About the Author(s)

Author: Dr. Bob Jansen
Position: Managing Director and Chief Scientist, Turtle Lane Studios Pty Ltd
Lead Researcher, National Archives of Australia, Digital Preservation Project
Affiliated Organisation: Turtle Lane Studios Pty Ltd, National Archives of Australia

Short Biography:

Dr Bob Jansen is the Managing Director and Chief Scientist at Turtle Lane Studios Pty Ltd, an innovative company developing advanced communication technologies involving synchonised video, audio, images and text. He is also the lead researcher for the National Archives of Australia Digital Preservation Project, Agency to Researcher (AtoR).

Dr. Jansen has over thirty years experience in the Information technology industry. He has lead teams and worked in artificial intelligence, database technology, advanced information systems, electronic publishing, multimedia. He is the author of over twenty papers published in peer reviewed conferences and journals.

Dr Jansen has worked in private, public and research sectors in Australia and internationally. He held the position of Principal Research Scientist at CSIRO and has consulted to medium to large Australian organisations such as the Australian Broadcasting Commission, the NSW National Parks and Wildlife Service, The National Library of Australia and Toyota Motor Corporation. He was instrumental in the development of the Events on Line architecture for delivering synchronised video, audio, images and text to user workstations, which he commercialised into the Testimony Software system.

Contact Details

Dr. Bob Jansen
Turtle Lane Studios Pty Ltd
PO Box 26
Erskineville NSW 2043, AUSTRALIA

Email: bob.jansen@turtlelane.com.au
Phone: +61- (0)414 297 448

The Future of the Book: Format & technology

Main Description

This paper argues that the long history of paper book technology provides lessons for electronic book technology. Paper has a five-thousand year history and to expect a forty-year old technology to displace it in the commercial market place is optimistic, to say the least.

Paper books are based on sophisticated mechanisms for authoring and communicating information. Electronic technologies have even more sophisticated ways of doing the same, but we are unsure of how to author to such sophistication in any engaging way. The inclusion on multi-media formats, such as video and audio, exacerbates the problem because we reach the position of overloading the reader. How can we author an engaging electronic book?

Dot-point authoring rules the waves. Yet dot-points are not very engaging for readers. The narrative style has been lost in electronic books. It is even impossible to design an electronic book in any permanent way because display technology forces the content to be chunked in ways unforseen by the author.

This paper argues that commercial viability of electronic books rests on breaking the nexus on format and technology. The format and technology of electronic books must be guaranteed for a significant period of time. Consumers will not put up with the *status quo*, with no one in the publishing value chain able, or willing, to guarantee longevity.

The paper introduces the Testimony Software system as an example of modern electronic book technology and shows how complicated things can get.

Short Description

The future of the book is not necessarily secure. Although based on a long history, paper-based publishing is under attack from information technology, especially for younger readers. However, electronic book technology is difficult to maintain over time and unless resolved, this may be its death knell.

This paper compares paper and electronic book technology and concludes that unless electronic books break the nexus with format and technology, widespread acceptance, and hence commercial viability, may be a long way off.

Keywords

Electronic publishing, testimony software, event on line, turtle lane studios pty ltd, national archives of australia, electronic books, format, technology, artists books, in formation chunking, dynamic chunking, reading, authoring, ator
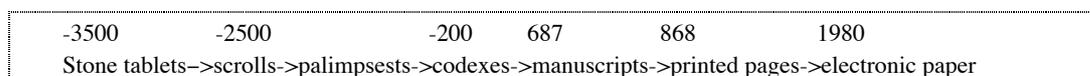
1   Backgound

The future of the book has been under discussion for many years now, since the introduction of the paper-less office concept in the mid 70's. However, like the paper-based office, the book has a rosy future. That does not mean, however, that the book will remain in its present state. In fact it seems fair to say that development of book technology has been more rapid over the last few years as convergence of media reaches new peaks and multimedia technology has become more robust and capable of delivering viable commercial 'books'.

Commercial viability of a non-paper-based book does depend on several developments, some of which are not being considered at present. The value chain of books is well known, yet electronic book publishers may be ignoring the two crucial elements in this value chain, the author and the reader, to their overall detriment. Authoring electronic books is not the same as authoring paper books. In a similar way, readers are not likely to buy electronic books that can not be guaranteed to work for any significant time.

This paper discusses some of these important issues. It looks at book technology from a historical perspective, compares the characteristics of books and electronic books, and discusses authoring electronic books. In providing an example of an electronic book developed by the author using Testimony Software, developed by him at CSIRO and Turtle Lane studios Pty Ltd, we show the major problem involved with electronic books: the complexity of the environment which must remain fully functional for the electronic book to continue to function.

The discussion ends with a description of a project at the National Archives of Australia in developing a digital repository for electronic objects. Lessons learned in this project are applicable to the longevity of any electronic product, be it an electronic book, digital images, digital video, or a combination of these technologies.

2   History of 'Book' Technology

-3500	-2500	-200	687	868	1980
Stone tablets–>scrolls->palimpsests->codexes->manuscripts->printed pages->electronic paper					

The technology behind books has been long in development. What started as engravings on stone, bone or metal, grew through scrolls and palimpsests, codexes, illuminated manuscripts into moveable metal type. This in turn led to word processors through the application of information technology and the rise of the desktop computer. Developments in IT led to the converging of disparate media, which became the multimedia technology and portable document format (PDF) and the E-Book format.

This timeline of development must be seen in perspective. The oldest know engravings on bone are dated at about -3500BC. So even allowing for inaccuracies in such dating, books are based on a development lifetime of approximately 5,000 years. This expanse of time has allowed the evolution of a sophisticated technology with an ever-increasing ability to communicate more and more complex information. Accompanying this increasing ability is, however, a decreasing stability. Rock is a very long-lived material, much longer that even the best quality paper today. As for electronic technologies, their lifetimes are currently measured in terms of years, despite the claims of disk manufacturers who claim hundreds of years. We are all too aware of the inability to read an electronic document, even in extreme cases, a week after their creation.

The web site, http://www.xs4all.nl/~knops/timetab.html, gives a fascinating view of the development of book technology.

So, books have had a long development that has produced a sophisticated medium for communication and recording. They are based on a stable technology because the pace of development has slowed down as technology development has slowed down, ignoring electronic book developments.

The one area that shows interesting developments of this technology is that of artist books.

Artists have a long history exploring ways of communicating via books. *In The Century of Artists Books*, Drucker (1995) provides a fascinating glimpse of artist books through the ages and demonstrates how long artists have been experimenting with this format. Keith Smith has a number of interesting publications. *Text in the Book Format* (Smith, 1995) and *Structure of the Visual Book* (Smith, 1984) provide remarkable examples of different books developed by artists. Suffice it to say, artists are probably the only ones at present exploring differing mechanisms of paper-book technology for communicating to readers.

3   History - Electronic 'Books'

1960	1970	1980	1990	2003(-ish)
Text editors -> Word Processing -> PDF -> Multimedia -> Interactive Digital IT				

Electronic books, on the other hand, have a much more recent nascence. IT technology has only recently reached the sophistication whereby it can replace paper-based technologies for communication, and even then only in some specialised areas, such as reference books. At best we can conclude that electronic books have been under development for about forty years.

They can provide sophisticated less-restrictive mechanisms for communication yet are more restrictive than books in their useful environment. The ability to represent information is almost unlimited, borrowing from television, film, audio, and written forms. The characteristics of the technology dictate a desktop environment for most products with true mobility a far cry from their competitors, paper-based books.

Yet they are built on an unstable technology, whose pace of development is ever increasing as information technology races to the future. Our grappling with the convergence of technology and media means that this pace of development will not slow down in the near future but will most likely increase to the detriment of the early adopters.

4   Characteristics of Books

So, what are the characteristics of books?

Well, they are readable for a very long time, depending on the material used in their creation. Archival paper is currently rated for hundreds of years. Acid free material enables books to retain their physical structure with little degradation for hundreds of years. Examples of books from the early years still exist as can be ascertained by visiting any major museum.

To read a book required no intermediate technology. This means that we can read a book almost anywhere, including underwater, if appropriate water-proof material is used in its development. Outer space presents no problems either, given appropriate material. If there is enough light, our books can be read!

There are internationally agreed sizes to the sizes of pieces of information presented in books. A1/A2/A3/A4/A5/Fullscap/US Letter are all common formats, which dictate the chunking of our information[1]. Our education system has made us all extremely efficient in handling these chunk sizes, to the extent that our scanning facilities appear to enable us to take in a whole page at a time without reading at all.

Books provide very good support for serial access. We all turn the pages to follow a narrative and happily start at page one and end at the last page with a feeling of satisfaction with an enjoyable read.

Books provide limited support for random access. Page numbers combined with tables of contents or back-of-the-book indexes provide for some random access but only facilitate access according to predetermined access points. It is interesting to note that neither tables of contents nor indexes would work without page numbers. Yet page numbers were implemented to assist the printer in reassembling books if dropped

Books also support sophisticated reader access. Isaac Newton is renowned for folding the corners of pages in his books so that the point of the fold indicated the item of his interest. We all have folded the corner of a page to indicate our position in a book and have all used fluorescent highlighters to highlight passages of interest. Try doing that to your favourite e-book!

Most importantly, books are extremely good at serendipitous access. Serendipitous access is an important yet generally unrecognised access method wherein we access the contents of a book by letting it fall open where it will. We all use it in a bookshop, where we pick a book from the shelf and, having no other knowledge of its contents, assess its interest to us by just reading from anywhere.

---

[1] Although it seems fair to conclude that the adoption of regular sizes was driven by publication methods, rather than any consideration for authors or readers.

There are types of books developed solely for serendipitous access. These 'coffee table' books have high quality images, usually little text, and are made for serendipitously browsing. They could be said to be advertorial material for travel destinations, cooking recipes, etc, and even can be said to have evolved into their own format, namely, magazines.

Books provide a standardised interface to their readers. With the exception of extreme examples, all books work the same way irrespective of the culture of their creation. Little differences do exist. Some books are read right to left, others left to right, some top to bottom, etc. Yet each of these consists of pages which are turned and are bound into a single physical form.

This standardisation of interface provides for an important piece of knowledge for authors. You know how your readers are going to interact with your book and this knowledge guides your authorship. You can also dictate how your readers will interact - if you don't provide indexes and tables of content, interaction is limited to serial or serendipity.

The long timeline of development has enabled the emergence of sophisticated and tried and tested protocols for addressing your reader. In *When Readers Become End-Users: Intercourse Without Seduction*, Dorne (1993) describes how nineteenth century authors woo readers using such phrases as 'Dear Reader' or 'Gentle Reader'. Using this language, readers were valued as the author's friend. Authors thought carefully about how to address the reader and this mode of address led to the expectation that the reader would stay with the author, if not from the first page, then at least following an argument from its unfolding to its conclusion.

5 Characteristics of Electronic 'Books'

So what about electronic books?

Well, comparatively, electronic books are readable for only a short period of time. We are all aware of not being able to read an electronic document within a short time of its creation. Word 3 for DOS is not supported by current versions of Microsoft Word and thus this format is unreadable without migration to a later format. Similarly, Wordstar documents would present difficulties for most readers. The numbers of unsupported formats increases everyday, as technology develops into newer formats and it becomes uneconomic to provide support for older formats. A problem facing book developers today is the myriad of available formats. For example, GraphicConverter (http://www.lemkesoft.de), a program for converting between image formats, supports over thirty different formats, and these only represent the common ones!

Electronic books require intermediate technology to read them. You can not simple get a stream of bits and read it without IT technology to interpret those bits and render some semblance of what the author intended. This dependency on intermediate technology limits the useful environment for electronic books. Even with the mobile technology available today, electronic books are not really enjoyable in bed! As for under water or in outer space, add a further layer of

technology to protect the electronics from adverse environments - removing you further from the book itself.

A quick survey of electronic books indicates that there is no standard interface and few cues and guides to assist the reader. Where are the front and back covers? How can we tell we're a third of the way through or have only a few pages to go? These cues are important in ensuring the reader remains engaged with their book and not 'lost in hyperspace'. The paradigms for engaging the reader are missing. We do borrow from other technologies, such as radio, TV and film but we have yet to answer the question, when does the applicability of these borrowed paradigms cease to be functional.

This lack of standards means that the author does not know how their readers will interact. We ourselves have many different mechanisms for engaging with electronic information spaces such as the Internet. We may go directly to what we want, we may browse or we may employ one of the many search engines. Essentially, we will utilise a mechanism that suits the task at hand. Since our tasks change over time, day to day or access to access, our mechanisms change accordingly. This makes it difficult to build access methods like in paper-based books. The lack of the visual clues and guides that we are used to with paper-based books lead to massive problems for users. Lost in Hyperspace is a particularly well-known problem yet few if any web sites make any attempt to assist the reader.

Electronic books excel at random access. If there is anything the IT industry can do very well, it is searching through copious amounts of content, albeit mostly textual. Admittedly, searching the web leads to 65,000 replies to any search. Yet given the number of web sites searched, 65,000 represents a small percentage. Yet it is still too large a number for any reader. The paper-based technology developed reference librarians, who are intelligent assistants to your searching problem. Searching engines are trying to address the trade off between intelligence and brute force. However, intelligent searching is still some way off. The electronic reference librarian has not been seen yet.

Whilst excelling at random access, electronic books are also good for serial and sequential access. Once you find a place to start, it is simple to follow a trail marked out by an author or publisher. Serendipity is overall lacking. Google (http://www.google.com) has a simple system for serendipity but it fails to take account of the reader's context.

Overall, it seems fair to say that for electronic books, electronic paper is common. Portable Document Format is the common exemplar of such a paradigm. With electronic paper, developers attempt to mimic paper, to produce a simulacrum of real paper, to varying degrees of success. They provide gutters and margins, headers and footers, tables of contents and indexes. Yet it is still difficult to find out where you are in these electronic books: what is the extent of the information space.

A major part of the problem is current display technology. Seventy-two dots per inch is a very low density information display, even at millions of colours. Higher density displays do exist and three hundred dots per inch is commercially available yet remains out of reach of most consumers. This is essentially a

technical and commercial viability problem. Once consumer demand is large enough, prices will fall to commercially viable levels and these types of displays will promulgate.

Readers of all types of information, whether paper-based or electronic, process the information before their eyes by splitting it into manageable chunks. Such chunking is reinforced, or maybe enforced, by our production processes and our education system. Adopting strict sizes for paper, enforces chunking to that size. Our education system trains us to efficiently use those standard sizes, even to the point of favouring some over others. Most of us can easily handle A4 or US Letter, but give us A3 or PDA/mobile phone sized screens and we begin to wilt. Simple observation of your own processing quickly indicates that we employ dynamic chunking for electronic information.

In dynamic chunking, we chose a chunk size suitable to the problem at hand and the display device in use. The problem for this strategy is that it is impossible to lay out your information to similarly adapt to such environments. To quickly capture your readers' attention, the important information should be laid out on a line commencing top left of the display, along a straight line to middle right and then along the line to bottom left. Research on head-up displays indicates that our eyes traverse this path automatically within the first few microseconds of us looking at a page. Yet if we design for A4 and display on PDA screens, this path is completely different and misses what the authors' determined what was important. In any scenario where you have very limited time to attract attention, this presents major difficulties. So, in dynamic chunking, chunk granularity is a function of the display device.

6 Technologies Involved

An interesting perspective on the viability of electronic books comes from a simple comparison of what is involved in accessing one, compared with accessing a paper book

On the paper side, you essentially need just your fingers and eyes. All books have the same behaviour. Pages are turned and eyes read the information contained. So, books can be widely read and manipulated by many people.

Contrast this with electronic books.

Electronic books can consist of multiple media, such as video, audio, images and text. In addition, most have programmed behaviour, which requires a backend that can interpret the programming instructions and behave accordingly. Video and audio require appropriate software to successfully render their content and maintain synchronicity. Then we need appropriate processing devices for hosting, delivering and displaying the content. To top it all, nowadays, we also require networking systems to store and deliver the electronic books to the reader.

The point is that for an electronic book to work, all of these technologies must be working correctly, whilst for paper books it is only the fingers and eyes.

From any perspective, it is clear that it is more complicated to keep electronic books working and this lengthy chain of dependencies indicates why so few electronic books are in regular use. It also explains why so many web sites are in

frequent use: in this case, the problem of keeping the technology working has been divested to the ISP or hosting provider, telco, etc. The reader *merely* has to keep his computer, software and network connection working[2].

7   Is the Future of the 'Book' at Risk

So is the future of the book at risk?

Well, the argument presented here is based around a simple question - is it still a 'book'?

This may sound trite, but in looking at the development of electronic books, you can see that we've appropriated the book paradigm, as a strategy for consumer acceptance and hence commercial viability. Yet in all such appropriations, the point is reached where the paradigm breaks down. When is it illogical to continue to be constrained by the appropriated paradigm? Is that time now, for electronic books and if so what will that mean for the whole publishing value chain?

In the case of electronic books, the paradigm is breaking down to such an extent that it is impossible, and probably commercial suicide, to guarantee readability of an electronic book for any length of time. The technical dependencies and volatility in the IT market would make any such guarantee an unlikely proposition.

Yet, why will consumers buy your electronic book if you can not provide any guarantee as to lifetime for the product. Would you buy a paper book that fell apart within a month? However, we expect electronic book purchasers to be cogniscent of any and all dependencies that our electronic books have with their environment, even to the point to getting them to download and install, either new software, or upgrades to their existing software.

So, to ensure commercial viability, one has to answer the question, how can we future proof our electronic books? A project, currently underway at the National Archives of Australia, is attempting to answer that very question.

8   National Archives of Australia - AtoR Digital Preservation Project

The National Archives of Australia's *Agency to Researcher Project (AtoR)*, is tackling the problem of preserving and providing access for the next one hundred years to electronic documents created in Federal Government agencies. This issue is similarly being addressed by many other archival institutions. The PADI web site, managed by the National Library of Australia (http://www.nla.gov.au/padi) provides a glimpse of international activity in this area.

---

[2] Which is daunting enough for most consumers given the rate of change of technology

In essence, the AtoR project is attempting to achieve:

• *Technology independence*, developing technologies that will enable the reading of electronic documents over 100 years; and

• *Format independence*, developing processes to convert unstable electronic formats into stable electronic formats.

A green paper on the project's approach has been published on the NAA's web site and is available from (http://www.naa.gov.au/recordkeeping/er/digital_preservation/summary.html).

As described in the green paper, the cornerstone of their approach is the use of archival data formats that are non-proprietary and specifically designed for long-term access across different computer platforms. Archival data formats are formats that digital data objects are converted into for preservation purposes.

The idea of creating their own data formats to meet the preservation needs of many record types is not as daunting as it first seems. Mark-up language technology, and specifically XML, allows quick and easy creation of non-proprietary archival formats.

Since the specification of the XML standard is freely available, the National Archives can create and maintain its own XML tools without dependence on a particular IT vendor and their proprietary knowledge. The preservation program can thus use XML as its technology base indefinitely. Even if the IT industry replaces XML with another data format technology in the future, we will still be able to create our own XML tools for as long as we wish because all the information needed to construct XML tools is publicly available.

Lessons learned from these types of activities will be of great commercial value to electronic book publishers. Adopting an 'archival' independent format should critically affect the useable life of an electronic book and make it more palatable for consumers.

The National Archives of Australia, through its AtoR project, has made a courageous decision. In adopting the approach described, they are guaranteeing that any digital record held in their archive will be readable for 100 years. To meet this goal, the AtoR project is developing a reader program, called Xanadu, that the NAA will support whilst access to their digital holdings are required. This step, taken by the NAA, is what is missing from commercial publishers. And yet, they could move along this road by adopting common standards and working together to ensure that electronic books remain readable

9   Testimony Software - An Example of Electronic Book Technology

No discussion on electronic books is complete without an example of the potential of the approach. Turtle Lane Studios Pty Ltd, an innovative Australian company, has developed an architecture called, Events on Line (http://www.turtlelane.com.au/TLS/SignificantProjectsData.html#anchor355453), that forms the basis of their Testimony Software system (http://www.attendonline.com/Testimony/testimony.html).

Testimony Software delivers synchronised, multiple, parallel media to a client.
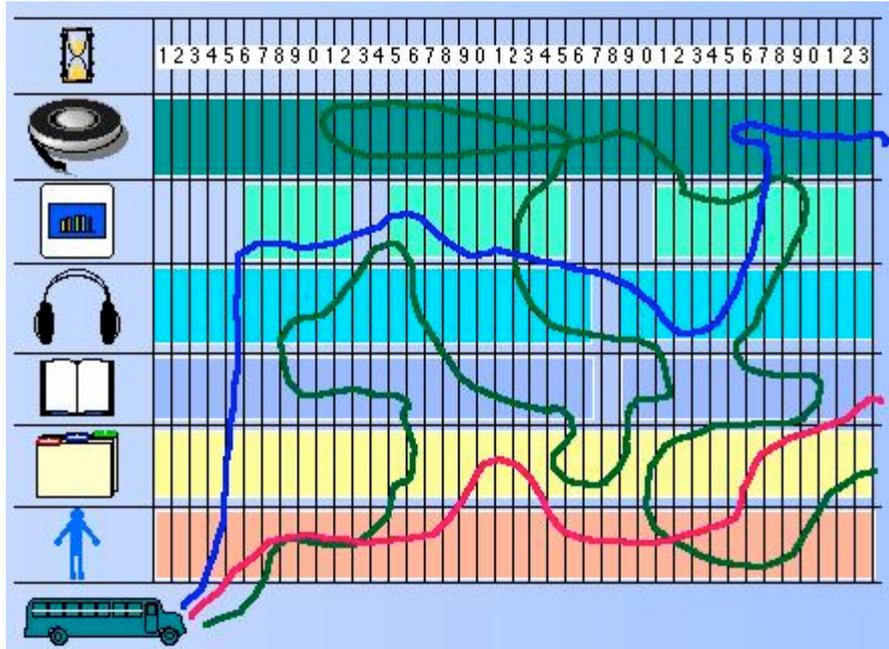
**Figure 1 *Architecture of Testimony Software***

The user drives the interaction, following either pre-compiled paths of interest or by blazing their own path through the information space. Information, in any format is made available the applicable time, so video, audio, images and text can all appear in the browser window as envisaged by the author. The software also enables the user to turn off selected media as required, so as to present the interface required by them at that time.

The software, in its third generation, has already delivered an number of products, including, a session on digital libraries presented at the Online and On Disk conference (1999), an CD ROM for the Federal Department of Information technology and the Arts, titled *The Digital Environment: New Technologies and Australian Culture*, a kiosk system for the NSW National Parks and Wildlife Service's Laperouse Museum, a multimedia kiosk for Toyota Motor Corporation of Australia, a demonstration for the National Library of Australia's oral history project, Holocaust survivor testimonies for the Sydney Jewish Museum and an art catalogue for Australian artist, Yvonne Boag.

As we can see from the example above, the user can interact with synchronised video, audio, images and text. A complete transcript of all audio is available for hearing impaired readers. The table of contents allows the user to access only points of interest to them. Text scrolls in time with the audio, images appear to illustrate the point under discussion, or can be chosen at random from a bag of available images. The user can commence any pathway at a paragraph of text, an item in the table of contents, an individual work or art, or by serendipitously toeing the water using the movie slider. Everything is indexed which provides for a myriad of pathways through the electronic space.

The interface is simple yet sophisticated to provide a real engaging experience for the user.

14

For an on-line example of the approach, please go to the Events on Line web site at http://showcase.cadre.com.au/eventsonline/. Demonstration CD's can be obtained from the author.

.
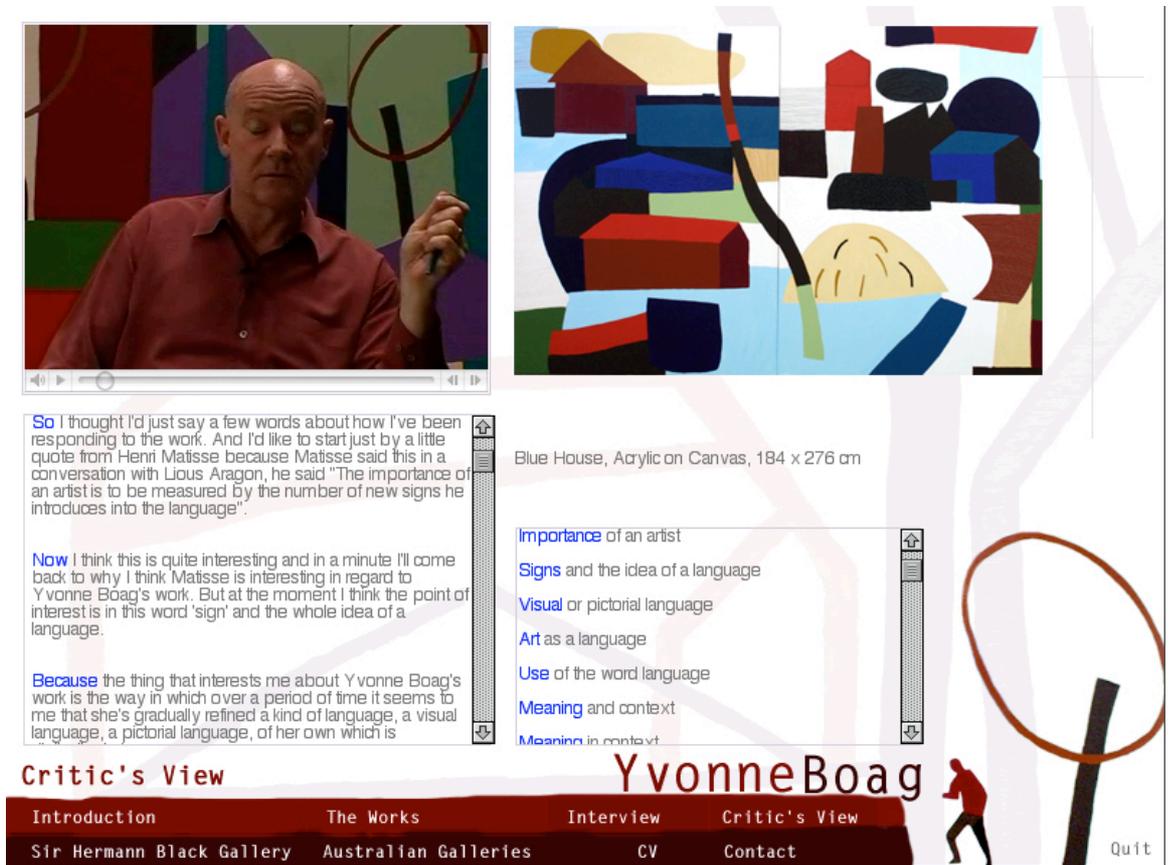
Figure 2 - *Art catalogue for Australian artist, Yvonne Boag, Critic's View section*

10 The Future of the 'Book'

So, what is the future of the book?

Well, in many ways, it is very rosy. After all, what is the likelihood of getting it all together in an IT sense. Book technology has proved very resilient and, undoubtably will remain resilient to all changes for the foreseeable future.

However, in some areas, electronic books have made remarkable gains. As described in other papers at the Book'2003 conference, electronic books and electronic book technology is in daily use in academic publishing. Students are increasingly familiar with electronic book technology and readily adapt to its idiosyncrasies and technological dependencies.

In areas such as fiction, it is difficult to say what effect electronic books will have. Whilst the Japanese read novels on their electronic book readers whilst on the train, the technology has so far made little impact in the western world - we are hooked on our paper.

From experience, it seems clear that the only way electronic books will improve is to break the dependence on technology and format. Longevity and technological complexity are two major factors impairing the wide take-up of electronic books. All in the publishing value chain must address these issues.

11 References

Dorner J, *When Readers Become End-Users: Intercourse Without Seduction*, LOGOS 4/1, 1993

Drucker J, *The Century of Artists' Books*, Granary Books (pub), 1995

Heslop H, Davis S & Wilson A, *An Approach to the Preservation of Digital Records*, URL= http://www.naa.gov.au/recordkeeping/er/digital_preservation/summary.html , December 2002

Smith KA, *Structure of the Visual Book*, Keith A Smith (pub), New York, Book No. 95, 1984

Smith KA, *Text in the Book Format*, Keith A Smith (pub), New York, Book No. 120, 1995