which was produced as a part of research and development project and featuring a number of innovative features.

At the time of writing (2007), it seems clear that the e-book industry is burgeoning. Publishers and libraries both treat the medium very seriously and sectoral interest is evident in primary, secondary and tertiary education as well as in, for example, the legal and health sectors. In the UK, the Joint Information Systems Committee (JISC) of the Further and Higher Education Funding Councils runs an e-Book Working Group in support of its collection development team, which has made available over 25 e-book titles or collections to tertiary institutions. Additionally, consortia such as North West Academic Libraries (NoWAL) and South West Higher Education Libraries (SWHELs) have made available substantial e-book collections to their member institutions.

Wikipedia, whose definition of e-book was quoted earlier in this article, is itself a special sub-genre of e-book: the social or networked book. Even in the two years since the Eco interview previously quoted, social or communal computing (generally dubbed Web 2.0) has extended to the e-book, so that we can define a sub-genre of born-digital as the social or networked book. These may be either simply developed within a social network such as a weblog, allowing readers to access, and comment on, chapters as they are written (e.g. L. Lee Lowe's *Mortal Ghost* at http://mortalghost .blogspot .com/ 2006/07/ chapter- one.html), or developed from the outset with the intention of developing the text according to readers' input. Lawrence Lessig's second edition of *Code and Other Laws of Cyberspace, Code v.2* (http://codebook.jot.com/WikiHome) is written in a wiki, while McKenzie Wark's *GAM3R 7H30RY* (http://www.futureofthebook.org/mckenziewark/gamertheory/) uses specialist software developed for the purpose. The end product may or may not be published subsequently in conventional paper form by a conventional publisher. The Institute for the Future of the Book has released CommentPress, a free, open source theme for the WordPress blog engine designed to allow paragraph-by-paragraph commenting in the margins of a text.

Thompson (2005: 2) has suggested that in 2003

> Publishers who had invested heavily in the 1990s in experiments in electronic publishing were counting their losses and closing down divisions. Third-party players who had hoped to ride the new wave of electronic publishing were scrambling to find alternative sources of revenue

and speaks of an 'evident faltering of the ebook revolution'. However, in May 2007, the Association of American Publishers (2007) noted a significant rise in e-book sales – a '24.1% increase in 2006 at $54 million, with a compound growth rate of 65% since 2002'. In a slightly earlier debate on how publishers viewed Google Book Search, an Oxford University Press Vice President said in their official weblog, that

> What we publishers have come to realize is that Google and friends have opened up the world to our content by showing us that discoverability and access leads to interest and opportunity. Every major media company is now thinking they need to figure out their share of the digital space. (Schnittman, 2007)

In 2007, evidence of use and interest from universities, schools and public libraries, as well as of the availability of title lists from aggregators, publishers and book suppliers, seems to rebut Thompson's claims.

REFERENCES

AlphaGraphics (nd) *Resources & Support: Glossary of Terms,* URL (consulted November 2006): http://www.alphabetapress.com/resources/glossary.html?db_item=glossaryterm&glossary_item=B

Armstrong, C.J. and Lonsdale, R. (1998) *The Publishing of Electronic Scholarly Monographs and Textbooks. A Supporting Study in the JISC eLib Programme.* Report G5. London: Library Information Technology Centre.

Ashcroft, L. and Watts, C. (2004) 'Change Implications Related to Electronic Educational Resources', *Online Information Review* 28 (4): 284–91.

Association of American Publishers (2007) 'AAP Reports Sales of Books Totaled $24.2 Billion in 2006' [Press release], 22 May, URL (consulted May 2007): http://publishers.org/main/PressCenter/2006BookSales.htm

Bazin, P. (1996) 'Towards Metareading,' in Geoffrey Nunberg (ed.) *The Future of the Book,* pp. 153–68. Berkeley, CA: University of California Press.

Borgman, C.L. (2000) *From Gutenberg to the Global Information Infrastructure: Access to Information in the Networked World.* Cambridge, MA: MIT Press.

Bush, V. (1945) 'As We May Think', *Atlantic Monthly* 176 (1): 101–8.

Chomsky, N. (2000) 'Explaining Language Use', *in* Noam Chomsky, *New Horizons in the Study of Language and Mind,* pp. 19–45. Cambridge: Cambridge University Press.

Cox, A. and Mohammed, H. (2001) 'E-books,' *FreePint* 80, URL (consulted November 2006): http://www.freepint.co.uk/issues/010201.htm#feature

Crawford, W. (2004) 'The Library Stuff', *Cites & Insights: Crawford at Large* 4 (13), URL (consulted November 2006): http://cites.boisestate.edu/civ4i13.pdf

Dearnley, J., McKnight, C., Connal, M. and Towle, G. (2002) *The People's Network, Public Libraries and E-books: Specification for Baseline Intelligence-Gathering Work.* Loughborough: Department of Information Science, Loughborough University.

Education for Change Ltd, University of Stirling Centre for Publishing Studies and University of Stirling Information Services (2003) *A Strategy and Vision for the Future for Electronic Textbooks in UK Further and Higher Education: A Study Prepared for the Joint Information Systems Committee (JISC) e-Books Working Group.* London: JISC e-Books Working Group. URL (consulted May 2007): http://www.jisc.ac.uk/uploaded_documents/Annex_E_E_Textbooks_Strategy_final_report.pdf

Esposito, J.J. (2003) 'The Processed Book', *First Monday* 8 (3), URL (consulted November 2006): http://firstmonday.org/issues/issue8_3/esposito/index.html

Feather, J. and Sturges, P. (eds) (1997) *International Encyclopedia of Information and Library Science* (2nd edn). London: Routledge.

Front Row (2005) BBC Radio 4, 3 June, 19:20.

Garrod, P. and Weller, J. (2004) *Ebooks in UK Public Libraries: Where Are We Now and the Way Ahead.* An issue paper from the Networked Services Policy Task Group, UKOLN, URL (consulted November 2006): http://www.ukoln.ac.uk/public/nsptg/e-books/

Gold Leaf (2003) *Promoting the Uptake of E-Books in Higher and Further Education.* London: JISC e-Books Working Group, URL (consulted May 2007): http://www.jisc.ac.uk/uploaded_documents/PromotingeBooksReportB.pdf

Hillesund, T. (2001) 'Will E-books Change the World?,' *First Monday* 6 (10), URL (consulted November 2006): http://www.firstmonday. org/issues/issue6_10/hillesund/

Hodgkin, A. (2003) 'Forget the E-book, it's the E-library that Matters', *Information World Review* 196: 13.

Kay, A. (2005) 'The Dynabook Revisited: A Conversation with Alan Kay', *The Book and the Computer,* URL (consulted November 2006): http://www.squeakland.org/school/HTML/essays/dynabook_revisited.htm

Landoni, M. (2003) 'Electronic Books', in John Feather and Paul Sturges (eds) *International Encyclopedia of Information and Library Science*, pp. 168–71. London: Routledge.

Landoni, M., Wilson, R. and Gibb, F. (2000) 'From the Visual Book to the WEB Book: The Importance of Design', *The Electronic Library* 18 (6): 407–19.

Landoni, M., Wilson, R. and Gibb, F. (2001) 'Looking for guidelines for the Production of Electronic Textbooks', *Online Information Review* 25 (3): 181–95.

Lebert, M. (2004) 'Michael Hart: Changing the World through e-books', URL (consulted November 2006): http://pge.rastko.net/about/marie_lebert

Lynch, C. (1999) 'Electrifying the Book', *Library Journal, Netconnect supplement* 124 (17): 3–6.

Lynch, C. (2001) 'The Battle to Define the Future of the Book in the Digital World', *First Monday* 6 (6), URL (consulted November 2006): http://firstmonday.org/issues/issue6_6/lynch/

McCracken, R. (2004) 'Agreements, User Licences and Codes of Practice', in C. Armstrong and L.W. Bebbington (eds) *Staying Legal: A Guide to Issues and Practice Affecting the Library, Information and Publishing Sectors,* 2nd edn, pp. 122–39. London: Facet Publishing.

McKnight, C. and Dearnley, J. (2003) 'Electronic Book Use in a Public Library', *Journal of Librarianship and Information Science* 35 (4): 235–42.

Malama, C., Landoni, M. and Wilson, R. (2005) 'What Readers Want: A Study of E-Fiction Usability', *D-Lib Magazine* 11 (5), URL (consulted November 2006): http://www.dlib.org/dlib/may05/wilson/05wilson.html

Midgley, S. (2002) 'The End of Books?', *The Guardian,* 9 April, URL (consulted May 2007): http://www.guardian.co.uk/Archive/Article/0,4273,4389510,00.html

Morgan, E.L. (1999) 'Electronic Books and Related Technologies', *Computers in Libraries* 19: 36–9.

OCLC (2004) *2004 Information Format Trends: Content, Not Containers.* Dublin, OH: OCLC, URL (consulted November 2006): http://www.oclc.org/reports/2004format.htm

Ohene-Djan, J. and Fernandes, A. (2003) 'Personalising Electronic Books', *Journal of Digital Information* 3 (4), URL (consulted November 2006): http://jodi.ecs.soton.ac.uk/Articles/v03/i04/Ohene-Djan/

Olson, R. (1988) 'Sidebar,' in Hennen, T.J. Jr. (1988) 'Public Librarians take a Cool View of the Future: Leadership, not Technology, may be the Issue', *American Libraries* 19 (5): 390–1.

Pace, A.K. (2004) 'Technically Speaking. E-books: Round Two', *American Libraries* 35 (8): 74–5.

Peurell, E. (1999) 'Print on Demand in Sweden: Four Projects and their Products, Problems and Prospects', in J.W.T. Smith, A. Ardö and P. Linde (eds) *Electronic Publishing '99: Redefining the Information Chain – New Ways and Voices. Proceedings of an ICCC/IFIP Conference,* University of Karlskrona/Ronneby, Ronneby, Sweden, 10–12 May, pp. 179–86. Washington, DC: ICCC Press.

'Publishing, history of' (nd) *Britannica Concise Encyclopedia* from Encyclopædia Britannica Premium Service, URL (consulted April 2007): http://www.britannica.com/ebc/article-28596?tocId=28596

Rao, S.S. (2003) 'Electronic Books: A Review and Evaluation', *Library Hi Tech* 21 (1): 85–93.

Rao, S.S. (2004) 'E-book Technologies in Education and India's Readiness', *Program* 38 (4): 257–67.

Sawyer, S.K. (2002) 'Electronic Books: Their Definition, Usage and Role in Libraries', *LIBRES: Library and Information Science Research* 12 (2), URL (consulted November 2006): http://libres.curtin.edu.au/libres12n2/

Schnittman, E. (2007) 'Playing Nice With Google', *OUPblog: the official blog of Oxford University Press,* 22nd March, URL (consulted May 2007): http://blog.oup.com/2007/03/playing_nice_wi/

Summerfield, M. (2001) *Online Books: What Roles Will They Fill For Users Of The Academic Library?* New York: Columbia University Libraries, URL (consulted May 2007): http://www.columbia.edu/cu/libraries/digital/texts/paper/index.html (no longer available).

text-e (2001) *text-e: Screens and networks: Towards a new Relationship with the Written Word. A Virtual symposium,* URL (consulted November 2006): http://www.text-e.org/

Thompson, J.B. (2005) *Books in the Digital Age: The Transformation of Academic and Higher Education Publishing in Britain and the United States.* Cambridge: Polity.

UNESCO (1964) *Recommendation concerning the International Standardization of Statistics Relating to Book Production and Periodicals,* URL (consulted November 2006): http://portal.unesco.org/en/ev.php-URL_ID=13068&URL_DO=DO_PRINTPAGE&URL_SECTION=201.html

Van Dam, A. (2001) 'Expert Declaration to United States District Court', Southern District of New York, 26 February 2001, URL (consulted November 2006): http://www.rosettabooks.com/casedocs/RH_Van_Dam.pdf

Whalley, W.B. (2006) 'e-Books for the Future: Here But Hiding?', *Ariadne* 49 (October), URL (consulted November 2006): http://www.ariadne.ac.uk/issue49/whalley/

Wiesner, K.S. (2003) *Electronic Publishing: The Definitive Guide*, 4th edn. Amherst Jctn, WI: Hard Shell Word Factory.

Wikipedia, the Free Encyclopedia. 'E-book', URL (consulted November 2006): http://en.wikipedia.org/wiki/Ebook

Wilson, T.D. (2000) 'Recent Trends in User Studies: Action Research and Qualitative Methods', *Information Research* 5 (3), URL (consulted November 2006): http://informationr.net/ir/5-3/paper76.html