

# “Publishing Trends in the Next Century”

Presentation to Local Publishers & Book Suppliers, Tuesday 13 January 98

On the Occasion of the Re-Opening of Singapore National Library at Stamford Road

Presented by William Claxton, Managing Director, iMedia (S) Pte Ltd.

## **Synopsis**

The future is online, not just for book sales, but delivery as well. Already widely used within the publishing industry, for proofing and prepress, the popular ‘Acrobat’ document format is proving to be ideally suited for online delivery of literary works. Acrobat provides faithful rendition of text, font, color and images, both in Latinized and double-byte languages.

Delivery of book contents online is potentially a multi-billion dollar market which, despite recent advances in desktop publishing and Internet, is still relatively untapped. There are a number of reasons for hesitancy on the part of book publishers, including the cost of conversion to electronic formats, and the lack of a proven revenue model for online delivery.

Perhaps academic books are more likely than others to overcome these obstacles. And a business model based on ‘free to read – pay to print’ would appeal to both librarians and book publishers, while delivering real convenience to users. There are a number of opportunities for testing this approach in Singapore, where strong infrastructure support exists for both IT and education, and where 80% of the publishers deal in academic texts.

## **Introduction**

The deployment and widespread use of Internet as a global information system has dramatically reduced the production and distribution costs associated with information dissemination. However it has brought into sharp focus the problems associated with incompatible formats for information content and catalog data. The work of standards bodies such as the Singapore National Information Infrastructure Standards Committee is to assure that incompatibilities in content format and/or access protocols do not inhibit the emergence of an information-based economy. These efforts show genuine promise for online publishers.

## **What’s in Store?**

In a vibrant information-based economy, information stores would contain a variety of documents or ‘information objects’. These would range from simple textual pages to complex graphical documents and even movies; one format doesn’t address all the needs. Separate catalogs would provide descriptive information about these objects, and would be searchable independently of the contents themselves. Exchange standards would help users determine which information object types are widely supported, robust enough for commercial use, suitable for evidentiary purposes, and so on. Similarly, standards will help publishers protect their investment in producing contents for online delivery .

While HTML and SGML have each been tried, both formats come up short in many respects; notably they lack the ability to faithfully replicate a printed edition. From the publishers’ perspective, these formats are expensive to create, and difficult to control (potentially resulting in one file for each page).

Today there is a genuine need for a 'final form' page fidelity standard that retains the page layout and pictorial information needed for the delivery of complex documents, containing a combination of text, graphics, images, and/or tables. Reports from the digital library community indicate that the number one request from users of online catalogs is 'show me the book.' The format of such books should be device independent, efficiently stored and transmitted, and meet the needs of a diverse population of document publishers and consumers. The format should also be suitable for routine archival use, for example allowing images and text to be extracted at any future date.

### **Who Minds the Store?**

But it is not just the publisher and consumer who will interact online. In any particular information exchange transaction, there are likely to be 3 parties: (a) an information provider, (b) an information consumer, and (c) an intermediary or 'broker'. It is the challenge of the information broker to provide the most comprehensive, descriptive and easy-to-use content catalogs. This intermediary could easily add value to a simple directory service by negotiating rights and charging for access, for example, by selling reprint rights.

It is in fact the information brokers who will set standards for document exchange. There will be a variety of formats required to support movies, audio and catalog exchange. Taking into consideration the peculiarities of delivering literary content online, the broker is clearly going to prefer literary content that is delivered in a single standard format.

### **Is there a Business Model?**

Once the choice of document format is agreed, publishers, brokers and consumers must decide on a revenue model. Commercial delivery of literary works online requires a mechanism for revenue generation to authors, publishers and information brokers. It must be cost effective for consumers, and provide ability to preview before buying.

There have been some attempts at document encryption and charging for access, but most are inconvenient and expensive to implement. IBM's much publicized 'Cryptolope' technology was recently abandoned because of cost and complexity. One approach that looks likely to be viable is based on the notion that it is not distribution but reprint rights, which should be the basis of charges. Books offered online using this approach would be free to read, but consumers would pay a modest fee for reprints, perhaps as low as photocopy charges.

### **What Type of Publications will Sell?**

There are many potential applications well suited to this approach, for example, the dissemination of tender documents. Tenders are bulky, and physical distribution requires centralized administration and revenue collection. The consumer often requires only a small portion of the document that is relevant to their needs. Faithful rendition is an issue due to legal interpretation of the document, especially where illustrations are concerned.

Another possibility is online delivery of academic texts and references. Academic users traditionally rely on libraries to obtain materials needed to supplement the basic text, and often resort to photocopying. All over the world, students have shown a willingness to pay for photocopies of scarce reference materials. And they need these services at all hours of the day and night.

### **What's the Price Point?**

To become a viable service, the price of such reprints should be equivalent to what is now paid for photocopies. Incidentally, this varies from country to country. So charging models will

likely need to be customized at specific online distribution points, with access limited by membership or similar commercial mechanisms.

The challenge in respect to collecting payment is that the amounts are quite small, just pennies per page. A number of alternatives presently exist to handle such payments, including cybercash, debit-based smart cards, and debit accounts (prepaid with a credit card or cash).

### **The Outlook for Controlled Reprints**

There is reason to believe that the controlled-reprints approach would be suitable for both academic and non-academic publishing as well. Controlled reprints will provide a cost-effective alternative to photocopying, offering unprecedented convenience to consumers. Both publishers and libraries will be able to earn revenue from such a service, in marked contrast to the current situation, in which reprint revenue goes only to the photocopy vendors.

The main challenges ahead relate not to document format and electronic commerce, but to providing sufficient security so that publishers will be willing to release contents online. As long as the reprinting process can be tightly controlled, there is little danger of a complete and redistributable electronic document falling into the hands of a competitor. In any case, someone who is out to pirate from the publisher can scan and convert a printed work into a distributable electronic edition. So high levels of security are counterproductive if they interfere with convenience, or fail to leverage the advantages of document standards.

### **Adobe Acrobat**

Adobe Acrobat is rapidly being accepted as the medium of choice for delivery of literary works online. Acrobat PDF is a graphical page-description language related to PostScript, and is based on Adobe's Imaging Model that is a robust way to represent graphics, text, and images in a device-independent and resolution-independent manner.

#### **How Does it Work?**

Adobe's Acrobat Portable Document Format (PDF) allows preformatted pages in full color to be interchanged via intermediate storage or over a network. These pages can originate from almost any application, independent of the hardware platform, operating system or even the fonts installed on the computer.

The portability of graphically rich documents could not be ensured without first solving the font distribution problem: incorporating embedded fonts for every document is cumbersome and prone to copyright violation. Adobe, the developer of Postscript, solved the problem by adapting Display Postscript to become a published file format eventually called 'Acrobat', and using their Multiple Master Font technology within these files. Font characters from a generic font are fitted on the fly to the font metrics of the original font used for any fonts that are not available.

#### **Is Acrobat an International Standard?**

The Acrobat Portable Document Format or 'PDF' was originally put forward by Adobe Systems Inc as a format for document exchange in 1993, and it has since garnered wide support from US Department of Commerce, NIST, and others. One of the attractive features of Acrobat is that it is well suited to a wide range of publishing and archiving activities. For example, Acrobat is gaining support as a format for legal evidence in several countries. The US Court of Appeals has recently adopted PDF as a standard for legal documents.

As part of the work of the standards committee that I co-chair, PDF has recently been put forward as a Singapore standard for platform-independent exchange of graphical documents, especially where faithful rendition is important. And in Singapore, a PDF document can already be submitted as legal evidence. All pleadings to the Subordinate Courts are routinely submitted in Acrobat format.

Not only is it becoming a defacto standard in some application areas, according to Adobe, the company is seeking to have PDF become a formal ISO-approved standard. Adobe hopes to announce this by mid-1998. Getting ISO endorsement will help PDF meet customer requirements for ISO9000 compliance and year 2000 issues in their document management applications. Adobe is also working with relevant authorities in the US to have PDF accepted as a national standard in for prepress applications.

### **How Does the Publisher Create a PDF file?**

The easiest way to create PDF files is to print them directly from within the originating application, for example, a word processor. This print-to-softcopy facility must be installed using the Adobe Acrobat product (or a licensed version of that product). Another means of creating PDF is conversion from PostScript to PDF, which can be carried out using software from Adobe and others; (a reverse process enables printable PostScript files to be recovered from PDF files).

Yet another means to create PDF files is to scan them from paper, and using Acrobat software, recognize and index the text. PDF is thus one of the very few formats that can provide a common representation of contents that originate on the computer and those which originate on paper. My company has been investigating the archival characteristics of such captured documents, for both National Library and National Archives of Singapore.

### **What Does it Require to View PDF Files?**

Separate, and freely distributed, viewer applications allow consumers to view PDF pages on almost any platform. There are also shareware browser extensions to allow PDF documents to be viewed within either Internet Explorer or Netscape's Navigator/Communicator. Functionality of these viewers includes panning, zooming, scrolling, skipping pages and traversing the document using hypertext links or a variety of navigational buttons.

### **What Makes Acrobat Special?**

One of the first impressions when using PDF is that text is handled in a very special way. When zooming into or out of a page, the text characters do not rasterize, they are vector based. And when a search is performed, the indicated word or phrase is highlighted on the page, no matter what typeface was used. Thus, a PDF document combines the textual properties of a native application file with the image properties of a scanned page.

A key feature of PDF files is the embedded document markups, not only hypertext links, but thumbnail icons of pages, chapter outlines, article threads and user-definable annotations. Chapter outlines enhance the useability of a document, for example by adding section headings or tables of contents. Thumbnail icons of document pages ensure fast browsing and random access. Article threads let the user step through a long passage, following text from page to page. Page annotations act as electronic Post-Its, and may be user specific. Hypertext links can be used to navigate within a document or to another document, locally or via Internet.

Markups are usually added manually after the PDF file has been created. This is done within a full-featured PDF viewer, such as Adobe's Exchange product, or using any of a growing

number of PDF post-processing tools from other vendors. Alternatively, the markups can in some cases be passed down from 'front-end' document-processing applications during conversion. Some of the document markups are very specialized, and a number of vendors have announced online forms-applications using PDF.

These features have made Acrobat PDF a popular choice for document exchange among publishers, especially those needing to deliver the same document in print, CD-ROM and via Internet. In fact, a large number of book and journal publishers have already chosen PDF to put entire literary works online. HTML and SGML are not well-suited for long literary texts, because of their relatively high production cost and lack of faithful rendition characteristics.

## **Commercial Deployment**

Why aren't more books already online? The answer is that publishers are unwilling to invest in creating one-off electronic editions which cannot be leveraged in multiple markets, and which are unlikely to return sufficient revenue to justify the cost of conversion. Furthermore, electronic commerce is only just beginning to become accepted by publishers, even for the sale of physical books. Finally there is the perceived danger of alienating the publisher's existing business partners, the distributors, channel resellers and retailers.

### **The Big Picture**

Publishers are certainly looking at a bigger picture than what the consumer sees. Publishers need to establish a digital workflow which allows them to edit, layout, proof, prepress, publish, disseminate and archive an editorial product which can be resold in a wide variety of increasingly fragmented markets. They need to make decisions about electronic document format in light of professional standards for editorial review, photographic reproduction, color accuracy, and print quality, which don't exist in today's multimedia or Internet products.

Acrobat is only one component of the solution which publishers seek. Acrobat addresses the entire workflow of publishing for both print and electronic distribution. But Acrobat must be complemented by database technologies that support large-scale document management, and by electronic commerce technologies that securely collect and distribute revenue.

### **Can Digital Copies be Secure?**

Document security is probably the number one concern of publishers, before they embark on commercial deployment of online books. Publishers will be concerned about the potential losses that might arise if a single electronic copy of one of their books is freely distributed worldwide via the Internet. There are, of course, many components of document security.

Online documents should be easy to control. A large number of HTML files becomes a problem in terms of document management. Yet an entire book can be easily delivered in a single Acrobat PDF file, and can still be served up a page at-a-time on the internet (just like HTML pages). Furthermore, publishers can set document attributes which control the ability of a PDF file to be copied, printed or annotated by the consumer. These attributes can also be enabled or disabled dynamically, for example in a pay-for-reprints service.

Online documents must contain indelible copyright information. An Acrobat document can easily do this today using embedded bibliographics and annotations. Not only can the document contain author, publisher, ISBN number and related bibliographic references; the document can be annotated so that any reprint contains a specific copyright notice in the footer of each page. It is difficult, but not impossible for another person to alter this information.

Online documents should also contain an indellible electronic signature. A digital signature can be injected into an Acrobat document to uniquely identify authorship. Pirates who create electronic documents by scanning printed ones would not have access to this digital signature, and thus could not make 'authentic' copies.

### **Reach Out and Touch the Book**

Obviously, digital access need not include possession. Online documents should be viewable without passing the entire document to the consumer. This makes them more secure and easier to browse (eg- just downloading the page being viewed). Acrobat documents achieve this using page-at-a-time technology installed on a Web server, similar to the way HTML pages work. While individual pages are cached on the consumer's computer during browsing, and could be captured from the screen itself, it will be difficult (and inconvenient) for consumers to copy an entire book in this manner. They will never get an exact replica.

Reprinting of online documents needs to be tightly controlled. The ability to copy a document that is sent across for reprinting is a real threat to publishers. To counter this treat, our company and others are developing technology that tightly controls the printing process, essentially sending the document directly from the merchant's server to consumer's printer.

### **Other Commercial Concerns**

Once the 'hard issues' of document security are addressed, publishers will have a myriad of other commercial concerns. For example, what affect will delivery of online books have on sales of printed editions? What are the arguments in favor of selling books online versus actually delivering them online? Will authors be willing to grant electronic publishing rights, and so on? I am confident that favorable answers will be forthcoming, and publishers will find little reason to hold back from online delivery.

It is interesting to note that from a cost of materials perspective, reprinting online editions turns the traditional publishing model upside down. The consumer actually bears the cost of paper and ink (plus machine wear and tear), while the publisher earns revenue for purely intellectual property rights (ie- licensing reprints). This is referred to as 'distribute and print' rather than 'print and distribute.' This is a cost reduction publishers will have a tough time ignoring.

### **Potential Local Applications**

As mentioned, both tender documents and academic texts seem promising markets for pilot implementations of the controlled reprints business model. There are a number of factors that make Singapore an ideal environment for commercial trials of this technology:

- high-bandwidth testbed available in Singapore ONE
- strong support for standardization within the National Information Infrastructure
- excellent Governmental support for IT and Education
- 100% of university students have access to Cash Card technology for payment
- 80% of local and international publishing operations focus on academic publishing
- Singapore's regional position as a logistics center, with a strong local presence of most international publishers
- English is the common language in education and business

In this environment, it is entirely feasible for a group of publishers to test the controlled-reprints business model, without fear of their publications being made accessible outside of Singapore.

## **Academic Requirements**

A common assertion made about online publishing schemes is that consumers do not want to read from the computer monitor. While it is arguably false (people do in fact read lots of text from their computer screens), the issue is relatively moot if the consumer is able to choose between reading onscreen or in print. Faithful rendition assures that the difference in display vs print is not qualitative, but economic: it is simply the cost and potential inconvenience of reprinting. This issue will always be a trade-off that can be measured in economic terms.

Thus, while students definitely want access to reference texts using a computer, they won't necessarily read the documents on a computer screen. For some there will be concern about eyestrain, and for others, it will be a documentary or aesthetic consideration that encourages them to pay for reprints. It has been observed from the usage pattern of online information that students prefer to print out relevant pages, to mark them up with a highlighter, compare them with a friend's assignment, and retain them for exam preparation.

### **Unique Requirements**

In addition, the academic community has some unique requirements when it comes to online delivery. Firstly they want consistent pagination, so that assignments can be given with reference to page numbers in either the printed or online edition. Consistent pagination is a feature related to faithful rendition, so technology such as Acrobat will be appropriate.

Academic institutions need to ensure that everyone has equal access to online materials. Publishers should not be seen to be profiting at the expense of those students who can't afford to pay for reprints, or who have no computer access.

The payment method must be accessible and convenient to all students. Fortunately, in this country, all the students at National University of Singapore and Nanyang Technological University already carry Cash Cards. These tokens are ideal for penny transactions via Internet. A home reader for Cash Card today costs only \$50, and public reader can be found all over the campus. In other countries, students must be given accounts that they can use to pay for access to electronic editions, which they could top up with cash payments in advance.

### **The Untapped Potential of Online Delivery**

One of the great advantages of online delivery, in academic terms, is the ability to provide for copious searching. A student can not only search within a publication; they can search across entire libraries of information. Searches can be constructed based on bibliographic data, or they can be based on the occurrence of a particular word or phrase in the text of the work.

A further advantage, only now beginning to be tapped, is the ability of online editions to be cross-referenced in citations and other links. This requires a sufficient mass of publications to be available in electronic editions, and the cooperation of libraries and information brokers in cross referencing. Standards for such links need to be established. But it is conceivable that a user in the future could link from one text to another, as easily as she surfs the web today.

## **Conclusions**

The cost of newsprint and other paper products have risen dramatically in the last 6 months. Perhaps this is just cyclical, or perhaps it is related to the destruction of forestry which proceeds unabated throughout the world, especially in tropical countries. And with the current global economic downturn, the cost of imported paper is being driven even higher due to currency imbalances. Many Asian newspaper publishers have already closed down because advertising revenue is unable to keep pace with rising costs.

In this context, it seems unlikely that the cost of paper books can maintain their current low pricing, and equally unlikely that publishers will for long resist the temptation to make money from online delivery of their vast literary holdings. Many of a publisher's out-of-print collections could easily be reissued online, for example. Production problems need not constrain these projects; even if layouts are not available in electronic form, pages from old books could be scanned for conversion into an electronic edition.

### **Where To Begin?**

While there are commercial issues to be resolved, such as the security of documents being distributed via Internet, these problems are not especially difficult to resolve. Publishers and information brokers will overcome these commercial obstacles by addressing the convenience of consumers, rather than nurturing false-hope for total document security. It just won't happen that electronic editions will be completely safe from piracy.

Similarly, pay-per-view schemes are unrealistic. It will not be easy to sell book reprints online if the consumer is unable to preview. Like in a bookstore or a library, browsing must be free. On the other hand, consumers are unlikely to read entire books online at today's display inadequate display resolution. A computer displays at less than 100 dpi, while a laser printout is at least 300 dpi, a 3 to 1 improvement.

So there is a real revenue opportunity today in charging for reprints. This will be attractive to publishers, because it is a pure licensing activity, and because it shifts the cost of paper and ink to the consumer. It will also be attractive to consumers such as academic users, who rely heavily on photocopying today to meet their information consumption requirements. Singapore is obviously an excellent environment to nurture such businesses.

### **A Paperless Future?**

In the more distant future, display resolutions will increase to rival today's laser prints, and bandwidth will be less of a constraint. Students will grow up using online books for much of their reference work, and will consider it natural to subscribe to personalized magazines online. Books will be brought to life onscreen with voice-over readings or commentaries, and pictures will be augmented with movie clips. Portable document viewers, some that emulate bound paper books, will become commonplace. Learning will be a lifelong activity.

In that context, we can expect to see the distinction between libraries and bookstores blur. Both will be brokers offering free access to viewing of books, and controlled access to reprints. Both will need to become competent in rights negotiation, electronic commerce and royalty distribution. Bookstores will deal in a wider range of information products, cater to obscure tastes, and may handle a lot of goods of questionable authenticity and origin.

Libraries will still specialize in sophisticated indexing and cross-referencing services. But it is conceivable that the main advantage to getting something from a library over getting it in a bookstore will be the guarantee of authenticity, and assurance that the publisher and author are getting a share of revenue. Arguably, this is exactly the opposite of today's environment!

Business models will change in that personal publishing is likely to become a real option for many authors. Librarians will become information consultants, perhaps hanging out their shingle for a fee, and they will help the consumer track down the canonical editions of particular works, among other services. Some consumers will pay a premium for highly filtered

and summarized information, or just a tip on what's worth reading! These materials may become the electronic equivalent of the Reader's Guide.

While the world will not be entirely paperless, many of our notions of what constitutes a book will be altered radically. HTML editions will be the truly softcover editions, without authenticity and looking slightly different on every computer. PDF editions or their typographical equivalent will provide a more substantial document, with embedded digital signature, linked table of contents, article threads, thumbnails, and other standard markups built in. And for the discerning collector, cloth-bound paper books will still be a worthy investment.

### ***About the Author***

William Claxton is Managing Director of iMedia (S) Pte Ltd, a leader in electronic publishing solutions and products, based in Singapore. Mr. Claxton is also co-Chair of the National Information Infrastructure Subcommittee on Standards for Information Sources and Exchange.

For further information, check out: <http://www.paperless.com.sg>