

An overview of the electronic publishing industry and trends A U.S. perspective.



Javier Cuadriello Rodríguez
in cooperation with
Pablo Amor Echeverri
Science, Technology and Education
Delegation of the European Commission
2300 M St.
Washington DC, 20037
Tel: 202-862-9500 Fax: 202-429-1766
www.eurunion.org
SUMMER 2000



C O N T E N T S

	Page number
1. INTRODUCTION	3
2. ELECTRONIC TEXT BASED DOCUMENTS	5
2.1. INTRODUCTION TO ELECTRONIC AND FLUID DOCUMENTS	5
2.2 TEXT BASED DOCUMENT FORMATS	7
2.2.1 TEX/LATEX	7
2.2.2. POSTSCRIPT AND ENCAPSULATED POSTSCRIPT (EPS)	8
2.2.3 HTML	9
2.2.4 ADOBE PORTABLE DOCUMENT FORMAT (PDF)	12
2.2.5 MICROSOFT OFFICE FORMATS	15
2.2.6 SGML	16
2.2.7 XML	17
2.2.8 XTM	23
2.3 COMPARISON	24
2.4 TOOLS FOR TEXT BASED DOCUMENT CREATION	25
3. DISTRIBUTION	25
3. 1 MULTIMEDIA PUBLISHING FOR WEB AND CD-ROM.	25
SHOCKWAVE AND MACROMEDIA DIRECTOR, MACROMEDIA FLASH AND FLASH PLAYER, MACROMEDIA AUTHORWARE AND AUTHORWARE PLAYER.	
3.2 WEB DEVELOPING TOOLS	30
3.3 PUBLISHING FOR MOBILE DEVICES	30
3.3.1 WAP, I-MODE.	32
3.4 ELECTRONIC BOOKS	36
3.4.1 PLATFORMS	40
3.5 ELECTRONIC PAPER	41
3.5.1 XEROX/3M	41
3.5.2 E-INK/ LUCENT TECHNOLOGIES	43
3.6 PRINT ON DEMAND	44

4. DIGITAL IMAGING	45
4.1 NEW IMAGE FORMATS	45
4.1.1 PORTABLE NETWORK GRAPHICS (PNG)	45
4.1.2 MULTIPLE IMAGE NETWORK GRAPHICS (MNG)	46
4.1.3 JPG2000	47
4.1.4 LURAWave	49
4.2 NEW VECTOR IMAGE FORMATS	49
4.2.1. SCALABLE VECTOR GRAPHICS (SVG)	49
4.3 IMAGING TOOLS	50
 5. MULTIMEDIA PUBLISHING	 51
5.1 DIGITAL MUSIC.	51
5.2 VIDEO	56

APPENDIX A. TOOLS FOR TEXT BASED DOCUMENT CREATION

A.1 SUN MICROSYSTEMS - STAR OFFICE / STARPORTAL	
A.2 MICROSOFT OFFICE, MICROSOFT WORD AND COLLABORATIVE TOOLS	
A.3 COREL WORD PERFECT OFFICE	
A.4 ADOBE FRAMEMAKER + SGML	
A.5 ARBORTEXT EPIC	
A.6 ADOBE ACROBAT / ACROBAT READER, ADOBE WEB BUY AND MERCHANT	
A.7 DTP PACKAGES - QUARK XPRESS AND AVENUE.QUARK	
A.8 DTP PACKAGES - ADOBE INDESIGN	
A.9 ADOBE INCOPY	

APPENDIX B. WEB DEVELOPING TOOLS

B.1 MACROMEDIA DREAMWEAVER/FIREWORKS/ DREAMWEAVER ULTRADEV.	
-------------------------------------------------------------	--

APPENDIX C. IMAGING TOOLS

C.1 ADOBE PHOTOSHOP	
C.2 PAINT SHOP PRO	

Appendix D - GLOSSARY

1. INTRODUCTION

Today any kind of document can be created, stored, distributed and displayed using some kind of electronic device. Paper is still seen as a useful interface medium, but computing has replaced it as the main vehicle for transporting information. Document creation has also been completely transformed and virtually all formal documents are produced using computers. However the linear nature of paper-mediated information continues to dominate systems-design thinking¹. Documents may be created in the computer, transmitted electronically, but ultimately most of them are still indented for printing. As an example word processors and desktop publishing packages (DTP) simulate a piece of paper on the screen, trying to represent as accurately as possible how the document will ultimately look when printed. However new mobile devices and cheap computers and Internet appliances are increasingly making more and more data to be accessed primarily in some form of electronic media. Publishing is no longer only concerned with producing printed pages, publishing is now also about managing content and adapting it to a variety of formats, ranging from the printed page to WAP-enabled phones, and this should only increase in the upcoming years.

Before computers and other electronic devices started invading our everyday life the definition of a document was much more clear; a piece of paper with text, graphics and pictures in it; a record or a movie could also be considered as a different kind of documents. Music, video, text and pictures could be clearly separated. New technologies have made the distinction blur. Today a given document, lets say a web page or a CD-ROM presentation can include text, music, speech and video. Furthermore new media allows new documents to lose the linearity imposed by paper, the hyperlink has added interactivity to the document; it gives the opportunity to access additional information and to move between different parts of the document.

The new electronic books present many advantages, such as search and annotation capabilities, however traditional books and newspapers do not seem to be under a serious threat, at least for the moment. Paper, with all its limitations, still provides a high resolution, cheap and portable format. An important factor is that graphics and text are still more comfortably read or viewed on paper. This is both a technological and a perception problem. From the technological side, screens lack the resolution of paper; while the resolution of the typical printer is 600 dots per inch (dpi) or more, the resolution of the average laptop or desktop screen is only 72 or 96 dpi. Also, screens are ultimately a direct source of light, which makes them uncomfortable and tiring to look at for extended periods of time. Improvements in display technologies that will increase screen resolutions, intermediate solutions between paper and a screens such as electronic ink and the proliferation of small electronic portable devices that allow to carry large amounts of information at all times may change all this. From the perception side, people has been reading on paper for a long time and we are used to it, for example most people find it easier to discover an error in a text when reading it in paper than when reading it on the screen. Even the position of the screen is different from the position for reading a book. This perception problem will probably be reduced as people gets used to access information in this new way, and as the next generations born in

¹ John Curtis, Cap Gemini, Ernst&Young

the “Internet era” grow up. For example younger people rarely print email messages, while older people or people new to email still do it often.

The World Wide Web provides a whole new opportunity for publishing and distributing information and publications, from electronic books to video and music. Most newspapers have had an online version for a long time and many are offered for free. Radio and even TV can also be transmitted thought the Internet. All this allows people to access information published in the other side of the World easily, now a reader in Frankfurt interested in American politics can read the Washington Post online, which out the need to go to a special newsagent to get it.

The distribution of music and video over the Internet also presents both a great opportunity and a great threat for publishers. The most famous representative of this new technology is Napster, right now in the middle of a highly publicized litigating against the music recording industry. But alternative, non-centralized systems such as Gnutella and other programs will make it very hard to stop illegal distribution of copyrighted material. Compressed music is the most popular multimedia file traded over the Internet due to the acceptable file sizes provided by the Mp3 and other compression algorithms. New compression algorithms for video, the possibility to override the protection in DVDs, faster internet connection in businesses, universities and thanks to cable modems and DSL also at home, may make video the next popular thing to trade over the internet.

It is hard to predict how technologies will develop, which formats and which media will prevail and which will become obsolete and disappear before reaching mainstream popularity. As recent history has shown, superior technologies can easily succumb against superior marketing strategies and promising technologies such as Virtual Reality can lose all their momentum. New opportunities and threats appear and spread rapidly and almost unexpectedly, such as the MP3 music format, the solid-state music players, Portable Digital Assistants (PDAs) and mobile phones.

Remembering the state of technology in the late 1980s and early 1990s gives us a perspective of how fast the way in which information is published and distributed has changed. Before graphical user interfaces, when modems had speed that make the transmission of pictures impractical and mobile phones were a rare luxury item; the World Wide Web did not yet exist and email was still relegated to the scientific community. A mere ten years have brought a radical transformation; email, the Internet and mobile phones are now an indispensable part in many people’s lives. It took 10 years to move for a small number of people with very slow modems reading text only pages to millions of people downloading music with Napster like utilities. For the moment the rapid growth rate in the computer technology does not seem to be slowing down and although the rate of increase of Internet connections and computer speed decreases there will still be big advances. Whatever technology comes after DSL and Cable or improvements in these technologies and new advances in wireless technologies will bring high quality streaming video to all kind of fixed and mobile devices.

This report describes some of the present and upcoming technology trends for document creation, transmission, storage, display and printing, file format and standards. It also describes some of the most popular, promising and interesting document creation tools.

2. ELECTRONIC TEXT BASED DOCUMENTS

2. 1 INTRODUCTION TO ELECTRONIC AND FLUID DOCUMENTS

Electronic texts present the opportunity to search, annotate and link the text to other useful information. Exploiting the whole power of electronic texts could radically change the way text and documents are created. Traditional ways of adding additional information in a text, such as parenthesis and footnotes, could become obsolete. Right now web pages and some other kinds of documents already give the opportunity to obtain information about parts of the text. For example, while reading an online article, when a given company is mentioned it provides a link to access information about the company. This could be taken even further. Researchers at Xerox PARC are studying what they call “Fluid Documents”. Fluid documents alter their layout, typography, and other graphical characteristics in order to present supporting material in the context of the primary material that it annotates. The document surface acts as a canvas in which traditionally static elements may move and change smoothly to make room for additional information. Documents and other information spaces often include multiple layers of information: there is some primary information that is presented directly, and then supporting information that may optionally be included. Examples of supporting information are non-mainstream material that is presented in a footnote; attribution of prior work via a reference; derivations of a formula; definitions of technical terms; illustrations and diagrams; or simply more detail about a particular point that the entire audience may not be interested in.

Researchers have been investigating user interface techniques for *fluid documents* for providing smooth, contextual access to additional information. The approach is to alter the graphics or typography of a document in order to show the supporting information in the context of the primary information it annotates. This alteration is animated, so that the reader’s attention is guided calmly to the revealed supporting material, and later guided back to the primary material. Figure 1 shows an excerpt from Shakespeare’s *Romeo and Juliet*, with the existence of annotations indicated by underlines.

But soft! What light through yonder window breaks?
It is the East, and Juliet is the sun!
Arise, fair sun, and kill the envious moon,
Who is already sick and pale with grief
That thou her maid art far more fair than she.
the moon is here thought of as Diana
Be not her maid, since she is envious.
Her vestal livery is but sick and green,
And none but fools do wear it. Cast it off.

Figure 1

In this instance, moving the cursor over an underlined phrase causes an annotation to grow from a tiny size to its full, readable size. At the same time, the primary text moves apart to make room for the annotation. Placing the annotation right at the annotated material not only allows the eye to quickly pick up the expanding annotation, but it also allows easy comparison to the primary material. When the reader is done with the annotation, he or she may simply move the mouse away and the annotation shrinks and the primary text closes back together.²

² Bay-Wei Chang, Jock D. Mackinlay, Polle T. Zellweger - XEROX PARC- Fluidly Revealing Information in Fluid Documents - Published in AAAI Smart Graphics Spring Symposium 2000.

Researchers in Xerox PARC have developed several fluid techniques: *fluid interline*, *fluid overlay*, and *fluid margin callout*.

Figure 1 uses the *fluid interline* technique for allocating space for annotations in textual documents, primarily modifying the position of the primary text to make space. This technique illustrates the central principle of fluid documents: when the user expresses interest in supporting information, the primary material must make (or find) space for that information, and the supporting information must present itself in a salient fashion.

The *fluid overlay* technique alters the primary text's color so that supporting material can be placed on top of it (see Figure 2). The supporting material also must display itself in a contrasting color. As in all the fluid techniques, the transitions are animated; so the text fades to gray while the gloss is growing to full size. This technique keeps the layout of the primary text unchanged.

The *fluid margin callout* technique alters the primary text even less than fluid overlay: the only intrusion is a line that grows from the anchor text to the margin. In this case, the gloss is placed in existing white space (see Figure 3). This example also shows that fluid links can include images.³

But soft! What light through yonder window breaks?
It is the East, and Juliet is the sun!
Arise, fair sun, and kill the envious moon,
Who is already sick and pale with grief
That thou her maid art far more fair than she.
Be not her maid, since she is envious.
the moon is here thought of as Diana
Her vestal livery is but sick and green,
And none but fools do wear it. Cast it off.

Figure 2

But soft! What light through yonder window breaks?
It is the East, and Juliet is the sun!
Arise, fair sun, and kill the envious moon,
Who is already sick and pale with grief
That thou her maid art far more fair than she.
Be not her maid, since she is envious.
Her vestal livery is but sick and green,
And none but fools do wear it. Cast it off.



Figure 3

³ Bay-Wei Chang, Jock D. Mackinlay, Polle T. Zellweger - XEROX PARC- Fluidly Revealing Information in Fluid Documents - Published in AAAI Smart Graphics Spring Symposium 2000.

2.2 TEXT BASED DOCUMENT FORMATS

2.2.1 TEX/LATEX

Tex and later Latex are languages to create high quality text based documents. Although they are old formats and they do not have a promising future, they are mentioned here for two reasons. First because they are still used by the scientific community and second because they present some interesting features, such as the separation between format and content, that have been recovered by some of the latest formats such as XML.

TEX

\TeX is a computer program created by Donald E. Knuth. It is aimed at typesetting text and mathematical formulae. Knuth started writing the \TeX typesetting engine in 1977 to explore the potential of the digital printing equipment that was beginning to infiltrate the publishing industry at that time. \TeX as it is used today was released in 1982, with some slight enhancements added in 1989 to better support 8-bit characters and multiple languages. \TeX is renowned for being stable, for running on many different kinds of computers, and for being virtually bug free.

LATEX

\LaTeX is a macro package which enables authors to typeset and print their work at the highest typographical quality, using a predefined, professional layout. \LaTeX was originally written by Leslie Lamport . It uses the \TeX formatter as its typesetting engine.

In 1994 the \LaTeX package was updated to include some long-requested improvements, and to reunify all the patched versions which had cropped up since the release of \LaTeX 2.09 some years earlier. To distinguish the new version from the old, it is called \LaTeX_2 .

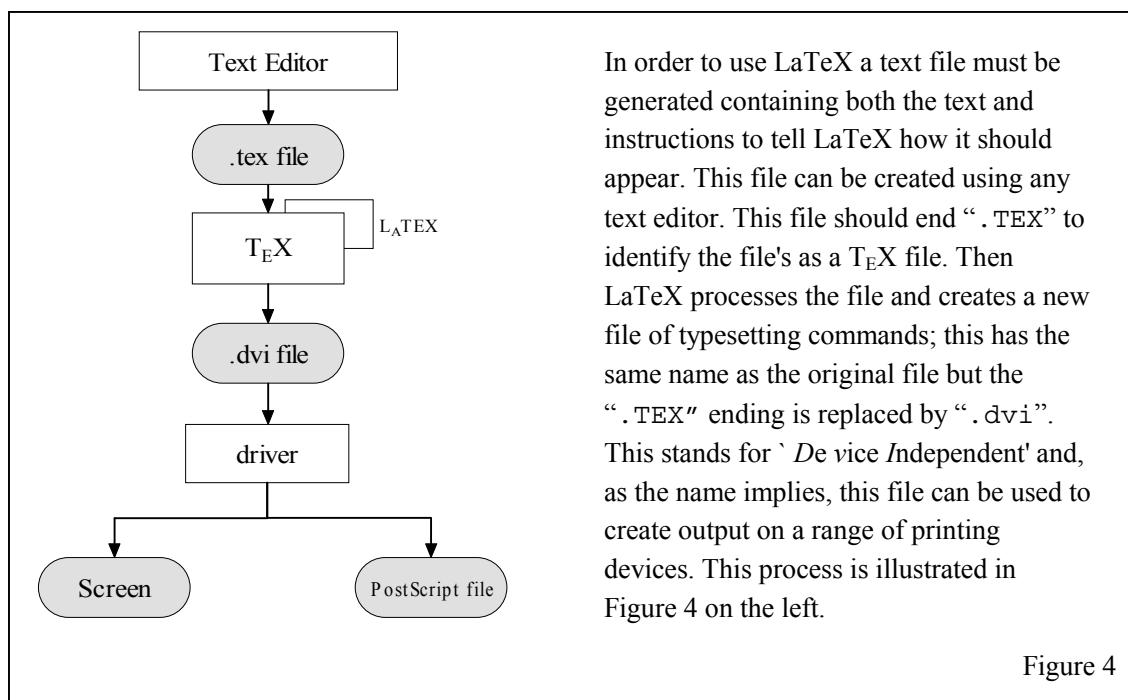


Figure 4

In a L_ATEX environment, L_ATEX takes the role of the book designer and uses T_EX as its typesetter. The author has to provide additional information, which describes the logical structure of his work. This information is written into the text as L_ATEX commands. Therefore there is a separation between the contents and the format, somehow in a similar was as the XML format (see section 2.2.7).

Generation a L_ATEX document is different from the WYSIWYG ⁴approach which most modern word processors such as MS Word or Corel WordPerfect take. With these applications, authors specify the document layout interactively while typing text into the computer. All along the way, they can see on the screen how the final work will look when it is printed. However, when using L_ATEX it is not possible to see the final output while typing the text. But the final output can be previewed on the screen after processing the file with L_ATEX. Then corrections can be made before actually sending the document to the printer.

The L_ATEX format has many desirable characteristics such as its separation between format and contents and its portability, however the lack of a WYSIWYG environment for writing the files has kept L_ATEX away from mainstream use. L_ATEX is relegated to the scientific community and many papers are still submitted to scientific publications in this format. L_ATEX defenders believe that the way L_ATEX forces to structure the contents produces better quality documents. However the knowledge and amount of learning required to produce a Latex document is far greater than in any modern word processor. As for the future of L_ATEX it will probably stay relegated mainly to UNIX systems and the scientific community still for so long, but it does not seem to be any chance of ever being widely accepted by the public.

2.2.2. POSTSCRIPT AND ENCAPSULATED POSTSCRIPT (EPS)

PostScript was introduced by Adobe Systems in 1985. PostScript is a “programming language” but unlike other languages, PostScript is designed to describe accurately what a page looks like, therefore it is known as a *page description language*, optimized for printing graphics and text (whether on paper, film, or on the screen).

In order to be printed a PostScript file needs to be processed by a RIP - a Raster Image Processor. A RIP takes in PostScript code and renders it into dots on a page. This processor is a combination of software and hardware, which typically lives in a printer. Therefore a PostScript printer is a device that reads and interprets PostScript files.

The main purpose of PostScript was to provide a convenient language to describe images in a device independent manner. This device independence means that the image is described without reference to any specific device features (*e.g.* printer resolution) so that the same description could be used on any PostScript printer without modification. In practice, some PostScript files do make assumptions about

⁴ What You See Is What You Get

the target device (such as its resolution or the number of paper trays it has), but this is bad practice and limits portability.

PostScript files can be viewed in every current computer system. Free PostScript viewers, such as GhostScript/GhostView (which run on Windows, Macintosh (ported by Glyptic Technologies), Linux and most versions of UNIX) can be found on the Internet.

For many purposes, PostScript is an old format to save, store and transmit files over the Internet. PDF (See Section 2.2.2) is a replacement for storing and transmitting PostScript files and EPS files in many, but not all, cases. However PDF is *not* a replacement for the PostScript language or the PostScript processors that live inside of printers, imagesetters, and platesetters. Adobe Illustrator files are also essentially PostScript files and they are best kept in that format for later editing, not in PDF.

Encapsulated PostScript

An EPS file is a PostScript program, saved as a single file that includes a low-resolution preview "encapsulated" inside of it, allowing some programs to display a preview on the screen. Programs such as Adobe InDesign (see Appendix A.8) do not need this preview because they have a RIP built in, which allows it to open PostScript files natively.

2.2.3 HTML

HTML is the standard format to display documents (web pages) over the Internet. HTML files can be read by browsers such as Netscape, Microsoft Internet Explorer and many others, but also programs like Microsoft Office can open (and create) HTML files. As an HTML file is just a text file and as there are web browsers for every operating system HTML becomes a truly platform-independent format.

HTML is a content-based or structural markup language. This means that codes (tags) are used to indicate the various parts of the document, such as headings, paragraphs, lists, etc. However it does not go as far as SGML or XML and tags cannot be defined. This markup also includes codes for forming *hypertext links*, which allows connecting to other document or pieces of information.

HTML does not maintain a completely fixed page layout and depending on the size of the windows or the resolution of the monitor the file will be displayed in a slightly different way, this is useful for web pages but it implies that documents cannot be printed with fidelity. This makes HTML unsuitable as an intermediate format for printed publications.

Although HTML is a text file and therefore can be created with any text editor, there is an amazingly large number of programs (such as Macromedia DreamWaver (see Appendix B), Adobe GoLive etc.) that allow the creation of HTML documents in a graphic, WYSIWYG⁵, environment. As these tools improved and the complexity of HTML increased they have become the most common method for creating HTML documents. Even Microsoft Word and Corel Word perfect can save as HTML pages

⁵ What You See Is What You Get

producing acceptable results for simple documents. Thanks to all these tools, very little knowledge is required anymore to produce a HTML document.

Although basic HTML can be displayed in any browser, more complex code becomes somehow browser specific. The two most popular browsers, i.e. Internet Explorer and Netscape, implement some of the new features in different ways. The main differences are in the way CCS are implemented (to be disused in Section 2.2.7 , XML) and some proprietary extensions.

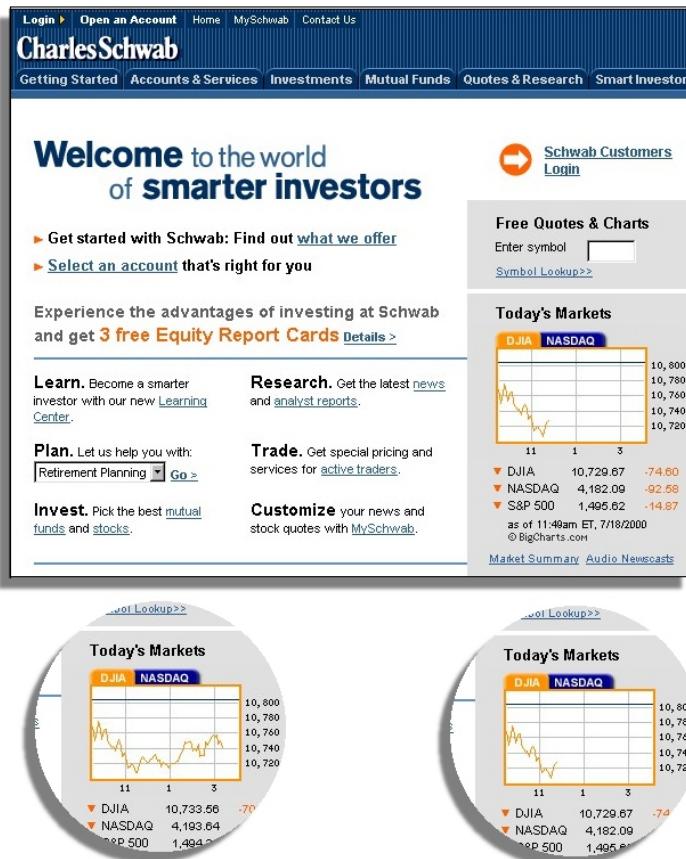


Figure 5

them without altering the rest of the document. As an example of this we have banner adds that change every time the page is loaded or graphs reflecting the changes in the stock market that are generated automatically outside the document (see Figure 5).

HTML 4.0 is an SGML application (Standard Generalized Markup Language). The HTML standard is maintained and developed by the World Wide Web Consortium (W3C). The W3C provides a on-line validator for HTML pages (<http://validator.w3.org/>). This system can tell if a given page is 100% compliant with the W3C standards. However, popular websites such as Cnn.com, Yahoo, the Financial Times, Altavista and many others, do not pass this test.

HTML is a very flexible format; images, audio, video, Java applets and scripts, Flash animations (see section 3.1) and a large number of browser plug-ins can be used inside a HTML document, although all these are always on external files. This can be both an advantage and a disadvantage. For one side the document is not self-contained, therefore to send all the linked elements have to be sent separately and placed in the correct directories so that the document can find them. Therefore, sending a graphics intensive HTML document by email is not practical. On the other side, for a page being displayed on the Internet, the fact that images, sounds and other elements are outside (even in another server) the HTML file gives the flexibility to change

The current version of HTML is 4.01. In addition to the text, multimedia, and hyperlink features of the previous versions of HTML, HTML 4.0 supports more multimedia options, scripting languages, style sheets, better printing facilities, and documents that are more accessible to users with disabilities.

Further developments in HTML by the W3C consortium focus on the better integration of HTML with the upcoming XML (See section 2.2.8, XHTML).

The “BROWSER WARS”

According with StatMarket.com⁶ 86 percent of Internet users worldwide utilize the Microsoft Internet Explorer Web browser. (June 18 2000). Reports indicate that usage of IE has steadily risen, going from 64.6 percent on Feb. 8, 1999.⁷ Therefore, Microsoft Internet Explorer seems to have won the so-called “Browser Wars” which began when Microsoft released Internet Explorer to compete with Netscape, then dominating most of the browser market.

Netscape’s new browser (the preview release of version 6, at the time of writing this report) incorporates several improvements and the new faster rendering engine for web pages. Netscape runs in every major operating system, including Linux and FreeBSD and all the most popular versions of UNIX. Netscape opened the code of its browser in 1998. The Mozilla project coordinates the efforts of developers all over the world that work on the parts of the Netscape code. Netscape incorporates an email client and newsgroup reader with an open email storage format compatible with older Unix format and Qualcomm’s Eudora. Although it does no longer dominate the browser market, many institutions and users remain loyal to Netscape.

IE integrates better than Netscape with MS Windows and is also available for Macintosh and UNIX (version 5.0 only at the time of writing this report). It also incorporates superior options for saving web pages to a local computer.



⁶ Part of WebSideStory, Inc

⁷ WebSideStory arrived at its 86.08 percent figure for IE by recording information about browser types collected only by sites using its HitBox technology. In August 1999, the company said IE's share was at 75 percent, based on tens of millions of visits to HitBox sites.

2.2.4 ADOBE PORTABLE DOCUMENT FORMAT (PDF)

In 1993, Adobe introduced the Portable Document Format (PDF) as an extension to PostScript. Essentially; PDF is a digital document specification, to ease the burden of cross platform distribution of visually rich documents. The PDF format preserves all of the fonts, formatting, colors, and graphics of any source document, regardless of the application and platform used to create it.

PDF files can be published and distributed anywhere: in print, attached to e-mail, on corporate servers, posted on Web sites, or on CD-ROM. Using Acrobat software, bookmarks, cross-document links, Web links, live forms, security options, sound, and video (from external files) can be added to PDF files for enhanced online viewing. PDF files are compact and can be shared, viewed, navigated, and printed exactly as intended by anyone with a free Adobe Acrobat Reader.

The free Acrobat Reader can be downloaded from Adobe's Web site and can be freely distributed by anyone. The reader is available for every current operating system. According with Adobe, more than 110 million copies have been downloaded or preloaded onto computers. Acrobat Reader can also run as a browser plug-in for viewing PDF files inside the browser. There are also alternatives to the Adobe Reader to display PDF files, such as Glassbook reader, optimized for displaying electronic books in PDF format. Glassbook has recently been acquired by Adobe Systems.

PDF is built largely on the PostScript language, but the PDF format presents several advantages over the plain PostScript format. PostScript was designed to describe a page but PDF can also contain information not only related to how a page looks, but also can describe how it behaves and what kind of information is contained in the file. A PDF file can contain fonts, images, printing instructions, keywords for searching and indexing, job tickets, interactive hyperlinks, movies etc. A PDF file is actually a PostScript file which has already been interpreted by a RIP (Raster Image Processor , See section 2.2.2 Postscript) and made into clearly defined objects. These objects are viewable on screen not in code, but in visual objects that everyone can see. Because these files are already interpreted by the RIP, they can be more reliable than an EPS or a plain PostScript file when printed. Additionally, because EPS files and PostScript files can be easily converted to PDF and viewed on screen, print service providers can benefit from seeing the file *after* interpretation but before they send it to their printing devices. This may allow them to see errors in a file before wasting paper, film, or plates.

In order for a PDF file to be printed, however, the printer still needs to render the PDF objects to the page, and a PostScript printer is still the most reliable way to do this. Some PostScript printers understand not only the PostScript language, but also PDF files natively. And some printers, using a technology called Extreme, actually convert *all* jobs into a PDF file prior to printing. Agfa, Creo, Heidelberg, and Scitex have all announced print workflows based on Extreme.

Adobe Acrobat- the tool for enriching and modifying a PDF file - can print PDF to non-PostScript printers by interpreting the PDF file into that printer's language. But these languages are not as reliable or accurate as true Adobe PostScript, so professionals use a PostScript processor for best results.

PDF can be used as a replacement file format for EPS, and that PDF can be used as a delivery format for sending complete publications to press. It is also suitable for soft-proofing, distribution on the Internet,

and file archiving, as it is completely self-contained. But to print PDF, it is best to print to a device with true Adobe PostScript capabilities to get the highest quality output.

The diagrams below illustrate how the PDF format can be used to increase the efficiency of the document work flow for both business or personal use.

Figure 6 - POSTSCRIPT WORKFLOW. -

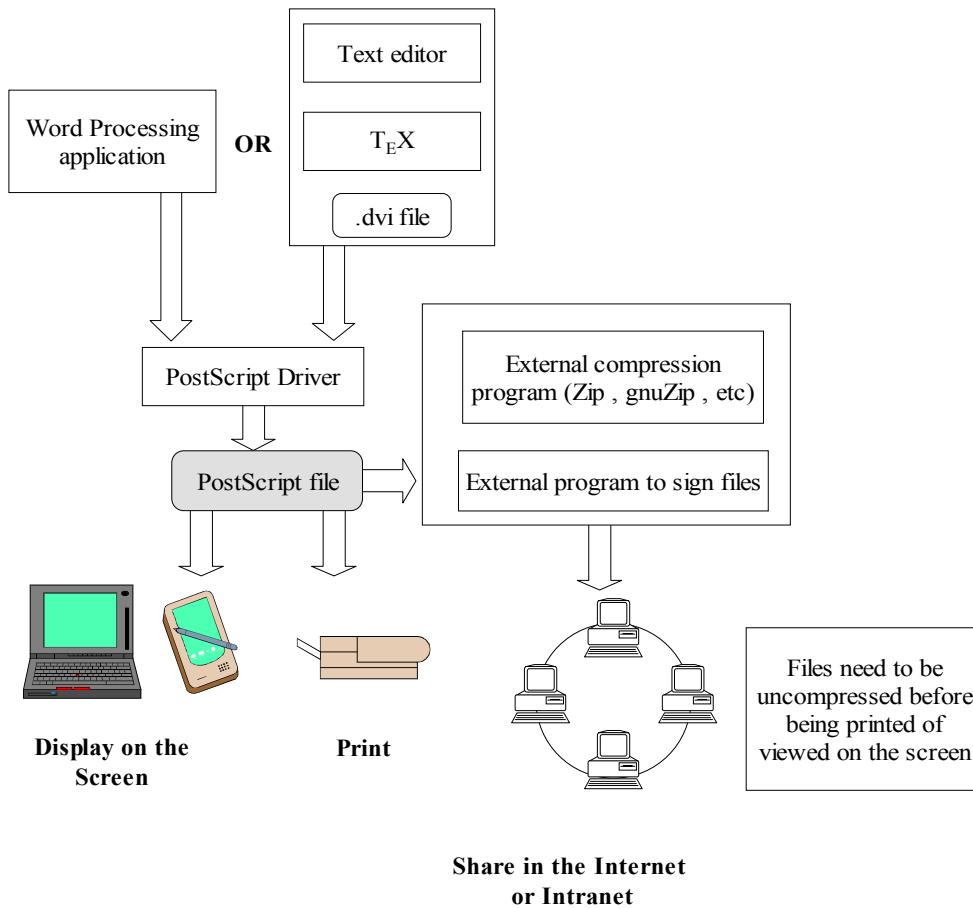
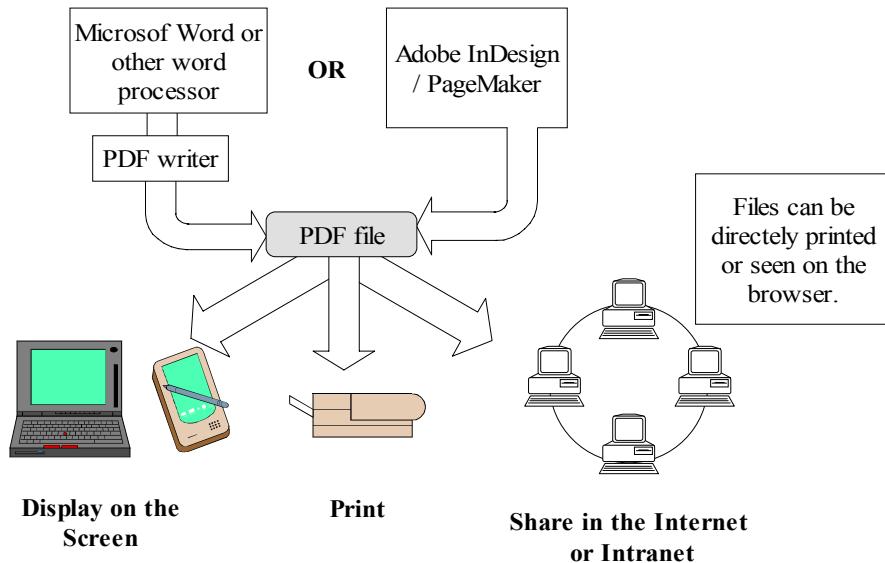


Figure 7 - PORTABLE DOCUMENT FORMAT (PDF) WORKFLOW



Until recently PDF files were created exclusively with the Adobe's software, therefore, this electronic format was associated with the software that created it and the software to read it, both property of Adobe Systems. However Adobe recently opened the PDF standard releasing a big part of the code and the API. New releases of Denebas Canvas, Corel Draw, Metacreations Freehand, etc support the creation of PDF files and a number of alternatives to convert document to PDF have appeared in the market. Furthermore Adobe has a service in their website that allows the on-line conversion of documents to PDF format for a subscription fee. However the most complete solution to create and edit PDF files with all the latest features continues to be Adobe Acrobat.

Compression

To reduce size, the PDF file format supports a number of industry-standard compression filters, including JPEG for images with selectable quality vs. size options.

- JPEG compression for color and gray scale images
- CCITT Group 3 , CCITT Group 4 , LZW (Lemple-Ziv-Welch) , and Run Length compression of monochrome images.
- LZW and Flate compression of text, graphics and indexed image data.⁸

As a result PDF files are usually smaller than their source, which makes them good for transferring through the Internet without the need of an external compression program.

⁸ Source: Adobe Systems Inc.

Portable Devices and Electronic Books. (See section 3.4)

PDF is a serious candidate as a standard file format for electronic books. According to Adobe a new version of the PDF specifications will allow text and images to reflow to accommodate different devices so that PDF files can be displayed in the smaller screens of PDAs and mobile phones. The possibilities of using PDF in these devices are discussed in more detail in section 3.4.

Security.

PDF files can include Digital signatures, (see later Adobe Acrobat in the tools Appendix A) and can be protected against printing, modifying and selecting and copying text.

Summarizing, the PDF format offers the following advantages over the plain Postscript format.

- Can include hyperlinks, videos, etc.
- More reliable than PostScript or Encapsulated Post Script (EPS) files.
- Smaller size without the need of external compression application. Convenient to send over low bandwidth networks.
- Can be digitally signed without the need of external program.
- Can be seen directly in the browser via the Acrobat Reader plug -in.
- Can create interactive forms.

2.2.5 MICROSOFT OFFICE FORMATS

The Microsoft Word document format (.doc) is widely used as a common format for document exchange. Excel (.xls) and PowerPoint (.ppt) file formats are also highly popular to distribute spreadsheet data and presentations respectively. They are de facto a common standard for document creation and exchange for many organizations and personal users. As most text-based documents are written in Microsoft Word it makes sense that its native file format is used extensively. Once again these formats are tightly linked to the software that creates them and the company that manufactures that software, in this case Microsoft Corp. Only recently some other software packages are able to read and to export to Microsoft Word files although import filters have usually been available in other Office suites such as Corel Word Perfect and Lotus SmartOffice.

Word files have the advantage over PDF files of being able to be modified by anyone that has Microsoft Office. Due to the dominant market position of Word this is a great advantage for documents in which several people have to collaborate, MS Word also includes several tools to make this collaboration easier (see Appendix A.2).

Word files incorporate limited security and protection features; a file has been protected against writing and/or reading with a password. However the security features in Word are very weak and overriding the password feature is relatively easy.

Word files can also be read by people that does not have a copy of Word by using the free Word Viewer, available for download at Microsoft's website. A major argument against the Word format is that it can

only be used in Macintosh and Windows environments. However considering the combined market share of these two operating systems for the desktop market (Microsoft dominates the OS market, with 93.63%⁹) it doesn't seem like a determinant reason. Furthermore other products such a Sun Microsystems StarOffice Writer which, apart from Windows, also runs under Linux, Unix, and Macintosh (Macintosh support is expected in late 2000) operating systems, can open and save as Word files supporting limited features (furthermore StarOffice Cal can import and export Excel Files and StarOffice Impress can import and export PowerPoint Presentations).

Word (and Excel) files can even be opened in Palm OS devices with programs such as "*Documents to Go*", by Dataviz and in PSION personal organizers running the EPOC32 operating system (See section XX) with the included text processor (also called Word). However due to the fact that the original source code from Microsoft is not available to any third party developers, the 100% of features of Office files are not fully supported by any non-Microsoft application. Complex Word files sometimes present small problems in formatting when opened in other computers or when imported to different versions of Word. Files also tend to be larger than PDF or HTML files.

Despite of any legal problems Microsoft may have and even if the anti-trust appeal ends up splitting the company in two, Microsoft Office will probably still remain the market leader for the foreseeable future, either in its present form or in any free or subscription based online version (such as ThinkFree.com). Therefore its format and any future developments of it, including any XML implementation, will continue to be a popular choice for writing, editing, collaborating and transmitting documents, spreadsheets and documents.

2.2.6 SGML

SGML, or Structured Generalized Markup Language, was conceived in 1986, as a standard for the advanced structuring of documents. It provides rigid, complex rules for document structure and format. The tools for creating and viewing SGML have been limited to a small, but dedicated user community. This community includes high-end publishers and government entities that rely on very structured data that needs to be shared in a tight community. Without going into too much detail about SGML, suffice it to say that it is the superset, foundation language for both HTML and XML. The reason for these "derivative" languages is that SGML is a metalanguage, or a specification that allows other languages to be created from it.

The main characteristics of SGML are:

- Infinite possibilities for expressing information i.e. infinite tag set.
- Allows document re-use
- Requires validation for completeness and correctness.

SGML has found its main customer base in organizations handling enormous quantities of documents - the U.S. Government Printing Office, IBM, the U.S. Department of Defense and Internal Revenue Service, and many publishers.

⁹ As of June 18 2000 according with StatMarker.com , a division of WebSideStory Inc.

2.2.7 XML

XML stands for Extensible Markup Language and is intended for applications that require functionality beyond the current Hypertext Markup Language (HTML). XML is being designed by a Working Group of the World Wide Web Consortium (W3C). This Working Group consists of about 14 companies and organizations with an interest in either providing or utilizing XML tools. This group includes Adobe, ArborText, DataChannel, Fuji Xerox, Hewlett-Packard, Inso, Isogen, Microsoft, Netscape, SoftQuad, Sun Microsystems and the University of Chicago.

The XML standard itself is completely open, the W3C members have early access to standards, but once the standard is complete the results are public and freely available on the web.

Due to the big support by industry and its desirable characteristics it seems inevitable that XML will become the primary means to deliver and exchange information in the near future.

The purpose of XML is to specify a highly functional subset of SGML (Standard Generalized Markup Language) that works well for delivering SGML information over the Web. XML has almost all of the capabilities of SGML that primarily affect document creation, not document delivery. The SGML capabilities that were dropped from XML are those irrelevant to the delivery of structured information over the Web. However, some of those capabilities are important to the creation of structured data. The reason for this is that XML was not originally designed to replace SGML. SGML remains an appropriate technology for creating and storing information. It is possible that subsequent revisions of XML will include some of the omitted SGML capabilities.

HTML and SGML have several limitations that XML intends to solve. These limitations are summarized below.

Main limitations of HTML

- *Extensibility.* - The HTML format does not allow an extensive mark-up to provide further information about the document, it does not allow users to specify their own tags or attributes in order to parameterize or otherwise semantically qualify their data. HTML has only a fixed set of tags, which serve mainly to specify formatting of the document in a limited way. This limits the possibilities for publishing the same information in different media without the need of human intervention, as the different parts of the document cannot be recognized by the computer and therefore cannot be properly re-formatted according to the desired publication media (e.g. browser, PDA, etc). It is also very difficult for different applications to exchange data from a HTML document, as once again, the computer does not understand what the data means. This lack of description of the content also results in limited searching possibilities.
- *Structure.* - HTML does not support the specification of deep structures needed to represent database schemas or object-oriented hierarchies.
- *Presentation.* - HTML has limited control over the presentation and the final page layout.

- *Validation.* - HTML does not support the kind of language specification that allows consuming applications to check data for structural validity on importation.
- *Maintenance.* - Maintaining data in HTML is relatively complicated and requires considerable human intervention.

Main limitations of SGML

- *No browser support.* – None of the mainstream browser, mainly Microsoft Internet Explorer and Netscape, provide support for SGML, however, from version four, they already include partial XML support. SGML offers many options and supporting all those options completely requires very complicated software. Even specialized SGML providers such as ArborText (See Appendix A..5) do not support a 100% of the options that SGML allows.¹⁰
- *No standard support for styles.* – Although some attempts have been made there no stylesheet format standard.

One of the most important limitations from HTML that is solved by XML is the lack of extensibility. As mentioned before, XML stands for “*Extensible Markup Language*”. First it is a markup language because it includes tags that provide information about the contents of the text. For example, in an XML document, title, data, summary, customer number etc. will be labeled, i.e. tagged as in the example below.

```
<Title> E-publishing </Title>
<date> 12/12/2000 </date>
```

XML is also *extensive* because an infinite number of new tags can be created. A Document Type Definition (DTD) is a set of syntax rules for tags. It tells what tags can be used in a document, what order they should appear in, which tags can appear inside other ones, which tags have attributes, and so on. Originally developed for use with SGML, a DTD can be part of an XML document, but it is usually a separate document or series of documents.

Because XML is not a language itself, but rather a system for defining languages, it does not have a universal DTD the way HTML does. Instead, each industry that wants to use XML for data exchange can define its own DTDs. If an organization uses XML to tag documents for internal use only, it can create its own private DTD. The Wall Street Journal Interactive Edition, for example, has a DTD specifying each edition, with information about pages, articles, summaries, bylines, etc.

One the main advantages of XML derives form its description of the information is that it makes easier the task of transmitting information between different application and database systems. For example, different computer systems can exchange and process XML-formatted purchase orders and invoices with little human intervention. This has important implications in for electronic commerce and B2B transactions and is one of the main reasons of the interest and industry support of XML.

Information formatted in XML can be also be easily identified by “smart agents” searching for information or can also be identified by human operators looking directly at the XML source.

¹⁰ Source: ArborText

Internet search engines will greatly benefit from this and search hits could be reduced if every webpage was written in properly formatted XML. However it will take a relatively long time for pages to take advantage of XML and it is not clear if private users will like and need the extra complexity of formation document in XML. Even if new intuitive tools are developed, up to a certain extent, the author will always have to specify what the different parts of the text are. Which takes extra time and effort. Intranets from companies and large organizations are more likely to benefit from XML searching capabilities in the short term.

XML is also extensible in the sense that the standard itself is being extended with several additional standards that add styles, linking, and referencing ability to the core XML set of capabilities. As a core standard, XML provides a solid foundation around which other standards may grow. W3C working groups are developing additional supporting standards for XML. *XLink* provides linking facilities that are far more sophisticated than those in HTML. *XPointers*, derived from the Text Encoding Initiative's (TEI) extended pointers, provide a way to consistently reference portions of documents.

Another important feature of XML is that it separates the contents of the document from the way it is displayed. Therefore the same content can be presented in different ways, without the need to re-write the information itself. The XSL (Extensible Style-sheet Language) stylesheet determines the way XML documents will be ultimately displayed. Different stylesheets can transform and format the same XML content to be displayed in different devices, from a browser in a desktop computer to the small screen of a mobile phone or a PDA. (See Figure 8)

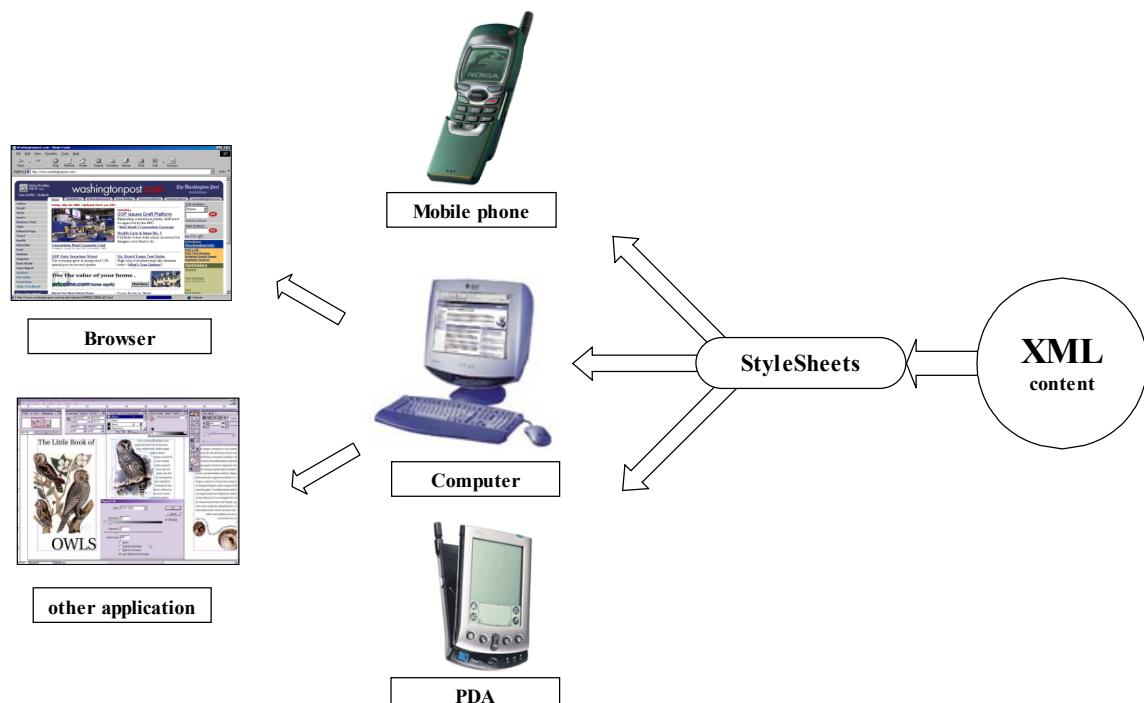


Figure 8

Separating the contents from the display format was already partially possible with HTML by using Cascading Style Sheets (CSS). Style sheets describe how documents are presented. The W3C has

produced two recommendations (CSS1 and CSS2), which are widely, although not consistently, implemented in browsers.

The fact that W3C has started developing XSL in addition to CSS has caused some confusion. The unique features are that CSS can be used to style HTML documents. XSL, on the other hand, is able to transform documents. For example, XSL can be used to transform XML data into HTML/CSS documents on the Web server. This way, the two languages complement each other and can be used together.¹¹

The W3C recommends using CSS when possible (See Table 1). The reason is that CSS is much easier to use, easier to learn, thus easier to maintain and cheaper. There are WYSIWYG editors for CSS and in general there are more tools for CSS than for XSL. But CSS's simplicity means it has its limitations. Some things cannot be done with CSS, or with CSS alone. Then XSL is needed, or at least its transformation part. According with the W3C, XSL should be used for everything that needs

transformations. For example, a list that has to be displayed in lexicographical order, or if words have to be replaced by other words, or if empty elements have to be replaced by text. CSS can do some text generation, but only for generating small things, such as numbers of section headers.

XML documents can be rendered in three different ways:

- If the document does not have to be transformed: use CSS.
- If the document has to be transformed: Use XSL-T, the transformation language of XSL, either:
 - Generate the style properties together with the rearranged text, using a sub-language of XSL called XSL-FO (XSL Formatting Objects)
or
 - Generate a new XML or HTML document and provide a CSS style sheet for that new document.

As mentioned before, XML is intended to have a prominent role as a data interchange format allowing different applications and devices to share information regardless of the maker of the device of the software. Therefore for XML to exploit all its possibilities it is particularly important that it is kept as a standard.

It has been proven impossible to keep HTML completely standard despite the efforts of the World Wide Web Consortium. The two main browser manufacturers failed to implement new features in a consistent way and added new features not compatible with each other. However due to the flexibility of HTML and the tolerance to errors by web browsers it has only been a minor problem. XML will impose more strict rules than HTML, which will not allow this kind of errors. Furthermore for XML there will be many more than two major players. Some of them may consider including new features, in particular Microsoft and its .NET strategy may try to incorporate features designed specifically for Microsoft

Table 1

	CSS	XLS
Can be used with HTML	yes	no
Can be used with XML	yes	yes
Transformation Language	no	yes
Syntax	CSS	XML

Source: W3C

¹¹ Source: The World Wide Web Consortium (W3C)

software or web services, in the same way they did with the Java programming language for example. Other companies such as Quark or Adobe will be interested in adding specific features for DTP packages if they were to integrate XML as their native file format in the future. However as long as the modifications remain open, anyone should be able to implement them in their own software solutions. However, even in the early stages of XML some small problems have already started to appear. Today XML is complex and needs specially trained operators. Nevertheless, PostScript drawing packages and other applications have demonstrated that it is indeed possible to make complex, programming-based processes reasonably simple to use. For XML to become as much a standard as Adobe's page-description language will require strong development efforts as well as broad end-user interest and education.

XML has a lot to offer as a data interchange tool, but for the main publishing market to move seriously to XML-based data structures, it will be necessary to re-engineer the applications and proprietary data structures extensively. This will not happen overnight. For example, while *avenue.quark* will outfit XPress 5 with XML import and export. However, there is little to indicate that the XPress file format will be rewritten to move closer to XML-based structures.

XML seems to reverse the common pattern of technology adoption that has driven much of the high-tech market. Practically all tools that have gained predominant market position have evolved from the ground up, starting as end-user applications and then becoming increasingly professional. If XML moves beyond vertical, high-end applications, its progress will represent the inverse of that standard operating procedure.¹²

The applications that will drive the acceptance of XML are those that cannot be accomplished within the limitations of HTML. These applications can be divided into four broad categories:

- Applications that require the Web client to mediate between two or more heterogeneous databases.
- Applications that attempt to distribute a significant proportion of the processing load from the Web server to the Web client.
- Applications that require the Web client to present different views of the same data to different users.
- Applications in which intelligent Web agents attempt to tailor information discovery to the needs of individual users.¹³

XML will be useful for delivering information to different devices and database like applications, this will undoubtedly make it a very useful tool for large organizations managing their data and to publishers targeting the new publishing media. However the home user does not seem to require most of this added functionality and complexity of XML. When writing a personal letter or a school report nobody would like to worry about DTDs or tagging the different parts of text. Most users do not even use most of the capabilities of Microsoft Word such as headings, captions etc. Therefore, there is no reason why

¹² Source: Pfeiffer Consulting 2000

¹³ Source: Pfeiffer Consulting 2000

home users should be initially interested in this aspect of XML. Allowing documents to be converted to a web format without losing important formatting aspects and the ability to transfer files among applications will be a more interesting feature for this segment of the market. Once again the full possibilities of XML will only be widely accepted and used when there is a large number of easy to use tools become available.

DOM and SAX

DOM and SAX are currently the two most popular APIs for manipulating XML documents. They differ significantly in provenance, in scope, and in programming style. They are not in direct competition with each other; each has strengths and weaknesses.

The Document Object Model (DOM) is a platform- and language-neutral interface that will allow programs and scripts to dynamically access and update the content, structure and style of documents. The document can be further processed and the results of that processing can be incorporated back into the presented page.

Unlike the DOM, SAX ("Simple API for XML") is not being developed by the World Wide Web Consortium. Instead, it was developed by an informal group of participants of the XML-DEV mailing list. SAX 1 has been fairly widely supported by providers of XML processing software. SAX 2, now being developed, is not yet widely supported, and at this writing diverges significantly from SAX 1, though it includes the SAX 1 APIs for backward-compatibility purposes.

2.2.8 XHTML

XHTML 1.0 is W3C's recommendation for the latest version of HTML. XHTML 1.0 is a reformulation of HTML 4.01 in XML, and combines the strength of HTML4 with the power of XML.

XHTML is a family of current and future document types and modules that reproduce, subset, and extend HTML 4. XHTML family document types are XML based, and ultimately are designed to work in conjunction with XML-based user agents. It is intended to be used as a language for content that is both XML-conforming and, if some simple *guidelines* are followed, operates in HTML 4 conforming user agents. XHTML 1.0 provides the following benefits:

XHTML documents are XML conforming. As such, they are readily viewed, edited, and validated with standard XML tools.

XHTML documents can be written to operate as well or better than they did before in existing HTML 4-conforming user agents as well as in new, XHTML 1.0 conforming user agents.

XHTML documents can utilize applications (e.g. scripts and applets) that rely upon either the HTML Document Object Model or the XML Document Object Model (DOM)

As the XHTML family evolves, documents conforming to XHTML 1.0 will be more likely to interoperate within and among various XHTML environments.

XHTML 1.0 is specified in three "flavors". Each variant has its own DTD - Document Type Definition - which sets out the rules and regulations for using HTML in a succinct and definitive manner. These three variants are summarized below:

XHTML Transitional - Most people writing Web pages for the general public to access will want to use this flavor of HTML 4. The idea is to take advantage of XHTML features including style sheets but nonetheless to make small adjustments to your mark-up for the benefit of those viewing your pages with older browsers which cannot understand style sheets.

XHTML Strict - Use this when you want really clean structural mark-up, free of any tags associated with layout. Use this together with W3C's Cascading Style Sheet language (*CSS*) to get the font, color, and layout effects you want.

XHTML Frameset - Use this when you want to use HTML Frames to partition the browser window into two or more frames.¹⁴

¹⁴ Source: World Wide Web Consortium (W3C)

2.3 COMPARISON

This section compares the 4 most popular and most promising document formats. All these formats have a place and can exist together.

Format	Advantages	Challenges	Most common usage
HTML	<ul style="list-style-type: none"> ○ Wide platform support can be viewed in any browser. ○ Good for delivering simple text ○ Conversion-to-HTML tools getting better 	<ul style="list-style-type: none"> ○ Viewing inconsistencies between browsers ○ Some features not available in early version browsers ○ No "document" concept, can't save or print files with fidelity 	<ul style="list-style-type: none"> ○ Delivery of static information: <ul style="list-style-type: none"> - Static information delivery . - "Output" from other systems (databases, XML, etc.)
XML	<ul style="list-style-type: none"> ○ Stores content as well as context (metadata) ○ Can display dynamic information as a document ○ XSL (eXtensible Style Sheet Language) could allow for more control over presentation. 	<ul style="list-style-type: none"> ○ Emerging standard, limited tools for authoring and viewing. ○ Presentation standard still in progress. ○ Limited industry expertise. 	<ul style="list-style-type: none"> ○ Delivery of dynamic data: <p>Object databases B2B E-commerce (Purchase Orders, Invoices)</p>
PDF	<ul style="list-style-type: none"> ○ Complete visual integrity (layouts, fonts, colors, pagination). ○ Platform and application independent. ○ Free Reader plug-in with extensive search and navigation features. ○ Optimized file compression and delivery. ○ Exact on-screen and printed output. ○ Ability to recreate paper documents into searchable, viewable electronic files. 	<ul style="list-style-type: none"> ○ Users need to install free Acrobat Reader. ○ May need print- and web-specific PDF renditions for high-end publishing applications. 	<ul style="list-style-type: none"> ○ Delivery of visually rich content: <ul style="list-style-type: none"> - Marketing materials - Published information - Documents originating from paper - Output from electronic sources (MS Word, Excel, PowerPoint, Quark, PageMaker, etc)
Word File (.doc)	<ul style="list-style-type: none"> ○ "De facto" standard for text base document creation ○ Allows editing and collaboration. ○ Can be imported by a large number of applications. 	<ul style="list-style-type: none"> ○ Proprietary format ○ Microsoft Office is the only application that can realistically use all the possibilities of the format. ○ Integration with the rest of MS Office software. 	<ul style="list-style-type: none"> ○ Creation of every kind of text based document with limited. ○ Collaboration in document creation.

2. 4. TOOLS FOR TEXT BASED DOCUMENT CREATION.

This section provides an introduction to some of the most popular and promising packages for the creation of documents composed of text and graphics. Details about these packages can be found in appendix A.

3. DISTRIBUTION

3.1 MULTIMEDIA PUBLISHING FOR WEB AND CD-ROM.

**SHOCKWAVE AND MACROMEDIA DIRECTOR, MACROMEDIA FLASH AND FLASH PLAYER,
MACROMEDIA AUTHORWARE AND AUTHORWARE PLAYER.**

The two most popular programs and formats to create both multimedia rich presentations for the web and/or for distribution in CD-ROM or DVD are Macromedia Director with its Shockwave files and Macromedia Flash. They both have similar characteristics and they can both be used to create CD-ROM/DVD presentations and multimedia web pages. Flash animations require the free Flash Player and Director animations require the free Shockwave player to display content on Internet browsers. Flash is more lightweight than Director and it is more oriented towards web publishing. Flash Player delivers low-bandwidth, high-impact Web sites and user interfaces. Shockwave Player delivers interactive multimedia product demos, e-merchandising applications, and rich-media multi-user games. Both Flash and Shockwave animations can be compiled in an executable file to be included in CD-ROMs or distributed in other ways. According to Macromedia on August 13th 2000 Shockwave has a penetration of 142 million people while Flash has a penetration 277 million people¹⁵. Another product of the Macromedia family for delivering visual rich-media is Macromedia Authorware. This product is specifically designed for creating Web and online learning applications to deliver on the Web, LANs, and CD-ROM.

Macromedia Flash

Macromedia Flash 5 uses vector graphics with bitmaps, audio, animation, and interactivity to create multimedia Web pages. It supports direct import from Macromedia FreeHand, Fireworks and provides powerful development tools for creating advanced Web sites and applications, including tight integration with Macromedia Generator. Flash 5 also supports import and export of MP3 audio. With MP3 compression, presentation can have long-form animations with voiceovers and background music but still keep files small enough for low-bandwidth delivery, animation frames can also be visually synchronized to streaming audio.

¹⁵ Details of how these numbers are calculated can be found at http://www.macromedia.com/software/player_census/

Macromedia Flash allows the integration of field entry forms for data gathering and e-commerce applications. Macromedia Flash Player can pass information to a Web server so that input from text fields can be used for password fields. All text entered into password-specified text fields will automatically convert characters to an indistinguishable form. Flash can be used to create Web application front-ends using *Get* and

Post actions to place text from and to a Web server easily. Flash can pass data to any CGI script for closer integration with Microsoft Active Server Pages, Allaire ColdFusion, or Macromedia Generator Enterprise Edition server.

From version 5, Flash allows developers to utilize XML structured data within their Flash-based Web applications for a broad range of e-commerce purposes. Using XML for rich data and Macromedia Flash for logical and intuitive user interfaces, companies can create sales forms, virtual shopping carts, customer surveys, and stock availability matrixes. Continuous XML connectivity allows immediate updates of any mission-critical displayed information.

Flash also offers Web-native printing, providing customizable development options for publishers. Macromedia Flash Web-native printing offers:

- WYPINWYS (What You Print Is Not What You See).
Printable content can be downloaded on demand, providing a faster site-viewing experience.
- Consumers can view content and applications suitable for the screen and output a separate design suitable for print.

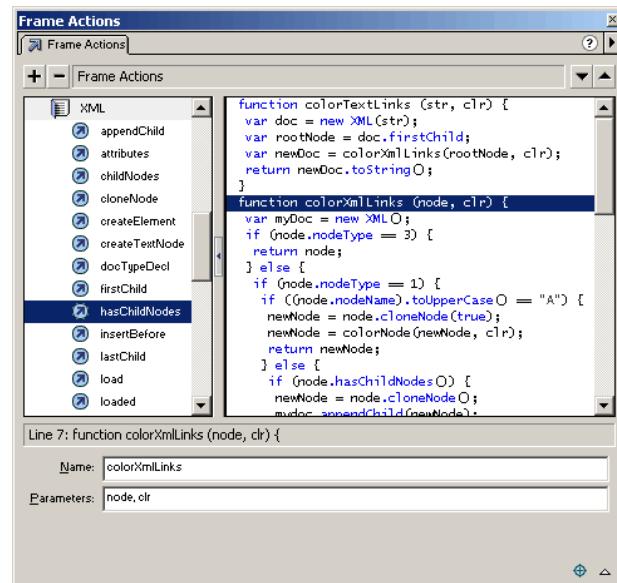


Figure 9- Standard XML objects in Flash 5 ActionScript enable dynamic Web applications and complex user interaction.



16

Figure 10 - This Hewlett Packard ad banner is 19K when viewed on a Web site and delivers an 8.5" X 11" sheet laden with product literature when printed.

¹⁶ Image from Macromedia.

Macromedia Flash 5 offers native support for importing, extending, and exporting Apple QuickTime 4 movies. Additionally, Apple has licensed the Macromedia Flash Player and built it into QuickTime, allowing QuickTime movies to include Flash graphics, animations, and interactivity. Flash 5 support for QuickTime 4 means that Web developers can incorporate compact Flash interface elements into QuickTime movies. Web users will be able to watch streaming video overlaid with Flash interfaces that stream over the Internet. Flash effects that can enhance a QuickTime movie include navigational controls, text effects, animation, titling, etc.

Macromedia Flash 5 also provides animators and developers the ability to publish RealFlash content for the RealPlayer G2 and RealPlayer 7 and 8. Flash 5 adds support in the Publish Settings to export all the necessary streaming RealAudio tracks, specially tuned Flash movies to add seeking support, and SMIL code to ensure syncing playback in the RealPlayer. In addition, with the new support for Macromedia Flash 4 content in RealPlayer 8, Macromedia Flash developers can create not only RealFlash animations but also streaming e-commerce experiences, such as adding items to a shopping cart while watching a video about a product.

Macromedia Flash Player

With Macromedia Flash content will play back consistently and reliably across a broad range of browsers, operating systems, and devices. The Macromedia Flash Player is currently distributed with every major browser, operating system, and Web-enabled appliance. Partners include Microsoft, Apple, AOL, Netscape, Prodigy, WebTV, RealNetworks, and Excite@Home. According to Macromedia more than 92% of online *users* can view Macromedia Flash content, without having to download a new player.

Macromedia Director

Macromedia Director is the market-leading multimedia solution for creating multimedia presentations, e-merchandising applications, and interactive entertainment. Director combines graphics, sound, animation, text, video, and multi-user technology with rich interactivity to create content that is easy to deploy as Shockwave content for the Web, or as multimedia for CD-ROM and DVD-ROM.

Macromedia Director can create anything from a simple navigation menu for a web page to simple games. Macromedia Director 8 Shockwave Studio streamlines the authoring process with centralized, automated functions that make it easier to manage assets, edit animations, build graphics, and create content for low- bandwidth delivery.

Macromedia Director incorporates the Lingo scripting language. Lingo scripting is a way to add interactivity, and data tracking, as well as script-driven animations and effects.

Director 8 Shockwave Studio has the ability to link scripts from external files. Scripts can be stored in a central file location or in a source-code control system for code management and version control.

Macromedia Director can import, control, and communicate with Flash

files to use Flash's animation, shape morphing, masking, and transparency. Director can also interact with live QuickTime streams like video broadcasts. QuickTime support includes streaming playback and MP3 audio in Shockwave Player and Live stream playback.

Shockwave can now download and use XML source as a building block for Web applications. Director parses XML elements, attributes, and character data, helping developers deliver complex constructs that can be controlled and manipulated without downloading large files.

Macromedia Shockwave Player

Shockwave Player is used to view Shockwave multi-user content. Shockwave Player is pre-installed with all MacOS and Windows 98 systems and it comes on most browser CDs. According with Macromedia over 109 million Web users already have Shockwave Player installed. Shockwave Player can be easily downloaded from macromedia.com. According to Macromedia, over 100,000 people successfully download Shockwave Player from Macromedia's Web site everyday. More than 109 million Web users—over 51% of people online—can immediately view Shockwave content, according to a recent independent study. Macromedia Flash and Shockwave have a stronger online presence than any other streaming video and media formats online, according to an October 1999 survey conducted by NPD Online Research, the parent company of Media Metrix.



Figure 11

¹⁷ Image from Macromedia.

Macromedia Authorware and Authorware player.

Authorware 5.1 has been optimized for creating interactive learning applications, allowing content authors to build in student response mechanisms such as buttons, hotspots (responsive to the cursor), hot objects, text entry, menus and logic-based events that generate an automated response. Authorware is designed to allow nonprogrammers to build interactive content through a drag-and-drop, visual environment that can be used over the Web, allowing student-teacher interchanges. The tool now includes hyperlinking of text documents and search and retrieval. Authorware also now includes a set of templates for quicker creation of instructional material, giving an instructor on the Web the means track student responses to material and otherwise track their progress.

Authorware can import and control Flash vector graphics and animations. It also supports QuickTime 4.0 to integrate new video, sound, and graphics formats into Authorware applications. In addition to supporting QuickTime video, QuickTime 4.0 adds support for MPEG, WAV, image formats, interactive QuickTime VR, and streaming video.

Cross platform compatibility.

In order to be widely accepted it is important that Flash, Shockwave and Authorwave animations can be accessed by as many people as possible. That is why all these products work at least under Windows and Macintosh OS, covering the great majority of web users. Flash is the most OS independent format as it also plays in Linux and UNIX systems. Table 2 details the operating system compatibility of these products.

Macromedia Web players	Win 95/98/NT/2000	Win 3.1	Mac PPC	Mac 68k	Sun Solaris	Linux x86	SGI IRIX
Shockwave 8 & Flash Players	•		•				
Shockwave 6 & Flash Players	•	•	•	•			
Shockwave 6		•					
Flash Player	•	•	•	•	•	•	•
Authorware 5.1 Player	•	•	•	•			

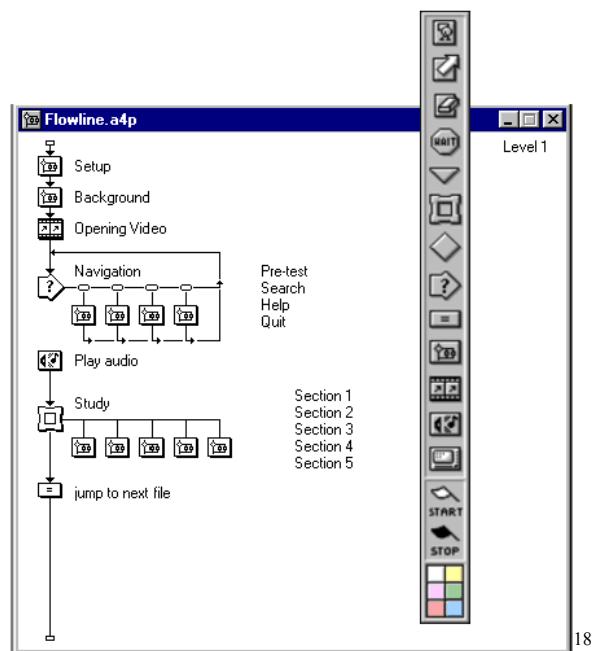


Figure 12

¹⁸

¹⁸ Image from Macromedia.

Table 2

3.2 WEB DEVELOPING TOOLS

See appendix B.

3.3 PUBLISHING FOR MOBILE DEVICES

In the next ten years one of the most important targets to be considered when publishing any kind of information in electronic format will be portable devices connected to the Internet. These include mobile phones, Personal Digital Assistants (PDAs) and a whole range of hybrid devices. The number of devices wireless devices continue to increase at an impressive rate; Swedish telecom equipment maker Ericsson in its 2000 industry-wide mobile phone sales estimate of 400-440 million units.¹⁹ In April 2000, research firm IDC estimated that by the end of 2002 more U.S. households will connect to the Internet through wireless devices than desktop PCs. When this happens, there will need to be a fundamental shift in the thinking of the Web community and the IT industry as a whole. Once there are more wireless Internet users than wired users, Web developers will consider the needs of users using traditional Web browsers and wireless users at the same time when creating websites. Today, Internet sites have to be modified for wireless users, but three years from now, it is conceivable some of them might have to be redesigned for traditional browser users.²⁰

European markets are currently one or two years ahead of most of the World in mobile devices, with only Japan offering more advanced services. Currently 117 million Europeans have mobile phones, most of them digital. Finland, Spain, the UK and Germany have already awarded the licenses for the third generation (3G) mobile phone. Six bidders have agreed to pay a total of €50.5bn (US\$46bn) for the license in Germany. Britain's auction raised a further €37bn and Italy is currently (August 2000) auctioning its licenses.

The United States has suffered from multiple standards but the adoption of a single third generation (3G) standard will create the platform for a true nationwide Mobile Internet service. Mobile penetration is currently 30%, relatively low compared with other regions. However, it is predicted that in the next two years, 80% of the mobile phones will have mobile data capabilities.²¹

¹⁹ Source: CNNfn

²⁰ Source: IDC, division of IDG.

²¹ Yankee Group.

Figure 13

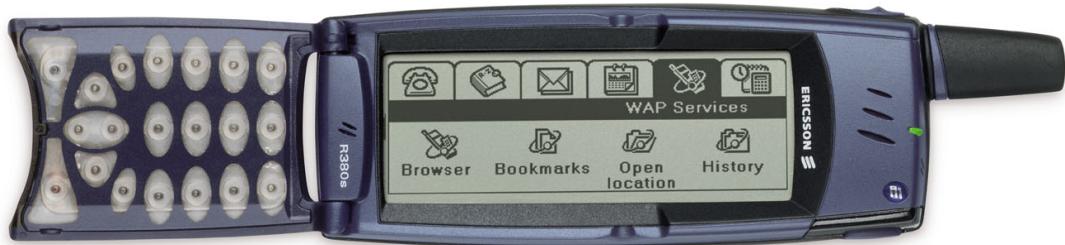
In this decade mobile devices should be increasingly able to connect to each other and to the Internet. Connection speeds will increase thanks to the new protocols and CDMA networks, with increased production the prices for connectivity and for the devices should be reduced and they should gain widespread use.

With the exception of the Nokia Communicator series and a few others, today there is still a clear distinction between a portable phone and a small PDA or portable computer. Although there are different opinions of whether these devices should converge (See Figure 14) or should remain separate but capable of communicating with each other via infrared or Bluetooth radio technology (See Figure 13). It seems that most devices will probably converge into one the next decade, but there is no need for a unique solution.



Ericsson MC218 and mobile phone²²

Figure 14



Ericsson R380 (source: Ericsson)

Unless small head mounted display devices similar to the Sony Glasstron (See Figure 15) become cheaper and smaller, handheld devices will always have small screens as they are by definition designed to be light, small and portable.

Therefore content intended to be viewed in these devices must be formatted to fit these screens. Several solutions are currently available to provide Internet style information to these devices. The most widespread are WAP and I-mode, described below.

Figure 15



Sony Glasstron²³

²² Source: Ericsson

²³ Source: Sony

3.3.1 WAP AND I-MODE

Wireless Application Protocol (WAP)

The Wireless Application Protocol (WAP) is an open, global standard for mobile communication. WAP allows mobile users to access and interact with different informational services. Ericsson, Nokia, Motorola and Unwired Planet (now Phone.com) co-founded the WAP Forum in June 1997 to provide a worldwide standard for the delivery of Internet-based services to the mass-market of wireless devices. These devices include mobile telephones, pagers, personal digital assistants (PDAs) and other wireless terminals. The WAP Forum controls and manages the continued development of the Wireless Application Protocol. Today, there are no less than 195 companies from around the world that have joined the WAP Forum.

WAP defines an XML (eXtensible Markup Language) syntax called WML (Wireless Markup Language). WML does for mobile devices what HTML does in full size computers. WML is specifically designed to display information on small screens; it supports links but does not support graphics. WML's basis in XML also positions it well as a future target markup language for automatic content transformation. Using an XSL style sheet, content written in XML-defined markup languages can be automatically translated into content suitable for either HTML or WML (See Figure 8 section 2.2.7). Likewise, content written in well-formed WML can also be translated to other XML-based markup languages, using a different XSL style sheet.²⁴

WAP empowers mobile users of wireless devices to access live interactive information services and applications from the screens of mobile phones. Services and applications include email, customer care, call management, unified messaging, weather and traffic alerts, news, sports and information services, electronic commerce transactions and banking services, online address book and directory services, as well as corporate intranet applications.

WAP is not available on the main wireless service providers in the U.S., Sprint's Wireless Web, ATT's PocketNet, nor Verizon's Mobile Web services are actually WAP. According to Scott Goldman, CEO of the WAP Forum There are currently no commercially available WAP services in the United States today. All of the above companies use a proprietary technology from Phone.com which uses HDML (Handheld Device Markup Language).²⁵

²⁴ Source Phone.com

²⁵ Source: InfoWorld Media Group Inc.

WAP has been an extremely over hyped technology, often ignoring the limitations of the current devices for data input, their speed and limited screens. Wireless phones can currently transmit data only at 9.6 kbps, this added to the small low resolution monochrome screens makes the use of graphics impractical and writing text using a conventional mobile phone keyboard is very restrictive. At the moment only a small percentage of mobile phones use WAP and the available WAP content is limited, although this, along with the data transmission rate, will change soon. Some systems try to convert normal Internet content to a format that can be understood by mobile phones. Passcall's GateWave's incorporates the "Intelligent Transformation" technology. Intelligent Transformation analyzes existing HTML websites and "intelligently" transforms them, in real time, to formats suited to end-users' mobile devices including WAP.

IDC predicts that by mid-2001, all digital cellular/PCS handsets shipped in the world will be WAP-capable and obviously more content should be available everyday. Looking at some of the new Japanese devices and with the fast development in technology in this area, it seems like the current WAP technology will quickly become obsolete (see Figure 16). As said before the WAP Forum was created in 1997, by 2002 it will mean that WAP's ideas are five years old, which is an extremely long time in the world of mobile communications.

Figure 16



Kyocera VisualPhone VP-210
(Designed only for use in PHS networks in Japan, provides video conference at 2 frames per second)

i-mode

The *i-mode* information service was launched in Japan in February 1999 by NTT DoCoMo. By the summer 2000 it has already captured four million subscribers, representing 13% of DoCoMo's subscriber base and nearly 8% of Japan's 49 million mobile users. DoCoMo has created a network of partners which offer specially formatted websites to fit into the small screen on the mobile handset. The service boasts 311 websites linked to its portal page, and 4,362 "unofficial" web pages created by private individuals.²⁶

I-mode is served by "compact HTML," or CHTML, which technically can allow users to also access desktop HTML sites, although it looks better if it has been written in CHTML.

i-mode is based on a per-packet charging (users pay a basic subscription fee per month and an amount per 128 bytes of information received). Subscribers are permanently connected to specifically designed Internet pages offering information, Internet browsing, translation services, and network games. *i-mode* has also experienced the reverse of the European trend with Average Revenue Per User (ARPU) actually rising.²⁷ NTT DoCoMo's *i-mode*, is a proprietary service only offered in Japan and cannot be made readily available on any other service carrier's network. In comparison, WAP boasts the

²⁶ Source: Financial Times, August 2000.

²⁷ Source : KPMG – Mobile Internet “Any Place , any time , everything”

worldwide support of over 500 major phone carriers and manufacturers that have been working together to ensure that their services are compatible with each other.

Broadband wireless services

In order to make the distribution of content to mobile devices in a less limited way than WAP, higher bandwidth for these devices is needed. New improvements into the wireless networks will allow to deliver multimedia content to mobile devices. This increase in speed added to future improvement in higher quality, low battery consumption screens will bring video and online books to handheld devices and/or mobile phones.

The first broadband services will be built on existing second generation (2G) digital networks. The first step will be the 2.5G services.

High Speed Circuit Switched Data (HSCSD) extends the current GSM systems to allow for improved data rates, and it is cost efficient to implement requiring only software upgrades. HSCSD is already online from some European Operators. HSCSD should increase wireless data access speeds on phone networks from 9.6Kbps to 57.6 Kbps.

General Packet Radio System (GPRS) will introduce packet based technology and will theoretically increase data speeds up to 115kbps. GPRS will also allow devices to stay connected to the network all the time, while users are billed only for the time spent actually sending or receiving burst of data.

Enhanced Data for Global Evolution (EDGE), will be a higher bandwidth version of GPRS, offering speeds up to 384kbps and enabling high-speed, mobile multimedia applications. EDGE will work in GSM and TDMA networks and should allow phones to work every region of the World.

The first third generation (3G) service is expected to be deployed in Korea by the end of this year, and similar service is expected in the United States and Japan sometime in 2001. However it will not reach the estimated bandwidth of 2Mbps and will require a new network.

Universal Mobile Telephone System (UMTS)

UMTS™ (Universal Mobile Telephone Service) is a Third Generation (3G) mobile system being developed by ETSI™ within the (International Telecommunications Union) ITU's IMT-2000 framework. IMT-2000 is now being standardized by the International Telecommunication Union (ITU) as the global standard for the third-generation mobile communications system. UMTS is a European System which is attempting to combine cellular, cordless, low-end wireless, local area network, private mobile radio, and paging system.

UMTS will offer the first true

3G IP native broadband native network. Although not available until next year at the earliest, it will offer speeds up to at least 384 kbits/s, with later stationary bandwidth up to a maximum of 2Mbits/s. This will run on CDMA networks and should be able to deliver video to mobile devices such as the Nokia concept in Figure 17. Already available CDMA chips also include the necessary processing power to handle video and other multimedia applications.

Figure 17



Nokia 3rd Generation mobile phone concept.²⁸

²⁸ Source: Nokia

3.4 ELECTRONIC BOOKS

Some single-purpose devices which only function is to display reading matter in a book like fashion have appeared in the recent years. However, none of the electronic book devices has become very popular, mainly because they provided a lower quality reading experience at an initial higher cost than traditional paper books.

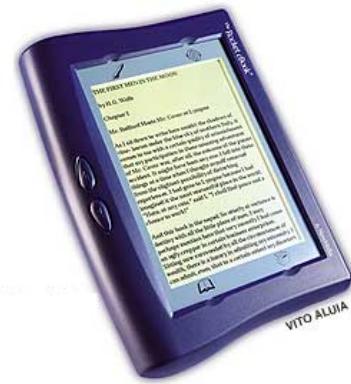
The first handheld electronic dictionary that was introduced by Franklin Electronic Publishers in 1986, displaying a mere one line of text at a time. Five years later, Sony's Data Discman displayed disc-stored books on a 8-centimeter screen, still too small for comfortable book reading. By 1999, the hardcover-sized SoftBook Reader from SoftBook Press and the paperback-sized Rocket eBook from NuvoMedia (See Figure 18) attracted media attention to the current generation of e-book. However these dedicated devices are still expensive, between US\$200 to US\$600. For those prices it is possible to buy either a Personal Digital Assistant (PDA) or even a low-end computer, which can perform many more tasks than only displaying electronic books. Another big drawback is that some readers' use proprietary file formats, requiring that additional versions of electronic literature be made available for downloading.

In January 2000, both SoftBook Press and NuvoMedia were acquired by Gemstar International Group. Gemstar aims to re-launch versions of both readers, manufactured by Thomson Consumer Electronics. They expect to sell about half a million units a year. However, even if these predictions are accurate this is still a small number compared with the number of other electronic devices capable of displaying an electronic text that are sold every year. The proliferation of lighter, cheaper and more powerful portable computers and personal digital assistants (PDAs) with better screen resolutions will probably mean that these single purpose reading devices will not reach widespread use for the moment. Therefore the term electronic book will refer more and more to a file containing a book.

It is important that there is a small number of formats to transmit electronic books if they are ever going to be widely distributed. Apart from proprietary formats to use with some specific electronic book devices the three main contenders are OEB, PDF from Adobe Systems (see Section 2.2.4 and Appendix A.6) and Microsoft Reader.

Microsoft released Microsoft Reader in the summer of 2000. In conjunction with the release, Barnes & Noble launched a MS Reader-specific on-line bookstore with 2,000 titles initially available (Barnes& Noble currently supports both MS reader and PDF format). Microsoft Reader has already gained support from important electronic book distributors. On May 2000 Microsoft announced that MS Reader would be the reading format for the iRead channel of iPublish.com at Time Warner Books and on August 2000 Amazon announced that they will sell electronic books over the internet using the MS Reader format.

Figure 18



Rocket eBook from NuvoMedia²⁹

²⁹ Source: MIT Technology Review.

MS reader will also be available for electronic book devices. Franklin Electronic Publishers Inc. announced on June 2000 that they will incorporate Microsoft Reader into their new line of next-generation multimedia eBook devices.

MS Reader is free for download in Microsoft's website and will be given away with every new copy of Windows. Microsoft Reader also gives the possibility for users to convert their Word documents to MS Reader format.

MS Reader features the ClearType technology that is supposed to improve the clarity of text on the screen (See Figure 19). To people accustomed to reading text on a computer for hours at a time it may seem clearer to read text in a normal word processor than with Microsoft Reader.

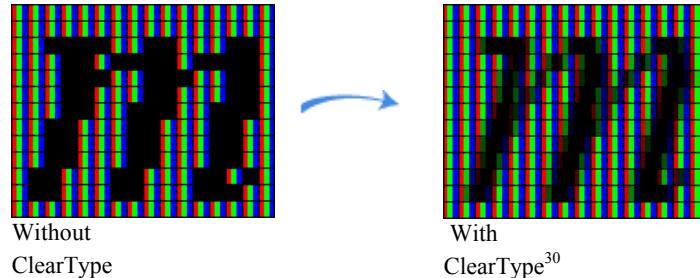


Figure 19

Microsoft Reader also provides copy protection for authors and booksellers. But while e-books rights management may be important to intellectual property holders it has been proven than any PC-based copy protection scheme can be cracked. It happened for example within two days of Stephen King's first electronic publication.

MS Reader has a relatively good chance of gaining a large user base. There are several reasons for this:

- It has Microsoft behind it and has support of big distributors such as Amazon.
- It will run in both Windows and PocketPC and will come as a standard with these operating systems.
- Converting files from Word is free.
- The ClearType technology (either for its real value as an improvement or as a marketing tool).

Another technology with interesting implications in the distribution of electronic texts is PDF, which is extensively discussed in section 2.2.4. With more than 110 million Acrobat readers already downloaded onto computers, PDF is currently the most popular format for e-book publication. PDF was specifically designed for preserving professional-quality documents across computer platforms and printers. PDF lets users print out the files with the confidence that they will look the same as in the screen. Adobe PDF is by far the most mature format for electronic books and supports the biggest array of options.

However, the software to create PDF files in a simple way is not free (there are freeware solutions but they have limitations and are not as user friendly) and will not come installed with Windows, as Microsoft Reader. Although this is not a major problem for big publishers, often the market develops from what the end users adopt, and if home users like the idea or making MS Reader files of their work it may have a serious impact in the choice of a standard. This may eventually force Adobe to make Distiller (See Appendix A.6) free and continue selling Acrobat as an advanced tool to edit PDF files.

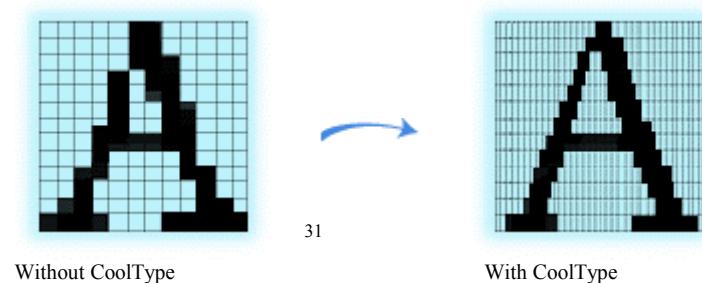
³⁰ Source: Microsoft Corp.

To counter Microsoft Reader, Adobe has recently added e-commerce encryption software called PDF Merchant, allowing rights to an electronic copy of a book to be assigned to a single computer (see Appendix A.6).

Adobe has also challenged Microsoft's ClearType with screen-enhancement routines of its own, which it calls CoolType; the competing technologies are very similar. CoolType creates clearer, crisper type using a font-rendering technique called color anti-aliasing, which works on digital liquid crystal display (LCD) screens such as those in laptops, handheld devices, and flat-panel desktop monitors. Unlike conventional anti-aliasing, which manipulates only whole pixels, CoolType controls the individual red, green, and blue subpixels on a digital LCD screen.

By adjusting the intensity of the

subpixels independently, the strokes of a character can be aligned on any subpixel boundary, thus achieving sharper, more precise smoothing along the edges of characters (See Figure 20)



Adobe has also responded to the competition created from Microsoft Reader by buying a leading developer of PDF-based book-reading software, Glassbook. The acquisition is intended to strengthen the PDF format in the electronic book business. Adobe will be now maintaining both Glassbook and Acrobat Reader, which essentially do the same thing, i.e. display PDF files. This is a common practice for Adobe that often maintains products with similar capabilities but different presentations oriented to different sectors of the market. According to Adobe, the Glassbook Reader will eventually be able to support multiple operating systems including Mac, Linux, and Palm, in addition to Windows. The Glassbook Reader will also display e-books that use the new Open eBook Publication Structure (*OEB*) format.

The PDF format also has important industry support. By mid-fall 2000 McGraw-Hill's own website, Primis Online, will offer one hundred textbooks in PDF format. From August 2000, Adobe also has an alliance with BarnesandNoble.com to become the exclusive e-book retailer of Adobe content.

iUniverse.com has also an agreement with Adobe to use the PDF format along with many of Adobe's publishing tools for both electronic book distribution and print on demand books (see section 3.6).

Another important format is the new Open eBook (*OEB*) standard created by a consortium of e-book hardware makers, traditional publishers, and other companies including Microsoft and Adobe. The main difference between *OEB* and PDF is that PDF locks in a book's formatting so it can be preserved intact across output devices; once created, it is not meant to be modified. This can be a drawback if an author or publisher wants to access parts of the text for excerpting or reconfiguring for a customized e-book, or for sampling or sale in smaller increments than book length. *OEB* allows an e-book's content to be reformatted on the fly, using a markup language that is essentially an extension of HTML. The Open eBook standard also makes it easy for dedicated reading devices to reformat text to fit their proprietary display configurations.

³¹ Source: Adobe Systems

The first published specification for OEB addresses neither security nor e-commerce protocols, leaving it to individual vendors to come up with their own approaches. This omission raises the possibility that the proposed standard could splinter into a variety of incompatible implementations.

Another interesting file format for electronic publishing is called TK3. Introduced by Night Kitchen—a New York startup—TK3 is the basis for a sophisticated literary software environment. The Night Kitchen TK3 Reader offers a book like reading experience on a desktop or laptop computer screen—complete with highlighting, corner-folding bookmarks, even Post-it-like “stickie notes.” TK3 also adds multimedia features such as sound and video. However, for the moment the TK3 format has not the support that Reader and PDF have from important companies like Microsoft and Adobe.

Although a unique standard would be desirable it is not completely necessary for electronic books to succeed. Ultimately, OEB, PDF and the MS Reader format could coexist, with the rival formats used for different output stages of the same e-book—OEB in the intermediate stages and PDF or MS Reader for final versions. This is why both Microsoft and Adobe are supporting the OEB standard. However the biggest benefit of OEB is for publishers, that will obtain a format that can be easily converted to both PDF and MS Reader if necessary. New electronic book reading software could also be able to read two or three formats (OEB will be an open standard and PDF is partially open) and automatic conversion tools could allow the online publisher to convert and deliver a book in the desired format starting from, for example, an OEB or XML version of the book.

Despite the hundreds of millions of PCs in use around the World, only a few hundred thousand of their users have downloaded e-books. The slow start is partly due to the perception that an e-book does not fully replicate the book-reading experience. Furthermore, people have to be used to the idea of downloading a book, instead of buying it from a shop. E-books are not necessarily trying to eliminate paper publishing. Most e-books attempt to replicate traditional books’ content and appearance and most e-books can be printed out with only minimal loss of information (primarily broken hypertext links). And for all their seeming differences, print and electronic publishers are releasing similar content. Eventually, digital downloads may be destined to become just one more format for readers, one more step on the convenience/cost continuum from hardcover to paperback to e-book.

At some point in the future, however, e-books and print are bound to diverge. Lurking amidst e-publishing today is the notion of multimedia books that seamlessly incorporate hypertext, sound and animation, some examples of this kind of “books” are already available. A hypertext branching narrative in a novel or a history book, for instance, would be impossible to reproduce in a book. However this does not necessarily represent a threat for paper or for the linear nature of written text. In the same way video did not make photography become obsolete, these new kind of multimedia “books” may become a completely different form of publishing.

For now, the advent of e-books means not replacing print, but supplementing it—redefining publishing economics and opening the way for authors whose work has been kept from appearing between book covers. Despite some of the over optimistic predictions about electronic books, it seems traditional paper books and book stores will co-exist with them for a long time. Most people alive today will prefer a hard cover paper book for the rest of their lives and Book Shops such as Borders and Barnes and Noble continue growing in the US and have started to expand to Europe.

3.4.1 PLATFORMS

Apart from specific software for mobile phones these are the most popular platforms for portable devices currently available. Although most of the devices that use them still do not include wireless data capabilities, they all can connect to the Internet and/or WAP information using add-on devices or connecting to mobile phones.

Furthermore, future versions of all these devices will include some kind of wireless connectivity and these operating systems are already ready to accept it.

The leading platform in the PDA market is Palm OS, especially in the United States. Microsoft's operating system for portable devices, Microsoft Windows CE, has evolved into Pocket PC. Windows CE was very criticized particularly for trying to simulate Windows for the desktop in a device which size and design required different ergonomic arrangements, Pocket PC has somehow improved this. It also includes the new Microsoft Reader to display electronic texts. The main challenger to Microsoft Pocket PC and Palm OS is EPOC, which is able to handle Wireless application protocol (WAP), can support Bluetooth and can use Java. EPOC comes with PSION organizers, the Ericsson MC218 and a few others. All these devices are still far more popular in Europe than in the US where they are many times ignored even in specific PDA publications. EPOC is supported by Symbian. Symbian is a joint venture between Panasonic, Ericsson, Nokia, PSION and Motorola. Its objective is to develop core software, application frameworks, applications and development tools for wireless information devices and creating standards for the interoperation of these devices with wireless networks, content services, messaging etc.

The Linux operating system has also been ported to run in smaller devices. It can currently run in some of the Pocket PC and PSION devices and new devices such as the Samsung *Yopy* prototype or the Agenda Computing



Figure 21 – Ericsson MC128³² and PSION REVO running EPOC32.



Figure 22 – Compaq iPAQ³³ and HP Jornada³⁴ running Microsoft Pocket PC

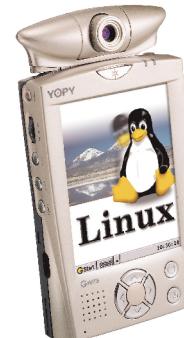


Figure 23 – Samsung Yopy prototype running Linux³⁵



Figure 24 – Palm V running Palm OS.³⁶ 40

AgendaVR3 already have Linux installed. Linux has the advantage of being completely open and although the number of applications for these devices is currently very small compared with Palm OS, EPOC or PocketPC/WindowsCE they are expected to increase rapidly.

3.5 ELECTRONIC PAPER

Electronic paper is a new technology that falls somewhere between paper and a conventional computer screen. It tries to combine the best qualities of paper and a computer screen. From paper it has:

- It has a user-friendly interface.
- It is thin and light weight.
- It is flexible; it does not break if it falls on the ground.
- It is not a source of light so it does not strain the reader's eyes as computer screens.
- Low battery consumption (Electronic paper does not need a backlight and can be read even if the power is turned off).
- It provides high contrast and wide viewing angle.
- Cheap (Electronic paper should eventually be very cheap it is produced in large quantities).

But like a computer display, it is also dynamic and re-writable.

Electronic paper could be used for many different applications, from newspapers that update themselves, to a new way to display electronic books downloaded from the Internet. It can also be used for wall size electronic whiteboards, bill boards and portable fold up displays. In fact J.C. Penney and Yahoo! Inc. have already used this technology in trials for advertisement displays.

The first electronic papers available will only be in black and white and will not be able to refresh the images fast enough to provide video. Therefore, they will not replace normal screens, at least for some time.

Electronic paper should be available anytime between the next three to ten years, then, as with any other technology, it will take some time until it escapes the initial high cost and low user base circle.

There are two different projects developing similar types of electronic paper. One of them is composed by the start-up E-ink with Lucent technologies as a partner; the other one is composed by Xerox and 3M.

3.5.1 XEROX / 3M

Xerox's Palo Alto Research Center (PARC) calls its electronic paper display technology "Gyricon". A Gyricon sheet is a thin layer of transparent plastic in which millions of small beads, somewhat like toner particles, are randomly dispersed. The beads, each contained in an oil-filled cavity, are free to rotate within those cavities. The beads are "bichromal," with hemispheres of two contrasting colors (e.g. black and white, red and white), and charged so they exhibit an electrical dipole (see Figure 25). When voltage is applied to the surface of the sheet, the beads rotate to present one colored side to the viewer.

Voltages can be applied to the surface to create images such as text and pictures. The image will persist until new voltage patterns are applied.

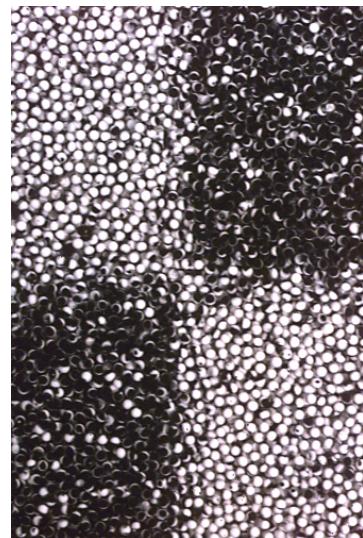
There are many ways an image can be created in electronic reusable paper. For example, sheets can be fed into printer-like devices that will erase old images and create new images. Printer-like devices can be made so compact and inexpensive that you can imagine carrying one in a purse or briefcase at all times. One envisioned device, called a wand, could be pulled by hand across a sheet of electronic reusable paper to create an image. With a built-in input scanner, this wand becomes a hand-operated multi-function device -- a printer, copier, fax, and scanner, all in one.

For applications requiring more rapid and direct electronic update, the Gyricon material might be packaged with a simple electrode structure on the surface and used more like a traditional display. An electronic reusable paper display could be very thin and flexible. A collection of these displays could be bound into an electronic book. With the appropriate electronics stored in the spine of the book, pages could be updated at will to display different content.

For portable applications, an active matrix array may be used to rapidly update a partial- or full-page display. Gyricon displays do not require backlighting or constant refreshing, and are brighter than today's reflective displays. These attributes will lead to Gyricon's utilization in lightweight and lower-power applications.

Xerox Corporation has selected 3M as the manufacturer to bring to market its electronic paper. The technology, supported by a portfolio of Xerox patents, has been prototyped at PARC on a limited scale, Xerox collaboration with 3M establishes a means by which the electronic paper material can be manufactured in the volumes necessary to meet market demand and to meet market demands and make the developments of a wide range of supporting applications commercially viable.

Figure 25



3.5.2 E-INK / LUCENT TECHNOLOGIES

The “ink” developed by E-Ink is a liquid that can be printed onto nearly any surface; within the liquid there are millions of tiny microcapsules, each one containing white particles suspended in a dark dye. When an electric field is applied, the write particles move to one end of the microcapsule where they become visible (See Figure 26). This makes the surface appear white at the spot. An opposite electric field pulls the particles to the other end of the microcapsules where they are hidden by the dye (see Figure 27). This makes the surface appear dark at that spot.

E-Ink has a partnership with Lucent Technologies; this allows them to combine the experience of both companies in digital paper and printable organic transistors.³⁷ Furthermore, E-Ink has received investments from Hearst Interactive media, Motorola and IBM.

Immedia Display products already produces world's first commercial electronic ink products. Displays are thin, lightweight, versatile and easy to read like paper. The "printed" messages can be changed by remote control.

E Ink foresees many applications for Electronic ink, including clothes, PDAs and watches. In fact Swatch made an early attempt to incorporate this technology in their watches. Thanks to technology produced by Lucent ‘s Bell Labs they believe that the price of electronic paper can be similar to the cost of traditional paper once it is produced in very large quantities.

Figure 26

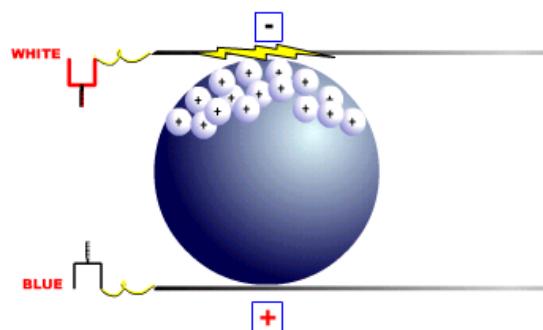
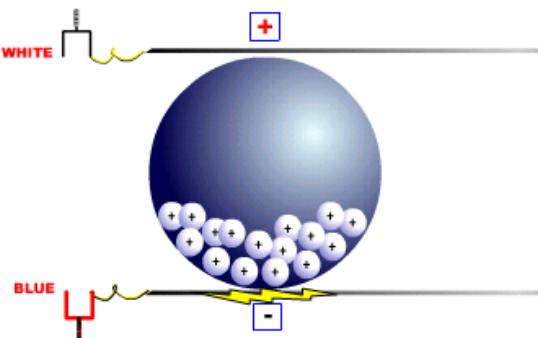


Figure 27



³⁷ Source: ZDNET News, October 1999.

3.6 PRINT ON DEMAND

Currently when books are published, a large number of them are printed at once and distributed to bookstores and other merchants, with the remainder stored in warehouses. If publishers greatly overestimate the number of books desired they must absorb the expense of unsold books. On the other hand, print runs using standard printing equipment are expensive unless done in bulk.

Print on Demand consists in printing books one at a time, at the time the customer wishes to purchase them. This may be more practical today than in the past due to the recent availability of high-end, high-volume publishing systems such as the Xerox DocuTech Network Publisher. If this effort does prove to be practical, it may dramatically change the book publishing industry. For instance, books would not need to go out-of-print, and could be customized when printed to have large type, use a special kind of paper, etc. As another example, small-run, specialized books, such as academic monographs, should become less expensive to produce. As a result, many more books will be available than in the past, with no additional inventory space required.

Print on demand makes use of an online, or virtual, bookstore. After making a selection at the virtual bookstore, one copy is printed for the customer, either at home if an appropriate printing device is available, or at a print shop. Some advantages of a virtual over a physical bookstore are: its contents are searchable, browsing can be done in multiple ways instead of the static shelf arrangement of a physical bookstore, a larger selection of books can be available, the bookstore is accessible anytime from anywhere, and it can provide auto-recommendation from an appropriate community of experts.

Print on demand services are already available although none of them have yet become very popular or attracted considerable media attention. Barnes & Noble teamed up with IBM in December of 1999 to provide print-on-demand facilities at book distribution centers. Barnes & Noble and IBM have been working with a pilot group of publishers as well as publishing portal iUniverse.com, in which Barnes & Noble (B&N owns a 49% interest in iUniverse.com). May more smaller companies also provide print on demand services at a smaller scale.

4. DIGITAL IMAGING

4.1 NEW IMAGE FORMATS

4.1.1 PORTABLE NETWORK GRAPHICS (PNG)

On Dec. 28, 1994, CompuServe (inventors of the GIF format) reached an agreement with Unisys (owners of the compression format used inside GIF) that required the payment of royalties whenever GIF functionality is used in a software application.

In consequence a group of independent software developers released the first draft of the “Portable Bitmap Format”, which the developers hoped would serve as an open-source, freely available alternative to GIF. In just two months, the concept reached final draft status, and officially became known as the PNG (Portable Network Graphics) format. The Portable Network Graphics (PNG) format was designed to replace the older and simpler GIF format and, to some extent, the much more complex TIFF format.

For the Web, PNG really has three main advantages over GIF: alpha channels (variable transparency), gamma correction (cross-platform control of image brightness), and two-dimensional interlacing (a method of progressive display). PNG also compresses better than GIF in almost every case, but the difference is generally only around 5% to 25. One GIF feature that PNG does *not* try to reproduce is multiple-image support, especially animations; PNG was and is intended to be a single-image format only. A PNG-like extension format called MNG (see section 4.1.2) was finalized in mid-1999 and is beginning to be supported by various applications such as Animation Shop included with Paint Shop Pro (See section xx), but MNGs and PNGs will have different file extensions and different purposes.

For image editing, either professional or otherwise, PNG provides a useful format for the storage of intermediate stages of editing. Since PNG's compression is fully lossless--and since it supports up to 48-bit true color or 16-bit grayscale--saving, restoring and re-saving an image will not degrade its quality, unlike standard JPEG (even at its highest quality settings). And unlike TIFF, the PNG specification does not allow implementors to choose what features they will support; the result is that a PNG image will be compatible with any PNG-supporting application.

Since there is currently no realistic method of transferring such large files to the majority of Internet users, PNG is usually seen in 8-bit or 24-bit versions. But, as access alternatives such as cable, DSL and satellite become commonplace over the next decade, Web designers will be able to use PNGs to produce color on the Internet with the same degree of accuracy that is currently seen only in high-end print.

One of the reasons for the limited availability of PNG files in the Internet is that Microsoft and Netscape, both failed to include PNG support in their 3.x browsers although the technology was already available. Both companies' 4.x browsers and the Windows version of Microsoft's 5 browser have implemented support for PNG but have failed to correctly implement PNG alpha transparency. Microsoft's WebTV browser is now in compliance with PNG's 24-bit color/8-bit transparency

guidelines, and Internet Explorer 5.x for the Macintosh is also reported to be in 24-bit color/8-bit transparency compliance.

Another reason for the still limited use of the PNG format is that only software vendors pay to use GIF files and the average web designer is not concerned with paying royalties. Therefore with the current browser support PNG graphics offer little improvement over GIF.

4.1.2 MULTIPLE IMAGE NETWORK GRAPHICS (MNG)

MNG stands for **M**ultiple-**I**mage **N**etwork **G**raphics. MNG evolved as an animated version of the PNG format, and carries with it PNG's feature set, including the ability to save true-color images and alpha channel transparencies.

On average MNG slightly outperforms GIF compression. But by using MNG tricks such as *sprites* and *nested loops* (see later), it should be possible to increase compression by factors of ten to 100. The main advantages of MNG over animated GIF are:

- Patent-free status of MNG versus UNISYS claims for royalties from GIF
- Platform-independent color via gamma, chromaticity and ICC profile information embedded in the MNG datastream.
- Good lossless compression.
- Moving sprites - Moving sprites can float above a background image and maintain full alpha channel transparency for pixel-perfect blending into the background. Sprites allow an image to be superimposed to the background and to create motion that superimposed image will be move that image from one point to another without having to store the background image.
- Nested loops - By using the sprite method of animation, entire image frames that are repeated can be removed.
- Transparent MNG includes support for a special implementation of JPG, called JNG (**J**PEG **N**etwork **G**raphics). This will allow JPEG style compression to be used inside of a MNG. JPEG (JNG) can be used when some quality loss in acceptable in the compression, this produces “higher compression”

Developers of the MNG format reckon that it could easily take three to six years for MNG-equipped browsers to saturate the marketplace.

4.1.3 JPG2000

The International Standards Organization's JPEG2000 committee has finalized specifications for a new algorithm that can compress images up to 200 times with a very small degradation in quality (see Figure 28) . The JPEG2000 specification, which will become ISO 15444 when it's officially approved in 2001. JPEG 2000 promises to keep the advantages of the current JPEG format, such as 24-bit color, tight compression, and cross-platform compatibility while adding some impressive enhancements of its own.

JPEG 2000 will use an entirely new compression scheme for the storage of image data. The current compression format, called DCT (Discrete Cosine Transformation), compresses image data into square blocks. The blocks are then stored in numerical order, representing data from the top to the bottom of the image. The new compression scheme, called Wavelet Compression, will store image data as a stream of information, with the stream containing more and more resolution as it opens or downloads. One immediate benefit to this “stream of information” concept is that it will help eliminate “JPEG Blotches” (the artifacts that appear around sharp edges in JPEG images). The information stream will also make it possible to download only as much resolution as is needed for a particular JPEG image.

JPEG 2000 offers additional features, such as the ability to add encrypted (coded) copyright information to a JPEG file. An encrypted copyright message inside a JPEG 2000 could survive common image editing procedures such as resizing and resampling and would be more difficult to remove than current “watermarking” technology.

JPEG 2000 also offers the ability to add “channels” of additional data to a JPEG file. This data could include color space information (such as CMYK descriptions) and ICC profiles (for eliminating PC versus Macintosh Gamma differences). A single image will be able to hold up to 256 compact data channels, so hardware and software vendors will be able to add information to these channels that could be particularly useful to their customers. For example, a software vendor could add a keyword library channel to a JPEG 2000 file, while a hardware vendor (such as a digital camera manufacturer) could add a camera profile that will help to ensure appropriate reproduction of a digital camera image.

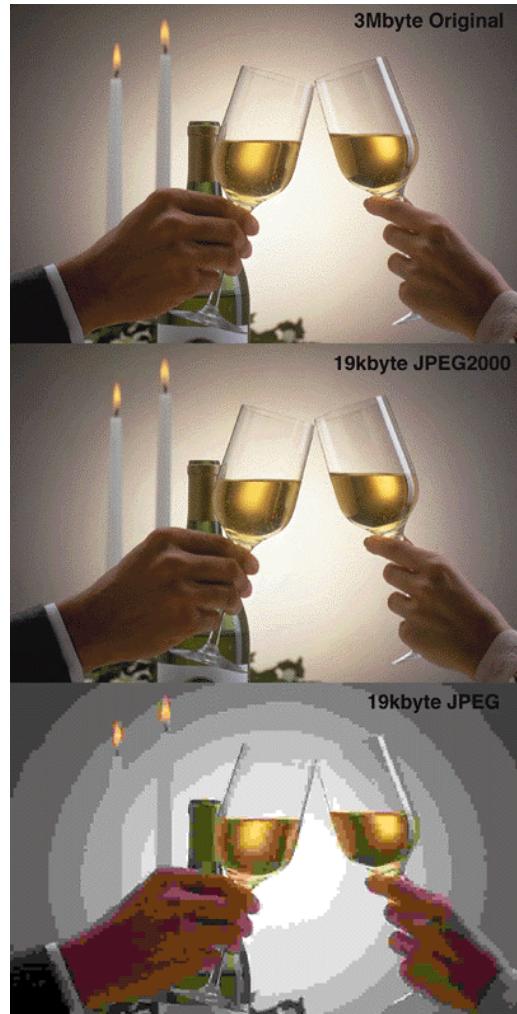


Figure 28

JPEG 2000 display ability will need to be added to both major browsers before it can be used as a replacement for the current JPEG format. This might cause a delay between when JPEG 2000 could become widely used versus when it actually *will* become widely used.

An important application for JPEG2000 is a handheld device combining both Internet applications and wireless and for digital cameras. All these devices will need new processors to handle the extra computing power required to compress images in this format. For the moment it seems that the best solution for small devices are dedicated JPEG2000 encoding processors.

Expectations that the JPEG2000 standard has a great interest for among chip vendors, software suppliers, camera manufacturers and wireless companies. But developers say the cost and complexity of implementing the compression technology in hardware is a big challenge to JPEG2000's mainstream acceptance.

The IMEC research consortium in Leuven, Belgium, is "currently developing hardware and software solutions for high-end JPEG2000-enabled products such as remote sensing.

Canon researchers are working on the architecture of JPEG2000 silicon for internal use. Introduction of the first JPEG2000 system products may be still one to two years away, however.

Hewlett-Packard already has several JPEG2000 pilot systems running in its Internet imaging lab, and plans to partner with wireless and camera companies to provide HP solutions for JPEG2000-based e-services and Internet applications.

Finally, the German company LuraTech, in Berlin, is aggressively promoting its wavelet-based software technology- much of it ready to comply with JPEG2000-on the software application developer market. The idea is to quickly make JPEG2000 plug-ins and viewers available as a standard feature of Web browsers and popular software programs such as Adobe's Photoshop. For the moment they have their own format (see Appendix C)

With the JPEG2000 format, depending on bandwidth or time limits, users can request a thumbnail, full-monitor resolution or print-quality resolution. Or a user can simply ask to access the image only at the appropriate resolution for his or her display. He or she can even request the delivery of only a particular region of an image, without transmitting the whole high-resolution picture.

A second part of the draft JPEG2000, which adds special-purpose extensions, will follow in 2001. JPEG2000 part one will be the royalty-free version, but part two can include various types of third-party extensions that may or may not involve royalties.

JPEG2000 is expected to become the accepted, universal file format for digital images, whether on the Web or from digital cameras, printers, faxes, remote sensors or wireless transmissions.

4.1.4 LURA WAVE

LuraWave is file format that uses wavelet-based compression to reduce file sizes while achieving better image quality than conventional methods such as JPEG and TIFF. Compared to the current standard lossy compression format, JPEG, the LuraWave product family delivers higher image quality at the same compression rates while reducing the amount of data required to represent and store an image. Images in the LuraWave format (.LWF) can be viewed over the Internet using the free LuraWave browser plug-ins.

LuraWave products are offered in configurations for individuals, networked workgroups, the Internet, etc. for all platforms (Unix/Linux, Windows, Mac).

All digital images that are currently stored in standard formats, such as TIFF, BMP, PPM, JPEG, etc. can easily be converted into the LuraWave image format by employing the various LuraWave products available online as Plug-ins (for Adobe Photoshop, Jasc Paintshop Pro, Netscape and Microsoft Browsers, Cumulus, Macromedia Director), stand-alone solutions (LuraWave SmartCompress) and software developer kits (C-SDK, OCX-SDK).

LuraTech is currently developing *LuraWave 3.0*, a new, improved image compression procedure. LuraWave 3.0 will improve on the extreme efficient LuraWave 2.0 procedure and provide compatibility with the forthcoming JPEG2000 international image compression standard.

LuraWave 2.0 already contains many of the features expected in the new JPEG2000 standard. LuraTech's wide experience in the field of wavelet-based image compression places the company in an ideal position to provide a leading implementation of the JPEG2000 standard.

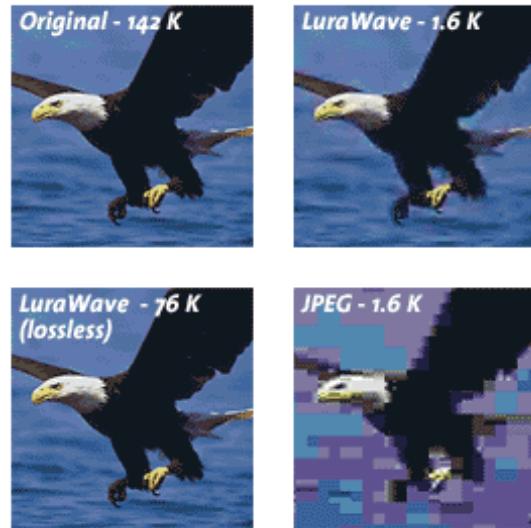


Figure 29

4.2 NEW VECTOR IMAGE FORMATS

4.2.1. SCALABLE VECTOR GRAPHICS (SVG)

SVG stands for Scalable Vector Graphics. SVG promises to bring the quality of vector graphics to the Web, and to do it in a form that is supported by all major graphics vendors and all major browsers.

The SVG initiative is part of the Graphics Activity arm of the World Wide Web Consortium (W3C). The most recent working draft of SVG included participation from companies such as Adobe, Apple, Corel, Hewlett-Packard, IBM, Macromedia, Microsoft, Netscape, Quark and Sun.

Given the level of widespread support for SVG, it may be fully implemented in browsers by the first quarter of 2001.

SVG is composed of a rich feature set, including support for traditional vector elements such as rectangles, rounded rectangles, circles, ellipses, lines, polylines/open paths, and polygons/closed paths. Support will also be included for text and for raster images such as JPGs and PNGs. As its name implies, SVG images will also be scalable, meaning that they can be rendered for display on a 96 dpi monitor or a 300 dpi printer, without any loss of quality. The scalable feature of SVG is also expected to support image panning, and the ability to zoom in and out of an area of an image, again with no loss of quality.

SVG will be capable of producing bitmap type effects such as lighting, shadows, blurs, noise and alpha-channel antialiasing on vector images. SVG will accomplish this by defining a preset group of bitmap filters that won't actually be applied to the image until the image reaches its destination (such as your viewer's Web browser). By applying bitmap effects in this way, the scalability of the vector image is maintained, along with the esthetics of bitmapping. If the viewer needs to output an SVG image rendered at 96 dpi to a printer, the original vector image can be sent to the printer, and the bitmap effects will then be added, this time calculated for 300 dpi instead of 96 dpi. The result is perfect image reproduction at both monitor and printer resolutions.

SVG allows interactivity and animation. SVG will be in full compliance with the World Wide Web Consortium's Document Object Model.

SVG is likely to avoid a problem of having the same file displayed differently on different browsers. This problem is caused in large part because the HTML Object Model was implemented by browser vendors before work on a standard had been completed. SVG will use the XML Object Model, which has been widely and consistently implemented since completion of the standard by the World Wide Web Consortium. As a result, much more consistent implementation is expected.

This feature is what will make SVG such a powerful animation format. Every aspect of an SVG file, including location, width, height, outline color, fill color and gradient type, is fully available in text format, and thus will be fully controllable.

4.3 IMAGING TOOLS

See Appendix C.

5. MULTIMEDIA PUBLISHING

5.1 DIGITAL MUSIC

Technologies enabling the simple and fast exchange of music have existed for decades. In the physical world, audio cassettes led the charge. Minidiscs (MD) and recordable compact discs (CD-R) made it easy to swap and pirate music with near-original quality. Movies, images, and text have had their equivalent of the audio cassette: video cassettes, video compact discs, photocopiers, scanners, email, optical character recognition, etc.

When the Internet and digital music compression format started to be used, music was mainly transmitted through Email, FTP, Usenet, IRC, ICQ and music WebPages. Finally the so called peer-to-peer systems, pioneered by Napster appeared.

The differences between traditional duplication methods and the Internet trading of digital files are the convenience and the speed and scale at which it can be done. Anyone, particularly university students and people with high speed internet connections, can easily download almost any music they want in minutes.

The MP3 and other compressed music formats allow to compress music to a size that makes it reasonable to store with the current storage capacity and more importantly makes it possible to transmit them across the Internet using the current bandwidth available for most users. For typical home Internet users (56 kbps modem), it takes approximately eleven minutes to download a high-quality MP3-encoded song³⁸. For a typical university student with access to the high-speed university network it takes about thirty seconds.

There are important differences between the US and Europe with regard to the downloading of digital files at home. In the US, most Internet Service Providers (ISPs) charge a flat rate, as does the telephone company for local calls. That makes the incremental download cost approximately zero. It is mainly a matter of patience and hard disk capacity. In Europe flat rate telephone access is relatively new and limited. However free ISPs are more common and provide a better service in Europe than in the US. Ultimately the flat rate telephone plans still provide an advantage to US users as they can leave their music downloads over night if necessary, and therefore getting music from the Internet becomes a matter of time, not money. New flat rate plans, DSL and cable connections should reduce the difference in the near future. Some times there is an initial installation charge for DSL or cable modem and then a monthly subscription fee between \$US50 to \$US100 in the US³⁹ and about €40 to €80 in Europe⁴⁰ depending on the speed of the service and the provider. Using that line, thousands of dollars worth of music can be freely downloaded.

The convenience of MP3 is undeniable. Today a 20 gigabyte hard disk costs less than US\$100. This can store approximately 5000 near-CD-quality songs, that is about 20000 minutes (fourteen days) of

³⁸ Popularly encoded songs (128 kbps at 44 KHz) are approximately one megabyte per minute.

³⁹ Prices from Verizon (Former Bell Atlantic and GTE) InfoSpeed.

⁴⁰ This considers prices from Terra Networks (Spain) and British Telecom (UK).

continuous listening. Portable MP3 players are about the size of audio cassettes. Compare that to compact discs, which store about seventy-four minutes of audio and do not fit into standard pockets. Using an MP3 or similar compression format it is easier to carry a large amount of music on a laptop hard drive or in an MP3 CD. MP3 provides a compression ration of 10 to 12. That means that 11 or more (not all music CDs are completely full) CDs can be compressed into a single recordable CD. Furthermore portable MP3 players are smaller than any other device to carry music. Typically these players have from 32 to 64 Megabytes of solid state memory allowing about an hour of CD quality music (see

Figure 30



Diamond Rio PMP600
Mp3/WMA solid-state
music player.

Figure 30). New portable players (see Figure 31) have their own portable hard drives with up to 6Gb of data storage capacity, which provide more than 100 hours of music.

There are several efforts to sell music legally in MP3 or several other formats. Formats such as Microsoft's WMA, MJUICE MP3, Liquid Audio and a few others incorporate copyright protecting features.

However protection schemes seldom work. Encryption for Microsoft WMA format files was broken almost immediately after its release. The process is simple: somebody would purchase the right to listen to the encrypted audio file, play it back through special software which records the decrypted audio file and he or she would have on the hard disk a permanently decrypted audio file.

Figure 31

(a)



(b)



- (a) The PJB 100 Designed by Compaq TM can stores 4.8 to 6 Gigabytes of Digital music.
- (b) The Creative Labs Nomad JukeBox stores 6 Gigabytes of music.

a video camera; these movies are then digitalized and sold on video CDs or distributed through ftp sites or IRC. Telesyncing of movies produces a significant loss of quality. However, other methods of capturing analog output before it leaves the computer are far better, for example the digital to analog to digital again conversion inside a computer sound card produces an almost negligible loss of quality. This is what end to end encryption would try to prevent. However to incorporate end-to-end encryption special more expressive monitors and speakers would be required and the system would still be vulnerable to “telesyncing” copy methods.

But much more important than the possibility of overriding copy protection there is the fact that virtually all the same multimedia content that can be purchased in protected formats, if not more, is available completely for free in services such as Napster and Gnutella, which has no central server that can be shut down (see Figure 32). Therefore a given Internet user has the choice of either getting a free song in a free format or pay for a similar file with a restricting copy protecting feature. As copying a digital file is hardly perceived as stealing in the same way as it would be stealing a CD from a shop, most users will take the free alternative. Furthermore Gnutella, Freenet, Napster, CuteMX, Scour Exchange and all the other file trading applications available provide a better user experience than any alternative currently provided by the recording industry. For example at the time of writing this report Amazon, CDnow and most of the most popular online retailers do not sell mp3 music (others such as Musicmaker.com already do although with a limited selection). This will change in a few months, if not weeks, but the downloadable music from these sites will still not match the Napster or Gnutella model: “*click an icon in your computer and access almost all the music you ever wanted... for free*”.

The recording industry has several options. It is not possible to prosecute every Napster or Gnutella user. It could be attempted to force all the manufacturers of MP3 portable devices to accept only SDMI protected music. But even if this was possible there are already free plans for mp3 players available in the Internet. Furthermore most of the MP3 encoded music is played on computers, and as long as music is released in the current music CD format it can be easily copied and compressed into MP3.

However some people believe that the use of “brute force” against the new distribution models will not succeed for all the reasons mentioned above and that the music industry should adapt and try to take advantages of the new distribution possibilities offered by the Internet and peer-to-peer trading systems.

One thing consumers demand above all others is product quality. Currently, Internet music varies greatly in quality. Some are extracted from music CDs which have scratches, others are encoded at low bit rates, diminishing the audio quality to unbearable levels. Yet others are encoded using low-quality encoding software, leading to diminished quality. The RIAA and its constituent record companies have an

Figure 32



Screen capture from Gnutella clone *Gnotella*.

opportunity to exploit their own brand name and quality control procedures to produce digital downloadable music with consistent high quality. May be users may be willing to pay a small amount to ensure they are not wasting their download resources.⁴¹

Another proposed method to limit the illegal distribution of music over the Internet consists in “*turning pirates into paying distributors*”. This was exposed by Gene Kam, Gnutella developer and founder of InfraSearch Inc. in the US Senate Judiciary Committee hearing about the intellectual property of music on the Internet on July 2000.

Exploiting new technologies often requires retooling. Automobile manufacturers, computer chip manufacturers, the United States Postal Service. All have had to continuously change their methods of doing business in order to remain competitive.

The idea is that a relatively simple infrastructure could be created that would allow users to share their files and be compensated for every trade and if the file is, for example a music file property of a record company, that company, would be also compensated or the resale of their intellectual property. The idea is that because users have a finite-capacity Internet connection they would seldom, if ever, allow someone to download from them for free when they could charge for the privilege. If people charged, for example, 1.50 USD for each track downloaded from their computer, and they split the revenues evenly with the record companies, the record companies would make 0.75 USD for each download. If that system was to acquire a similar number of users as Napster it could generate a very large profit for record companies, especially considering that they would not have the additional cost of manufacturing and distributing the CD.

Of course there are several problems to this system. First, it is not clear if this would be ultimately successful and if users would eventually revert to completely free systems. Furthermore recognizing a given track would involve either a different file format than MP3 that provided the added information about the artist, which would indubitably be a great inconvenient; or a more complex system that would allow the program to analyze and identify the music and its owner.

Mp3.com provides another interesting model to distribute music, which could be applied for video and electronic texts. Although MP3.com has many artists there is a fundamental problem with it model. A traditional recording company acts as a big marketing machine for their artists. Traditional record companies can afford to produce expensive music videos and promote their artists in radio stations.

Lars Ulrich from Metallica summarized the role of recording companies very simply:

*“Because what really, essentially, is a record company? A record company is really essentially a bank, a bank that funds a bunch of money to make records, and videos and promotion, publicity appearances and so on, and they take that shot that one day the artist is going to be so successful that they're going to first of all get all their money back, second of all make a profit.”*⁴²

⁴¹ From Gene Kan's statement for the Senate Judiciary Committee Forum: “ Utilitarian view of Intellectual property and music on the Internet.”

⁴² From Gene Kan's statement for the Senate Judiciary Committee Forum: “ Utilitarian view of Intellectual property and music on the Internet.”

Recording companies are in fact exactly like venture capitalists. They fund, promote, and advise their portfolio artists.

To discover an artist in MP3.com it is necessary to download a file (or listen to the Real Audio streaming version) from an initially unknown artist. It can take a long time before a user can find an artist that he or she likes, and this has to be done by trial and error. If Mp3.com was to do the same thing promoting their artists as traditional recording labels they would have to select a small group of them and leave the rest with no promotion. All the spending of promotion should be recovered by selling the music, instead of giving it away. Basically that will mean that MP3.com would become a traditional recording label, with the only difference that their music would be sold in the Internet as digital files rather than on CD in traditional shops, and traditional record label are also moving in that direction. For example Warner Music, which will become the world's largest music company if it completes its merger with EMI, said on September 2000 that they will start selling music on the web in November 2000, beginning with singles and internet-only tracks from artists like Barenaked Ladies and Paul Simon. The company, a division of Time Warner, plans to expand the digital music offerings to more than 1,000 digitised albums and singles shortly after the November launch. The music will be available for download from online retailers, including Walmart.com and Amazon.com. The programme will launch initially in the US and Canada, but Warner expects to expand it internationally later. Other large labels have also made part of their catalogues available for digital download, including Sony Music, EMI, Seagram's Universal and BMG Entertainment. It is only a matter of time before all the content is available in some kind of compressed form

Both artists and recording companies can potentially benefit from digital music if an adequate system is adopted. For artists, digital music provides a chance to reach directly into a global audience. They have the opportunity to capture nearly 100% of the gross sales of their product. For recording companies, there is an opportunity to reduce the marginal cost of distribution to nearly zero, and to expand the scope of distribution to the entire Internet.

5.2 VIDEO

With the impending globalization of inexpensive broadband access, the problems and opportunities that already exist for the music recording industry will extent to the video and movie industry.

Unlike MP3s, which provided the framework for CD piracy, historically there have been several barriers to user-to-user electronic movie piracy.

Before DVDs, to capture a VHS video into digital format it was required to have special video capture hardware, also at the time before DVD was widely available computers did not have the necessary power nor internet connections had the necessary bandwidth. DVDs have always been encrypted using the Content Scrambling System (CSS). CSS prevented anyone from accessing the contents of a DVD except for DVD players that had the key built in. But earlier in the year 2000, two European hackers discovered that Xing inadvertently exposed the CSS encryption key in their latest DVD player. Within weeks the anonymous team released a small software program named DeCSS, which allowed anyone to access the contents of a DVD. Despite numerous lawsuits against sites that posted the program, DeCSS could not be stopped. Its code has been posted everywhere and there are even T-shirts and songs that include the code.

However, the large size of DVD files still makes it unpractical to transfer them across the Internet with current Internet connection speeds, even with DSL or cable modems.

New formats such as MPEG 4 can compress movies to fit in a single 650-700 Mb CD-ROM. Video CD technology can compress a movie to fit 2 CDs with relatively acceptable quality for small screens and has been available for several years. Although the quality of these movies is not as high as the new DVD movies 650 megabytes can be downloaded in about 40 minutes with a fast Internet connection.

A new technology, named *DivX*⁴³ has the potential to do for video what MP3 did for audio. DivX format is in no way related to the failed Divx movie format, which was marketed as an alternative to DVD. The new DivX is actually a combination of a hacked version of the MPEG-4 codec for video, and MP3 for audio. MPEG or Motion Picture Experts Group file formats, are standard compression formats that use mathematical algorithms to dramatically reduce video and music file sizes. Perhaps the most well known format, MP3 (MPEG, audio layer 3), employs algorithms designed to reduce music file sizes without a significant loss of quality. Generally compression formats must balance video quality and file size. Increasing the former increases the latter, and vice-versa.

The DivX MPEG-4 codec can compress a full-length DVD movie that is about 2-3 gigabytes down to about 650mb to 700mb (small enough to burn onto a standard recordable CD).

However movies are harder to rip from DVD and encode into Dix format than MP3 files and the process is very time consuming, it can take more than 10 hours even for a fast computer (650Mhz). Sometimes Dix movies do not properly prepared or have excessive compression and lack quality. Better

⁴³ The “;-)” smiley face in the name is used to differentiate and to mock the original Divx movie format

software and faster computers will soon make this process of creating MPGE4 movies from DVDs faster and easier.

Playing DivX movies is easy. The codec if free to download. Once installed, PC, MAC, Linux and BeOS users can watch DivX encoded .avi and .asf movies in Microsoft Windows Media Player or other video players.

There are currently more high compression systems being developed. One them is the MPEG4-based codec called 3ivX is being developed by some of the same people that created DivX ;-). The goal is to optimize compression so the file size is smaller, higher quality and fully streamable. Above that the codec will be supported on all platforms (Windows, Macintosh, BeOS, Linux) and the player will be open source. Further developments of the MPEG compression are also being developed (MPEG-7 and MPEG-21) although some of them are mainly concerned with other issues.

The movie industry seems to be reacting faster than the music industry, as they have the advantage of been able to learn from the mistakes of the music industry, which due to the relatively modest space requirements of music compared with video, are destined to precede the movie industry in the distribution of multimedia content over the internet for some time. As a recent example of the movements of the video industry, Enron Broadband Services and Blockbuster have signed a 20-year exclusive agreement to deliver movies on-demand over DSL lines. The companies say that the agreement will result in a service that delivers a variety of secure, on-demand entertainment. Enron and Blockbuster plan to introduce the entertainment on-demand service in several U.S. cities by the end of the year 2000. Beginning in 2001, Blockbuster expects to extend the service to other domestic and international markets.

APPENDIX A - TOOLS FOR TEXT BASED DOCUMENT CREATION.

A.1. SUN MICROSYSTEMS - STAR OFFICE / STARPORTAL

Sun's StarOffice suite was originally built by Star Division Corp. of Germany as a group of object-oriented programs for word processing, spreadsheets and presentation graphics. Sun Microsystems acquired Star Division in August 1999 and began offering StarOffice as a free product. Despite being free, Star Office includes much of the functionality of Microsoft Office and/or WordPerfect.



The latest version at the time of writing this report- version 5.2 - supports files formatted with XML and features greater support for file formats used in the market-dominant Microsoft Office.

StarOffice is currently available for Windows, Linux, Unix (Sun Solaris) and OS/2 (only StarOffice version 5.1 at the time of writing this report).

Sun recently released StarOffice's source code and created a website - OpenOffice.org – to centralize the development efforts. The code was released under the GNU General Public License and the Sun Industry Standards Source License (SISSL). This is a similar move to the one Netscape adopted some time ago. Sun expects to improve the product with the voluntary help of developers all over the World. StarOffice was already shipped as a free product, even for non-personal use, making the source code available does not have a big effect on Sun's revenues.

StarPortal

Sun Microsystems will also release a second version of StarOffice, called StarPortal. This version is designed to run on servers, instead of individual desktops and to be accessible with a Web browser.

This will allow users to edit and access their document from any computer connected to the Internet without the need to carry their own computers, software and/or files with them.



Sun Server family.



Sun Ray enterprise

The portal includes a group of device templates, which can be thought of as a set of directions on how to present StarOffice documents, or parts of them, to specific devices such as a PDAs or cell phone screens that use the Wireless Application Protocol (WAP). Using the templates, the server can read an incoming browser request, select the relevant documents or data, and format them to match the client's screen.

According to Sun, in the future, StarPortal will import and export documents formatted in XML. One benefit of XML is that it will let the portal move parts of a document, instead of the entire file, back and forth between client devices and servers.

Sun's revenue strategy with StarPortal is twofold. First, companies that want to offer it as a service will want to pay for around-the-clock technical support that Sun will provide. Second, though the server component of StarPortal will run on Windows and Linux as well as Sun's Solaris. Therefore Sun will increase its revenues by selling more servers and by supporting Solaris and StarPortal software. StarPortal fits perfectly into Sun's strategy of network computers, connected to their servers.

According to Sun Microsystems the portal version is designed for businesses and application service providers. Whether any organization will adopt this system remains to be seen.

Summary of Star Office's features:

Latest version August 2000: version 5.2, version 5.1 for OS/2

- Free for personal and non-personal use.
- Integration between all StarOffice applications.
- Java™ and JavaScript™ technologies, OLE, DDE.

Database application supports ODBC, SQL, Active Data Objects (ADO).

Email client and browser supports POP3, IMAP4, VIM, NNTP, SMTP, LDAP, HTTP, FTP

HTML and RTF XML and iCAL import/export calendar support.

Import /Exports Microsoft Office compatible documents.

Runs on Solaris Operating Environment on SPARC and Intel Architecture platforms, Microsoft Windows 2000, NT, 98, 95 and Linux.

Available in Danish, Dutch, English, French, German, Italian, Polish, Portuguese, Russian, Spanish and Swedish.

A.2. MICROSOFT OFFICE, MICROSOFT WORD AND COLLABORATIVE TOOLS

Microsoft's Office Suite bundles several applications, which always include Word, for text processing, Excel for Spreadsheet creation, PowerPoint for presentations and Outlook an email client and organizer program. Depending on the version it may also include Access, Front Page and Photodraw. (See Table 1)



This review concentrates mainly on Word, as it is the tool most closely related with the subject of electronic publishing. However many of the benefits of Microsoft Office come from the integration between its applications, such as Outlook and FrontPage.

Standard	Small Business	Professional	Premium
<ul style="list-style-type: none">• Word• Excel• Outlook• PowerPoint	<ul style="list-style-type: none">• Word• Excel• Outlook• Publisher• Small Business Tools	<ul style="list-style-type: none">• Word• Excel• Outlook• Publisher• Small Business Tools• Access• PowerPoint	<ul style="list-style-type: none">• Word• Excel• Outlook• Publisher• Small Business Tools• Access• PowerPoint• FrontPage• PhotoDraw

Microsoft Word gained a lot of popularity from its early days by being one of the

first text processors running on Mac OS and Windows and offering a true WYSIWYG approach to document creation. Microsoft Word and MS Office suite spread quickly stealing market share from its major rival, WordPerfect, according to USA Today¹ just from 1995 to 1998 it grew from 43% to 80%, whereas the percentage using rival product WordPerfect dropped from 61% to 21%. This has been partly due to its ease of use but also to Microsoft efficient marketing strategy. Furthermore, the almost lack of protection made it easy for people to create illegal copies of it and increase the customer base very quickly. Otherwise, MS Office is an expensive piece of software, which many home users would not have been able to afford.

Table 1

MS Office's latest version (at the time of writing this report) for Windows is 2000 and 98 for the Macintosh OS, with a new Macintosh version expected for 2001. Version 2000 actually corresponds to Word version 9, however new releases are now numbered by the year of release according to Microsoft's practice started with Windows 95. At the time of writing this report Microsoft has already released the first non-public beta of the next release of Office.

Word 2000 includes a series of tools to allow collaboration on a document by various people. For example it can track changes and corrections in the text that can be accepted or rejected by other users. Notes and text and voice comments can also be added to the document.

¹ Source: USA Today 22 Jan 98

Several Web-based collaboration services have been incorporated in the latest release.

Web Discussions, let users post and attach comments within any Office document, or inside frames at the bottom of the screen. And Web Subscriptions allow users to easily track collaborative efforts.

They can subscribe to be notified of changes made to any Office 2000 or HTML document on a Web server. The conditions and frequency of updates can be determined by the user. These updates are automatically sent by e-mail.

These services are based on Web Server Extensions built into FrontPage 2000, the Web page management application within Office 2000. The new service allows users who do not have a Web server onsite collaborate on Office documents over via an ISP that supports the new program.

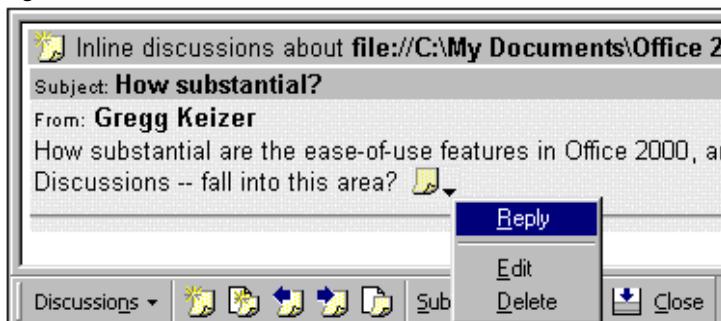
Support for the FrontPage Server Extensions allows users to publish documents directly to Web sites hosted by servers that support these extensions, as well as add the ability to collaborate within these documents. Furthermore Office 2000 makes extensive use of both HTML and XML as a file and data sharing format, intended to make the software more Web-friendly.

Microsoft is working on a similar server-based service to StarPortal for Office. And a California start-up called ThinkFree.com is already offering a Web-based office suite that is written in Java and mimics the features of the Microsoft software.

The web base office will allow customers to use Microsoft's Word, Excel, PowerPoint, and other Office applications via the Internet, instead of having the Microsoft software installed on their PCs. This all integrates with the upcoming .NET platform and strategy. Although review the whole .NET concept goes beyond the objective of this report a brief description can be found below.

Microsoft has not officially announced pricing for the hosted applications. Whether competition will force Microsoft to make the service available for free still remains to be seen.

Figure 1



MS Office 2000's Web Discussions tie message threads to specific documents, even specific parts of documents.

Summary of features:

Latest version August 2000: version 2000 (9), version 98 for Macintosh

- De facto industry standard for text based document creation.
- Ability to import and export to HTML documents.
- Multiple third party applications interact with Office
- Good integration between different applications bundled with Office.
- Collaborative tools and web integration
- Smart dictionary automatically recognizes the language in the document. Standard support for English, Spanish and French.
- Office 2000 *Collect and Paste* allows users to store up to 12 pieces of information from different documents on the clipboard.
- Available for Windows and Macintosh,
- Smaller version available for PocketPC/Windows CE devices. Documents can be read by other portable devices running Palm OS (with a third party application) and Epoc32.
- Upcoming on-line version.

.NET



The .NET (dot net) strategy, which Bill Gates described as the company's biggest shift in direction since the move from DOS to Windows, is expected to be rolled out in stages over the next two to three years. The new approach relies on XML -- to create a set of interactive applications stored "in the cloud" rather than holding most programs and data on individual computers and network servers. Information would then be delivered via a variety of devices, from mobile phones and PDAs to full-powered computers in an appropriate form for each device. Microsoft predicts that software will increasingly become a service, rather than a product.

From the point of view of its business strategy, the company seems to be taking some risks by moving to a subscription model where the majority of their revenues will come from incremental payments as people use the tools, while many of the development costs are paid for up front². Furthermore the .NET strategy involves every division of Microsoft and it is uncertain how this strategy could be accomplished if the company is finally forced to break up.

Microsoft's .NET strategy comprises the following:

Microsoft .NET platform—Includes .NET infrastructure and tools to build and operate a new generation of services; .NET User Experience to enable rich clients; .NET building block services, a new generation of highly distributed services; and .NET device software to enable a new breed of smart Internet devices.

Microsoft .NET products and services—Includes Windows.NET, with a core integrated set of building block services; MSN™ .NET; personal subscription services; Office.NET; Visual Studio® .NET; and bCentral™ for .NET.

Third-party .NET services—A vast range of partners and developers will have the opportunity to produce corporate and vertical services built on the .NET platform³.

More information about the .NET platform and application can be found in Microsoft's website at :

<http://www.microsoft.com/net/whitepaper.asp>

² Source: Richard Sherlund of Goldman Sachs in Wired.com

³ Source: Microsoft Corp.

A.3. COREL WORD PERFECT OFFICE



Corel bought WordPerfect from Novell in February 1996. This word processor once held an 80% Market share, however the relative neglect by Novell, Microsoft's aggressive marketing strategy and the somehow easier to use MS-Word made it lose its leading position. Corel continues to support and release new versions of WordPerfect. Word Perfect version 2000 is available for Windows and Linux. Version 2000 is, in the same way as Microsoft Office, version 9, renamed 2000 for marketing reasons. Furthermore version 8 is available for UNIX operating system and free for personal use for Linux. This makes WordPerfect a popular tool for the small Linux market in which Corel is placing a lot of interest.

WordPerfect, as Microsoft Word, usually comes as part of an integrated office suite, in this case WordPerfect Office, including applications such as Presentations, Quattro Pro for spreadsheet creation and Paradox for database creation. Word Perfect also incorporates some differentiating features like speech recognition in some editions and the ability to publish in HTML and XML.

Although the word processing market is mainly dominated by Microsoft Word and Office, Corel Word perfect still holds the niche of the legal market. Several years ago, Corel Corporation, realizing that it was losing the word processing war with Microsoft in the legal arena (especially at the larger firms), decided to create a suite targeted directly at the legal industry and delivered its WordPerfect Suite 7 - Legal Edition in 1997. A year later they came out with the WordPerfect Suite 8 - Legal Edition. The two legal editions contained the standard Corel Suite components (i.e., WordPerfect, along with: Presentations, Quattro Pro, Paradox), along with several legal applications: Amicus Attorney, a case management tool; HotDocs, a document assembly application; a table of authorities generator (LEXIS-NEXIS's FullAuthority in the first edition, West's CiteLink in the second); and enhanced document comparison.⁴

According to a 1998 survey by the American Bar Association, the most recent one available on February 2000, WordPerfect held a 60% share in both large and small law firms with Microsoft Word holding most of the rest and Lotus' office suite commanding a low single digit share in the small firm market.

As significant fact of the presence of WordPerfect in this market, the U.S. Dept. Of Justice ordered 55000 copies of Corel Law Word perfect in February 2000, almost 30% of them the special version meant for the legal profession.⁵

⁴ David Greenwald. LexTech Inc.

⁵ ClieNT Server News, G2 Computer Intelligence Inc. February 2000 , Issue 335, Article 28

Summary of features:

- Latest version August 2000: version 2000 (9), version 8 for Unix
- Web technologies from Trellix® and NetPerfect®
- Additional speech recognition software in some editions
- Increased import and export compatibility with Microsoft Office
- Same WordPerfect® file format since version 6.1 so older files don't need to be imported.
- Supports HTML, ODBC, Java™, SGML and XML Publish the same document to paper, HTML and XML.
- Imports Microsoft Office documents.
- Special version for Law professional available.
- Available for Windows, Linux and Unix operating Systems . Documents can be read by portable devices running Palm OS (with a third party application) and Epoc32.

A.4. ADOBE FRAMEMAKER + SGML

Adobe FrameMaker and Adobe FrameMaker +SGML are Adobe's tools for the creation of structured content. These products allow publication to the Web, CD-ROM, and print- while retaining the document's structure and organization.

FrameMaker+SGML tries to let the user focus on the content without having to learn the complexities of SGML or XML. Adobe Framemaker includes Quadralay WebWorks Publisher Standard Edition software for enhanced HTML and XML capabilities. It also supports creation of structured PDF files and includes extensive support for templates that lets the user focus on the content, while Adobe FrameMaker automatically formats it in a WYSIWYG environment.

XML and SGML creation

When publishing to XML, a cascading style sheet automatically can be generated automatically. XML tags can then be mapped from the FrameMaker paragraph style names, or, with FrameMaker+SGML, generated directly from the document's element tags - producing high quality XML that matches the original structure of the document.

HTML creation

Quadralay WebWorks Publisher Standard Edition also adds powerful and sophisticated HTML publishing features to Adobe FrameMaker and Adobe FrameMaker+SGML. WebWorks Publisher Standard Edition offers complete control over HTML generation and takes care of mapping FrameMaker styles to HTML styles, turning cross-references into hyperlinks, and converting graphics to GIF, JPEG, and PNG formats.

PDF creation

Windows and Macintosh versions of Adobe FrameMaker and Adobe FrameMaker+SGML include Adobe Acrobat Distiller 4.0, for publishing to Adobe Portable Document Format (PDF). When publishing to PDF, Distiller lets users specify settings such as compression, font embedding, and color profiles. Distiller 4.0 also enables the creation of structured PDF files.

When a FrameMaker or FrameMaker+SGML document is saved as a PDF file, cross-references and hypertext commands automatically become links, and bookmarks can be automatically created.

Author, title, subject, and keywords of a document can be specified before exporting it to PDF. This information appears when the PDF file is printed, and it is also accessible to some search engines for PDF files posted on the Web. FrameMaker and FrameMaker+SGML also allow placing a PDF page directly into a FrameMaker document. This enables the user to capture a legacy paper document or other information in PDF and leverage it in the FrameMaker publication.

Collaborative Document creation

For long documents, built-in cross-reference, table of contents, and index generation capabilities produces work that is easily navigable. Conditional text support enables the creation of multiple variants of a document in a single source. For efficient document review cycles, FrameMaker and FrameMaker+SGML include revision functions such as change control bars and a summary of changes.

Framemaker files can be read online by the FrameViewer without having the full program FrameMaker program. FrameViewer is available for UNIX and Windows. FrameViewer is not a free product but at the time of writing this report its price was US\$49 versus the US\$1,449 of the full FrameMaker+SGML. Furthermore as new browsers will display XML formatted files therefore viewing files created with FrameMaker should not represent a problem.

Companies such as Cisco Systems, Motorola , Hitachi, FedEx and Federal Express use Framemarker for some of their multichannel publishing needs.⁶

Summary of features:

- Latest version August 2000: version 6.0
- Available for Windows, Macintosh, IBM AIX, HP-UX Unix , Son Solaris, Beta version 5.5.6 for Linux at the time of writing this report.
- Creates PDF, SGML, XML, and HTML content
- Automatic generation of a cascading style sheet automatically. With FrameMaker+SGML, XML tags can be generated directly from the document's element tags.
- Files can be read by the FrameViewer without having the full program FrameMaker program. FrameViewer is available for UNIX and Windows. FrameViewer is not a free product.

⁶ Source: Adobe Systems Inc.

A.5. ARBORTEXT EPIC

Epic is a modular product line that helps enterprises to provide personalized, dynamic and easily searchable content for electronic publishing, e-commerce and B2B e-marketplaces. Epic provides an XML-based solution that enables to create, manage, personalize and share content.

Using XML, the several modules of the Epic product line can tailor the content to make it available in different media, including the Web, print, CD-ROM and wireless devices. It can capture existing content from word processing and desktop publishing software and also offers search functions taking advantages of the XML format. The Epic product line has many components, which are summarized in Table 2.

Table 2

Design Tools	Creation Tools	Management Tools	Transformation Tools
<u>Client</u> Epic Architect ACL Designer	<u>Client</u> Epic Editor / Epic Editor LE Epic Interchange	<u>Client</u> Adapter for Oracle IFS Adapter for Documentum	<u>Client</u> Web/Wireless Composer CD-ROM composer Print Composer <u>Server</u> Epic E-content Engine (E3)
<u>Server</u> Extent	<u>Server</u> Extent		

7

Epic's Design Tools: Epic Architect, ACL Designer .

Epic Architect offers both rapid prototyping capabilities and powerful application development functions. Epic Architect's key functions include:

- Graphical data modeling/DTD design – through a graphical user interface, users can develop and revise DTDs (Document Type Definitions) to guide authors to create valid XML content. Also a prototype of a DTD can be made by creating sample content in Epic Editor and importing that content into Epic Architect, which will automatically create a DTD based on the content.

Stylesheet development – with Epic's "Turbo Styler" capability, multiple stylesheets can be configured, for producing content on the Web, wireless, print and CD-ROM.

Tag alias design –. Organizations that create information in multiple languages can use tag aliases to help their authors create content in their native language regardless of the language of the DTD.

⁷ Source: Arbortext

Conversion configuration – to convert information from content formats such as Microsoft Word, Epic Architect provides the capability to associate style names with tag names. Because conversion rules can become complex, Arbortext makes it easier to set up conversions with a easy graphical user interface.

Creation tools: Epic Editor / Epic Editor LE, Epic Interchange

Epic offers several ways to create or capture XML content. Authors can create content directly in XML, or they can continue to use their existing word processing and desktop publishing tools and convert to XML as needed.

Epic Editor / Epic Editor LE – For directly creating and revising XML content. Epic Editor's key capabilities include an intuitive user interface and an easy and effective way to create complex, compound documents. These products run client-side or, through the Extend option, can run server-side with browser-based access

Epic Interchange –Epic Interchange converts between XML and popular word processing and desktop publishing formats. Epic Interchange converts Microsoft Word documents to XML and converts XML to Word. In September 2000, Epic Interchange will be upgraded to convert FrameMaker files and import Interleaf documents.

Epic Onword – In November, Arbortext will start shipping Epic Onword, an add-on to Microsoft Word that lets authors using Word perform XML-to-Word and Word-to-XML conversions entirely within the Word user interface.

Forms Engine – To create content using Web-based forms. For content that is appropriate to a forms interface, Epic's Forms Engine can translate these forms to XML.

Management tools

The system that holds content is called a "repository." Arbortext works with many different types and vendors of content repositories, from file systems to content management systems to databases.

Content management systems and databases introduce several benefits over standard file systems, including checkin/checkout, access control and version control. Eventually, all valuable business content will be stored in repositories that provide such controls.

The Epic E-Content Engine includes support for both Documentum and Oracle. Arbortext offers Repository Adapters to Documentum and Oracle as optional add-ons to Epic Editor and Epic Editor LE.

Transformation tools: Epic E-Content Engine (E3), Composers.

Epic provides powerful capabilities for transforming a single source of content for multiple audiences and multiple devices.

Epic E-Content Engine (E3) –is a server-based product that accepts many different types of content, stores that content in a repository, extracts that content according to each individual's needs, and

transforms it into Web, print, wireless and other forms.

Its main capabilities include:

Conversion between XML and legacy content formats (Microsoft Word today, FrameMaker and Interleaf in September; this is an optional add-on to E3)

Conversion from HTML forms to XML (this is an optional add-on to E3)

Personalization of content for specific groups or individuals

Automatic transformation from a single source of content to multiple media: Web, print, PDF, eBooks and wireless devices.

Integration with content repositories and Web platforms.

Composers – three composers are available as add-on options to Epic Editor: Print Composer, Web/Wireless Composer and CD-ROM Composer.

Print Composer – transforms XML to produce printed documents or PDF files (if Acrobat Distiller is installed). It can use multiple stylesheets to apply different appearances to the same content. Print Composer provides typical desktop publishing capabilities such as multiple columns, headers and footers, and hyphenation, cross-references and automatically generated tables of contents and indexes.

Web/Wireless Composer – transforms XML content for the Web, for wireless devices that use the Wireless Markup Language (WML), for eBooks that use the Open Electronic Book (OEB) language, and for Palm Pilot PDAs. The Web/Wireless Composer automatically generates navigation aids including a table of contents and index, and it also separates large documents into browser-sized chunks.

CD-ROM Composer – publishes XML onto CD-ROMs and provides updated information to CD-ROM users over the web.

Epic is also available with user interfaces in many languages, including: English, French, German, Spanish, Japanese, Chinese and Korean. Epic can be run in a client-server configuration or a browser-server configuration through the optional "Extend" product. Almost all of Epic runs on the following platforms: Windows (9x, NT 4, 2000), Unix (CompaqTru64 Unix, Hewlett-Packard HP/UX, IBM AIX, Sun Solaris).⁸

⁸

Epic Interchange runs only on Windows; availability on Unix is planned but not scheduled
Print Composer runs on Windows NT 4.0, Windows 2000 and Unix; it does not run on Win95 or 98
CD-ROM Composer runs only on Windows

Epic Onward, when available in November 2000, will run only on Windows
E3, which runs under a web server, initially runs on Windows NT under Microsoft IIS and on Sun Solaris under Apache; additional platform/web server combinations are planned but not scheduled.

A.6 ADOBE ACROBAT / ACROBAT READER, ADOBE WEB BUY AND MERCHANT

Adobe Acrobat is the tool to create and modify PDF files by Adobe Systems Inc. It is the only tool that can make complete use of all the possibilities of the PDF format.

Adobe Acrobat Reader can be downloaded for free from Adobe's web site. Acrobat Reader can only be used to view PDF files and to fill PDF forms, but cannot make modifications to a file. Adobe recently began shipping Acrobat Business Tools, a less expensive version of the product (sold only in volume), which strips out most of the graphics tools many users do not need.

From version 4.0, Acrobat includes several features that facilitate late-stage touch-up of Adobe PDF files. Minor text changes can be made using the Touch-Up Text tool.

Adobe Acrobat also includes Adobe Distiller and Adobe PDF writer; these tools allow almost any application that can print to create PDF files.

PDF Writer

Many applications can produce PDF files directly using the PDF writer as a printer driver. This is available for Microsoft Windows and Apple Macintosh OS. A printer driver normally converts operating system graphics and text commands (QuickDraw for the Macintosh and GDI for Windows) into commands understood by the printer. The driver embeds these commands in a stream of commands sent to a printer, the PDF Writer converts then to PDF operators and embeds them in a PDF file as shown in Figure 2.

Distiller

Some applications produce PostScript page descriptions directly because of limitations in the QuickDraw or GDI imaging models or because they run on DOS or UNIX computers, where there is no system-level printer driver. For these applications, PostScript page descriptions can be converted into PDF files using the Acrobat Distiller application as show in Figure 4. The Distiller application accepts any PostScript page description and in some cases produces more efficient PDF files and maintains the formatting more accurately. From any windows program Distiller can be chosen as a printer, when printing to distiller it will first create a PostScript file and then convert it to a PDF file.

PDF Writer

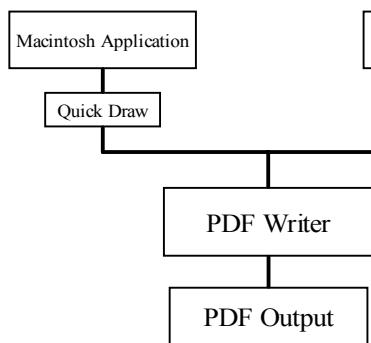


Figure 3

Distiller

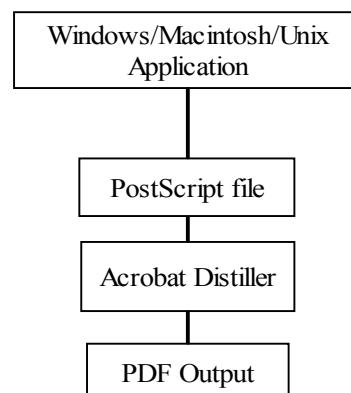


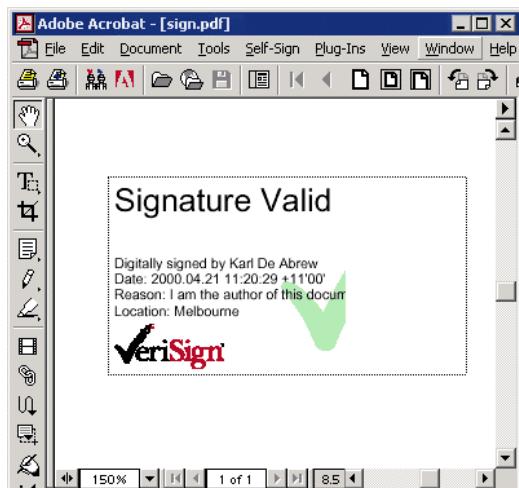
Figure 4

Once the file is created only minor modifications in the text are allowed from Acrobat. Thumbnail views, hypertext links and bookmarks sounds and videos can be added later.

Security Features

The latest version of Adobe Acrobat (4.05 at the time of writing this report, August 2000) comes with a Digital Signatures feature. These allow the user to include an electronic sign-off into the electronic document. The signing party is authenticated via a password. This is also a great way of securing your document against later changes.

Figure 5



(Screen capture from Adobe Acrobat showing a VeriSign certificate)

To use a digital signature a personal certificate must be purchased, from VeriSign for example (see Figure 5).

From version 4, Acrobat provides a default plug-in for supporting digital signatures. This Self-Sign plug-in uses a "self-trust" digital signature model. So while it can be used to prove that a document has not been altered, it cannot guarantee the identity of the author.

Luckily, solutions from several of Adobe's digital signature partners are available for Acrobat and even come in the distribution CD-ROM. It includes offerings from companies such as Baltimore Technologies, Communication Intelligence Corporation (CIC), Coastek, Entrust Technologies, PenOp, Silanis Technology and VeriSign.

Signing the document is as easy as to drag the digital signature from the tool bar.

Forms

Adobe Acrobat 4 can create interactive forms that anyone with Acrobat or the free Acrobat Reader can fill out online. Adobe Acrobat also includes built-in JavaScript™ capabilities, forms have fields that automatically verify, reformat, and calculate data as they are filled in.

Acrobat forms can be viewed with Netscape Navigator and Microsoft Internet Explorer. They also work with standard Web server CGI scripts, so HTML forms can be replaced with Acrobat forms with little or no change to existing scripts.

Acrobat forms can also act as an interface to database systems, enabling the creation of one-to-one marketing materials such as direct mail or database queries tailored for individual users. With dynamic page templates, forms can automatically expand to accommodate varying amounts of input, and then prompt delivery of customized catalogs, brochures, and other publications generated on the fly directly.

Web Buy / Adobe Merchant

PDF Merchant is a server-based technology that can be integrated into eCommerce and transaction systems. It encrypts Adobe PDF files and works with Web Buy (see later) to manage the distribution of and access to the secure PDF files. It makes it possible for publishers, distributors, and retailers to encrypt volumes of PDF files and sell them over the Web. PDF Merchant also lets users securely exchange any type of electronic document, not just for commercial purposes.

Content owners using PDF Merchant can also specify standard permissions for PDF files, including privileges for printing, changing the document, selecting text and graphics, and adding or changing annotations and form fields. For example, a textbook publisher may allow annotations, but disallow changing or copying the file.

With Web Buy, users can purchase and view digital content that has been produced by publishers using Adobe PDF Merchant. Web Buy lets users download encrypted files from the Web and unlock them to read on their personal computer or electronic reading device. Adobe Web Buy is a fully integrated component of the latest versions of Adobe Acrobat and Acrobat Reader (versions 4.05, English only, at the time of writing this report).⁹ Glassbook Reader, another program to display PDF files intended for electronic books, also incorporates the capability of buying e-books directly from the program.

Web Capture

Adobe Acrobat's Web Capture tool can convert a Web page or an entire Web site to a single Adobe Portable Document Format (PDF) file maintaining links and graphics for convenient offline viewing, printing, and storage.

Summary of the different versions and functionality of the Adobe Acrobat product line.

	Acrobat Reader	Acrobat Business Tools ¹⁰	Acrobat 4
View and Print PDF files	•	•	•
Fill and submit PDF forms	•	•	•
Download and Read Electronic books	•	•	•
Mark up documents electronically		•	•
Create PDF files from WebPages		•	•
Digitally sign documents		•	•
Create PDF files			•
Scan Paper into PDF files			•
Create PDF forms			•

⁹ Source: Adobe Systems Incorporated

¹⁰ Only available in the US and Canada.

Summary of Features:

- Convert any document to Adobe PDF
- Mark up and annotate PDF documents
- Apply security options and digital signatures
- Build -in capability for commercial and non-commercial document distribution.
- Creates PDF Web forms
- Allows to do late-stage text and image editing on PDF files
- Reuse text, graphics, and table data from PDF files
- Handles PostScript® 3™ graphics
- Available for Windows / Macintosh and UNIX.

DTP packages

DTP packages are intended to create high quality pages for brochures and documents where a high quality presentation is important. They are somewhere in between a word processor and vector imaging program.

A.7. DTP PACKAGES - QUARK XPRESS AND AVENUE.QUARK

QuarkXPress has earned a reputation as the best professional page layout application on the market. It is, in principle, intended only for professional use. QuarkXPress has a large array of illustration, color handling, and text tools that generate flawless output. However, the program still does not allow the output of completed documents to the Web. According with Quark, version 5.0, due out later in 2000, will include this capability. Version 4.1 has added support for importing and exporting HTML, as well as the capability to import PDF documents. QuarkXPress will not save documents into PDF format directly and Adobe Acrobat Distiller is required. The HTML export tool saves only text flows, not entire documents, so it be cannot used to export a completed publication to the Web directly.¹¹

avenue.quark

Quark recently introduced *avenue.quark*, which will also be part of XPress 5.0. *avenue.quark* is an extension for XML import and export . *avenue.quark* lets users tag the content of QuarkXPress documents and then extract that content in XML format . The next version of QuarkXPress, version 5.0, which is scheduled to ship at the end of the year 2000 will include HTML, XML and PDF export features.

¹¹ Source: CNET.com

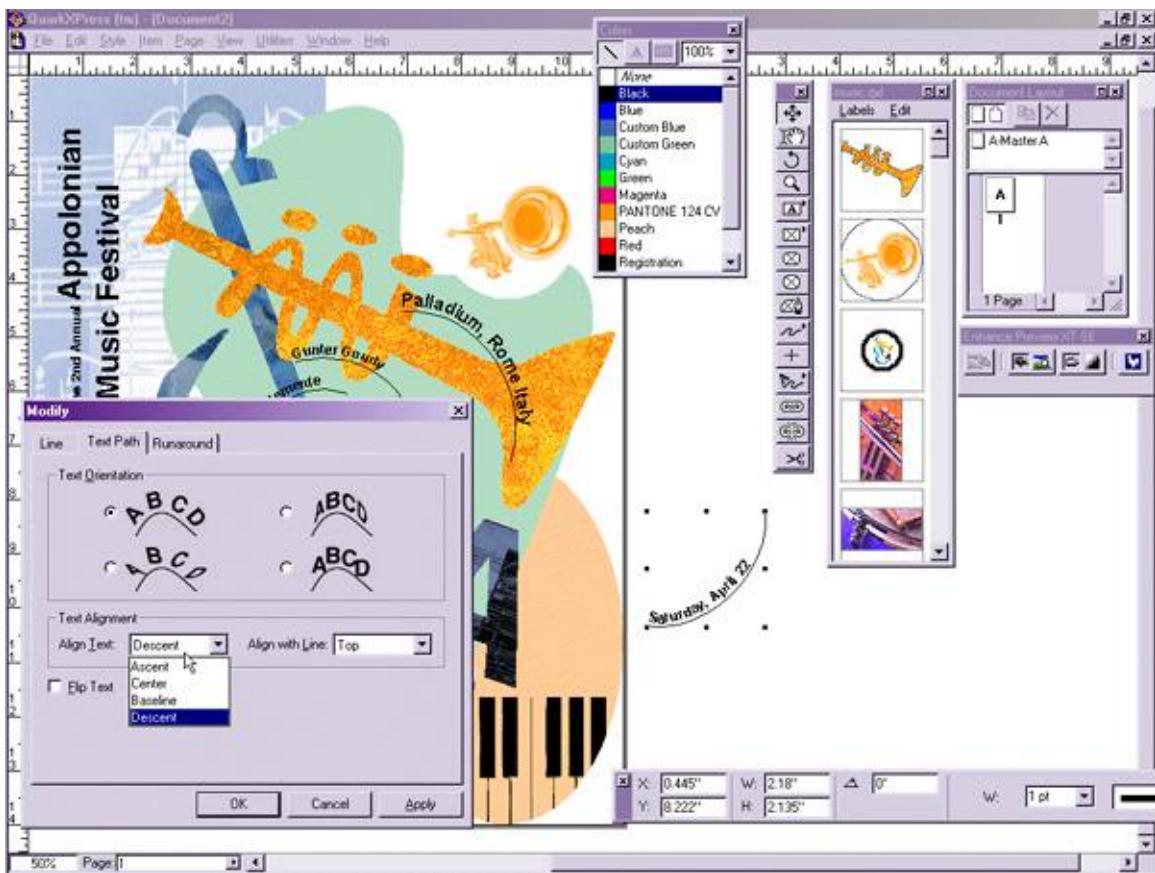


Figure 6 - QuarkXPress screen capture.

A.8. DTP PACKAGES - ADOBE INDESIGN

Adobe InDesign is a page layout program that integrates with Adobe's other design programs such as Illustrator and Photoshop. InDesign is oriented for high-end graphic designers, production artists, and prepress professionals. It is intended to be the successor of Adobe's PageMaker and PageMaker *Plus* and it represents Adobe's attempt to create a competitor for QuarkXPress. The InDesign software was built based on an open, object-oriented architecture. This open architecture creates a core on which third-party developers and system integrators can deliver custom publishing solutions for magazines, advertising agencies, catalogs, retailers, design studios, and newspapers. Many important features in InDesign, such as the ability to create tables, come from third party plug-ins.

InDesign supports PostScript Level 2 or higher, for text, objects drawn in InDesign, and imported bitmaps. InDesign also opens QuarkXPress and Adobe PageMaker *Plus* files, it can save and import PDF files.

Summary of features:

- Compatible with Quark Xpress and PageMaker file formats
- Can create PDF output.
- Available for Windows and Macintosh

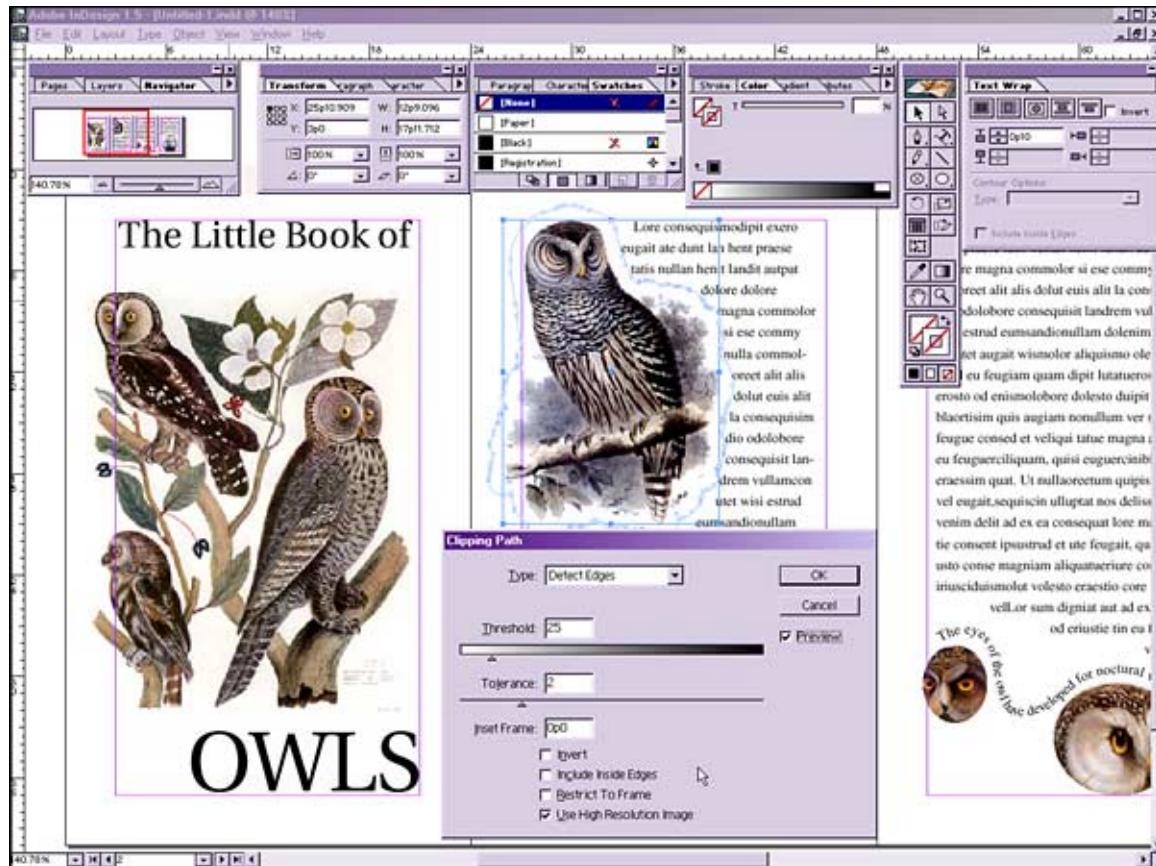


Figure 7 - Adobe InDesign 1.5 Screen capture.

A.9. ADOBE INCOPY

Adobe InCopy is an editorial application built on Adobe InDesign. InCopy is designed specifically for writers, editors and copy-fitters. In addition to InDesign's document creation features, InCopy provides a combination of change-tracking tools that improve collaboration in the document creation process. It provides a simple way for editors to track changes, collaborate with designers, and fit copy easily. It also supports collaboration using WebDAV server technology. WebDAV technology allows for collaboration with other applications that support WebDAV servers. Briefly, WebDAV stands for "Web-based Distributed Authoring and Versioning". It is a set of extensions to the HTTP protocol that allows users to collaboratively edit and manage files on remote web servers. WebDAV support also

enables versioning control to assist editors and designers to efficiently manage changes to a publication in real time.¹²

InCopy is available for Mac OS 8.5, Windows® 98/NT 4/ 2000/ Millennium (Me).

¹² Source: Adobe Systems

APPENDIX B – WEB DEVELOPING TOOLS

B.1. MACROMEDIA DREAMWAVER/ FIREWORKS / DREAMWEAVER ULTRADEV

Macromedia's Dreamweaver is one of the most popular WYSIWYG design tools for the development of web pages (see Figure 1). Macromedia's Dreamweaver 3 Fireworks 3 Studio bundles together Dreamweaver 3 and Fireworks 3.

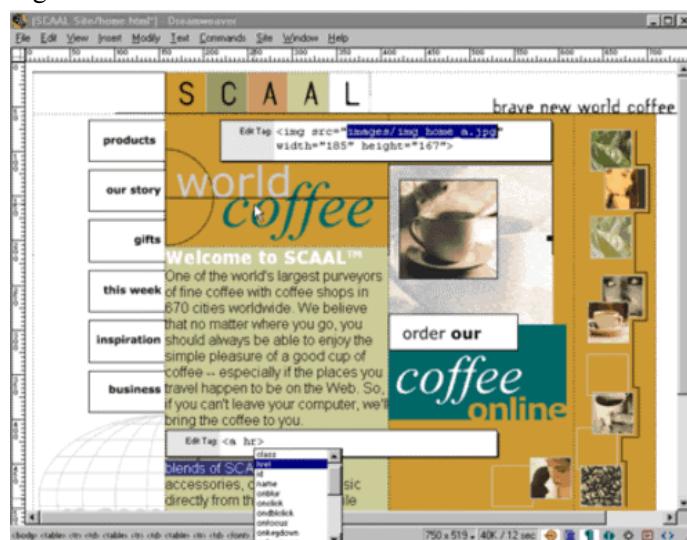
Fireworks is a graphics-production package that is positioned specifically for creating Web images. Some graphics programs let users create rollover buttons with highlights or beveled text and then add the appropriate JavaScript. Fireworks also allows users to work with these images as "states" (MouseOver, Click, etc), which are then optimized for a Web palette with the code automatically inserted in the document or Windows clipboard.

In order to maintain compatibility with earlier generations of browsers, Dreamweaver lets users convert layers to tables, and also includes some new code-cleaning features, including a wizard that specifically cleans up HTML pages created by Microsoft Word. Dreamweaver allows transferring files via FTP by dragging them from one window to the other. The Library palette lets users drag assets, templates, styles, or complete style sheets to and from the Document window. The Behaviors module performs like an object-oriented library, allowing users to drop events or actions onto text or graphics.

The program also includes an HTML source window and provides tight integration with Allaire's text-based editor HomeSite, which is included in the package. Several other bundled programs are integrated with the suite, including Headspace's Beatnik MIDI plug-in and Bradbury's TopStyle Lite Cascading Style Sheet editor.

Macromedia's Dreamweaver UltraDev is an integrated visual design tool for database-driven Web applications, such as e-commerce, personalization, and dynamically built pages. UltraDev combines Dreamweaver 3 and Drumbeat 2000, a development program that produces Active Server Pages (ASP), and IBM WebSphere-specific Java Server Pages (JSP). UltraDev improves those output

Figure 1



Screen capture from Macromedia Dreamweaver

options by adding support for Allaire ColdFusion Markup Language (CFML) as well as support for a more universal JSP format that can run on a wider variety of Web servers.

Dreamweaver 3.0 has its own programmable document object model (DOM) to let developers extend the behavior of Dreamweaver itself with a XML and JavaScript.

APPENDIX C - IMAGING TOOLS

C.1. ADOBE PHOTSHOP

Adobe Photoshop 5.5. provides an integrated solution for professionals who need to produce the highest quality images for both Web and print. Photoshop 5.5 includes Adobe ImageReady 2.0, a Web production component that looks, behaves, and interacts with Photoshop (see Figure 1).

Photoshop includes a set of advanced features for optimizing and compressing Web graphics. For more advanced Web graphics tasks—such as JavaScript rollovers, animation, and image slicing—a “jump to” icon quickly transfers the image to Adobe ImageReady.

Photoshop allows the creation of advanced JavaScript rollovers. An image can be optionally sliced and then exported directly to Adobe GoLive, where rollovers are defined, or the complete rollover behavior can be defined within Photoshop. In the latter case, Photoshop will export complete HTML and JavaScript code that is compatible with Adobe GoLive.

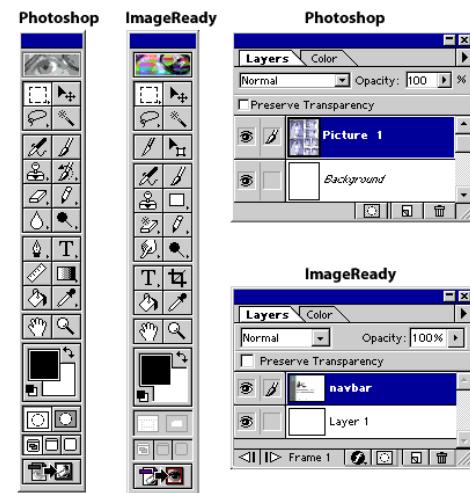


Figure 1

C.2. PAINT SHOP PRO

Paint Shop Pro, made by Jasc Software. Paint Shot pro includes Animation Shop for the design of animations for the Web. PSP includes much of the functionality of Photoshop, such as layers, at a much lower price. Furthermore PSP can use most of the plug-ins designed for Photoshop.

PSP allows to easily view and manage images using thumbnail previews and also allows to easily share images online for free by uploading photos directly to StudioAvenue.com

Thanks to its new vector illustration capabilities, Paint Shop Pro can also double as a basic drawing program. Building on the vector text and drawing tools first introduced in version 6, Paint Shop Pro

7.0 adds more functionality to these tools. PSP can create shapes, lines, and text in vector format. PSP also includes a library of styled lines with dots, dashes, and arrowheads.

PSP includes several tools to enhance digital pictures. In addition to these enhancement filters, Paint Shop Pro 7.0 comes with a set of photo restoration features that allow to fix old scanned photos. For example, remove scratches, restore damaged areas, and liven up faded images.

Paint Shop Pro 7.0 lets users preview how the mage will look using different formats and compression options. The web design features include the Image Slicer, which allows users to slice a large image onto multiple smaller images. The Image Rollover Creator tool can create graphic images for Web sites that change when you move the mouse over top of them. In addition, the Image Mapper tool enables to create image maps for Web site navigation, while the JPEG/GIF/PNG Optimizer compresses and optimizes images for use on the Web.

APPENDIX D - GLOSSARY

ADSL	Asymmetric Digital Subscriber Line
ASF	Advanced Streaming Format
B2B	Business to Business
CDMA	Code Division Multiple Access, a spread spectrum technology that offers 10 to 20 times the capacity of analog networks. Adherents claim that CDMA will eventually provide voice quality and coverage superior to TDMA.
CGM	Computer Graphics Metafile
CSS1 & CSS2	Cascading Style Sheets Level 1 and Level 2 W3C Recommendations. Full support for level 1 and partial support for level 2. URL: http://www.w3.org/TR/REC-CSS1/ , URL: http://www.w3.org/TR/REC-CSS2/
DHTML	Dynamic Hypertext Markup Language
DOM	Document Object Model (DOM) Level 1 W3C Recommendation. URL: http://www.w3.org/TR/REC-DOM-Level-1
DSL	Digital Subscriber Line
GSM	Global System for Mobile, the European standard for digital wireless networks. Based on TDMA technology, GSM uses SIM cards for subscriber identities - and these chips can also double as smartcards, opening up new commerce options.
HTML	Hypertext Markup Language
MP3	MPGE-1 Layer 3 compression
MS	Microsoft corp. - www.microsoft.com
PDA	Portable Digital Assistant
PDF	Portable Document Format
PS	PostScript
SGML	Standard Generalized Markup Language (SGML)
SGML	(Standard Generalized Markup Language)
SVG	Scalable vector Graphics
TDMA	Time Division Multiple Access, a digital technology that divides spectrum by assigning a different time slot to each user on a channel. A TDMA network delivers triple the caller capacity of an analog network.
W3C	World Wide Web Consortium
WMA	Windows Media Audio
WYSIWYG	What you see is what you get
XLink	XML Linking Language W3C Working Draft. Support for simple XLinks. URL: http://www.w3.org/TR/xlink

XML eXtensible Markup Language
XPointer XML Pointer Language
W3C Working Draft. IDLoc XPointer. URL: <http://www.w3.org/TR/WD-xptr>

