Developing and Publishing High-Quality Electronic Documents on CD-ROM

Steve Cunningham
California State University Stanislaus
http://www.cs.csustan.edu/~rsc/
cunningham@siggraph.org

Notes for Eurographics '97 tutorial, Budapest

Contents

Overview and Credits

Electronic publishing is the enterprise of producing content in electronic form that can be delivered by electronic media. These notes explore the opportunities and challenges in this kind of publishing and help the reader evaluate the different kinds of approaches that may be taken in this area.

The common theme that runs through all this material is that the editorial content and the needs of the reader must come first, and the publisher must choose the publication process that fits the needs of the content and reader. Thus we will find that the publisher may sometimes publish on paper, sometimes on the Web, sometimes on CD-ROM, and sometimes on a blend of two or more of these,. The publisher may sometimes publish with one authoring system and sometimes with another. The key is to know the content and the audience well enough to provide the best way of reaching any particular audience with any particular content.

Like any other area of publishing, electronic publishing must be built on a solid editorial base. This base ensures consistent content quality and focus, and is key to creating value for the audience. However, content that is published in each medium (print, sound, video, film, etc.) tends to have its own editorial nature, while electronic publishing can contain and integrate all these media, so the editorial task for electronic publications is probably more challenging than for each individual medium.

The fundamental questions for electronic publication production are the formats for content components, the tools for producing content in these formats, the authoring tools for integrating these components and for presenting the material to the audience, and the media used for delivering the integrated content to the audience.

Electronic publications are quite new and many people are uncomfortable with the notion of publishing in this way. However, electronic publications have become much more capable in the last year or two and are now at a point where they can begin to displace or complement paper publications for many areas. Technical publication is probably the area that offers the greatest promise for this change. Persons within the technical community are among the most likely to have access to the tools needed to read and use electronic publications, and are among those most likely to be comfortable with this new way of thinking about communication.

For this community, both HTML and Acrobat are extremely capable technologies for creating and delivering publications that include the added functions that make electronic publications so attractive. The choice between them should probably be based on two things: the degree to which high-quality page layout and design is important to the publication, and the degree to which the added value of applet-based content is important. Acrobat is much stronger in supporting

high-quality layout, but HTML offers access to applet capabilities. The additional production factor of prepress page layout tools (for Acrobat) or HTML authoring tools (for HTML publications) is probably less important because either of these can be acquired and staff can be trained as necessary.

The other dimension of electronic publishing is how to deliver your publication to your audience, and here we see the two choices of CD-ROM or networks. The choice here is probably clearer, at least for now, because of the stability and low real cost of the CD-ROM medium as contrasted with the instability of the networks and the often hidden costs of maintaining a high-function Web site. However, this is today's answer and we are not at all sure that it will be tomorrow's answer, because the networks are quite capable of being enhanced and increasing site automation may be coming.

Whatever direction you choose, however, we can guarantee that developing an electronic publishing capability will be exciting and sometimes frustrating, but you will find that the results will pay dividends in the amount and quality of information you can publish and the value your audience will get for it.

This tutorial focuses much more strongly on technical types of publications than general educational or entertainment publications. These notes examine the two main publication formats for technical content and the two primary delivery media for that content.

These notes draw heavily on the book, *Electronic Publishing on CD-ROM*, from O'Reilly & Associates, co-authored with Judson Rosebush. They also owe a great deal to my experience with publishing the electronic proceedings of the annual ACM SIGGRAPH conference. Much of the section on disc manufacturing was provided by Steven Langer and Breck Rowell of Disc Manufacturing, Inc., and their help is gratefully acknowledged. I would like to express my appreciation to Judson and to SIGGRAPH for the opportunity to work with them.

The Problem

Publishing is a familiar and historic activity, going back hundreds of years. The publisher may fairly be said to be one of the backbones of the whole growth of societies and cultures, and publishing is not an enterprise to be taken lightly.

In the introduction above we noted that we must meet the needs of the information to be published and of the audience. Information comes in many types, and each type has its own demands. Some are archival, while others are continually evolving. Some want to have free and open access, while others want to be restricted or even private. Some need high-quality design, while others find this a minor issue. Some have limited media needs, particularly those that are mostly text, while some need to be expressed through a great deal of media. And some are fundamentally passive, while others need to be presented very interactively.

Similarly, an audience has its own set of demands on published material. People have their preferred working styles, with some wanting to have material that can be heavily annotated or that can be available away from the desktop. Some people want simply to get information, while others want to have an experience. Some material is needed only rarely while other needs to be always available.

The range of demands on published material imposed by this extremely wide range of content and audience needs has put a significant strain on the traditional print publishing world. Print is being found less than satisfactory for some content and audiences are looking for something better.

There are important reasons why electronic publishing has become a significant tool for many applications. Electronic publishing allows us to step beyond some of the production and distribution costs and media limitations of print. Further, traditional paper publishing is an expensive endeavor and in many ways is limited by the print and paper medium.

Many drawbacks of print become evident when we look at the nature of things we need to communicate to our audience.

We want give our audience visual information and experience, so we must publish images. But producing images in print requires halftones or separations, and these compromises make it virtually impossible to reproduce images faithfully.

We also want to publish materials with integrated animated sequences and video, and it is clearly impossible to do this in print unless you want to publish "flip books" with an image on each page.

In areas that use computer graphics for exploration and analysis, such as engineering and the sciences, as well as in training materials, we want to publish interactive graphical simulations so that the reader can interact with the concepts that are being presented.

Finally, it can be important to provide your audience with original data sets for some studies so they can make their own tests to confirm or expand on the work.

Each of these reasons argues strongly for a medium beyond print. Some of them can be satisfied by other media, such as video for animations, but it is clumsy to present a subject through a number of different, separate media. They can all be satisfied in an integrated way by electronic media.

The production costs of publishing are simple to state.

Paper, particularly high-quality paper as needed for printing with color, has become very expensive. Some publishers now see the cost of paper making up as much as 30% of the entire publishing budget.

If you want to include color in print, as we must in computer graphics, you must deal with the cost of producing color separations and of using color presses; these costs are so high that most book publishers only include color very sparingly, as many authors can attest.

Once you have published something in print, you must ship it to your audience or store it for future shipments, both of which are expensive because paper is heavy and bulky, and has high costs for shipping and warehousing.

Finally, if you have a successful publication and need to reprint it, you must repeat all the printing costs

for the reprint job. Print remains the most widely used publication medium, and there is a great deal of value in print publications, but these costs have made many people in the business wonder how long that will continue to be true.

Electronic Publishing Opportunities

In the last few years we have seen the rise of two new kinds of technologies that make it possible for us to view electronic publishing as a workable enterprise. The first is the set of technologies for creating, storing, combining, and playing back individual kinds of electronic information and creating real electronic publications. These technologies are often lumped together under the name of "multimedia." The second kind of technology is the set of delivery media for presenting this content to a publication's audience.

The technologies for creating, storing, and playing back the individual parts of a document are as varied as the parts themselves. For each, there are questions of encoding and compression, because the parts must be stored digitally and it is important to store them efficiently. A brief list of the kinds of parts we would need and some of the supporting technologies is given below, and you will notice that the wide variety of available (and often overlapping) technologies will make it necessary to select from these options. This list cannot be complete, but it includes the most common and most important pieces.

- Text: ASCII and Unicode encodings; LZW compression
- Images: JPEG, GIF, TIFF, PICT, TGA encoding for color images; CCITT Group IV and JBIG for monochrome; JPEG, RLE, LZW compression
- Text and image layout: PostScript and PDF
- Text components: SGML and HTML
- Digital movies: QuickTime, Video for Windows, MPEG encodings; Cinepak and Indeo compression (codecs)
- Sound: AIFF, WAV, MIDI encodings; ADPCM compression
- Simulations and interactions: VRML, Java

Selecting the technologies you will use from this list is often done when you select your authoring system, the primary tool you use to assemble finished documents from the various parts that are created individually.

Authoring systems

An authoring system is a language and/or toolset which allows you to produce electronic and multimedia documents that can be read by appropriate delivery tools. An authoring system allows the author or developer to take the assets for a publication—the text, images, video, and sound—as well as the layout and interaction design for the publication, and combine these components into a final product that gives the user access to the publication's contents. An authoring system integrates the work of the specialized tools you use to create the individual types of media in your publication (e.g., a word processor for text, a QuickTime creator for movies, etc.).

Authoring systems have not developed as fully as individual media technologies, however. For electronic publications that involve high levels of user interaction and dynamic presentations, it is difficult to find tools that allow an author to use all these kinds of components fluently. Fortunately, publications for the technical community are not as demanding as those for entertainment use, and we find that there are reasonable tools for integrating the necessary kinds of content for these publications. These kinds of authoring systems will be discussed later in these notes.

Delivery systems

Media technologies to deliver finished publications to your audience are pretty stable. Basically you want technologies for getting large amounts of digital content to the individual reader. The traditional medium for distributing digital content has been the diskette, but with only some 1.4 MB of storage capacity it is far too limited to be useful for publishing (though some new diskette technologies may remove this barrier). Many parts of an electronic publication, such as digital animations, will be larger than a diskette can hold; digital video clips are often tens or even hundreds of megabytes long. At this time are only two delivery media are feasible for electronic publishing: the CD-ROM and the networks. Both are capable of delivering any kind of digital content and both are becoming widely available.

Some people question whether there is enough of a potential audience to support a real electronic publishing program. We have two sets of data on this: of all ACM members, 79% have CD-ROM players and 83% have WWW access; of technical program attendees at the annual SIGGRAPH conference in 1994, 93% had CD-ROM players and 94% were on the WWW. So at least among the technical community, the audience for these publications is definitely there!

CD-ROM

CD-ROM players have become a very common feature of personal computers, and it is becoming unusual to see a new PC that does not have such a player installed. Over 22 million CD-ROM players are now in use worldwide. If every computer being sold does not yet contain a CD-ROM drive, it will in the very near future because CD-ROM is the preferred distribution medium for anything over a few megabytes in size, especially software. CD-ROM will be supplanted by the Digital Video Disc (DVD) system, which was launched earlier this year, but this was delayed by problems with copyright from the movie companies whose products are expected to drive large-scale consumer acceptance. As of mid-1997, built-in DVD-ROM players are already starting to be included in high-end desktop systems. These will play CD-ROMs as well as digital movies and will further the spread of CD-ROM capability in individual computers. It is evident that the world has made a cultural commitment to CD-ROM.

The data on CD-ROM use show that the dominant market for CD-ROM is now in the United States. Japan's pattern of adopting CD-ROM is very close to

the United States. Europe is on a parallel growth path with the United States but is three to four years behind the US curve. The numbers in other countries should grow quickly as the CD-ROM and DVD devices reach the same proportion of computer desktops as is seen in the United States.

Networks

Like CD-ROM, the capability and availability of computer networks have grown explosively and the networks now provide an important medium for distributing electronic publications. The Internet has grown from a couple of hundred hosts in 1981 to 4,000 in 1986 to 500,000 in 1991 and to almost 10 million at the start of 1996 (with a projection by the authors of about 12 million by the middle of the year). These numbers are from the Network Wizards at

`http://www.nw.com/`

and are accepted as definitive. They are summarized in Figure 1. The number of Internet users is also growing very rapidly, with 11 million households online at the end of 1995 and 18 million online at the end of 1996. These numbers are fairly compatible with projections suggesting that 200 million persons will be using the Internet by the year 2000.

By far the largest country on the Internet is the United States, with over 60% of the hosts as of the middle 1990s, and Australia and New Zealand are have comparable per-capita Internet use. Most of the rest of the world, especially Europe, is on the same growth path as the USA. However, Japan has fewer than 2% of the overall number of Internet hosts, apparently because there is no government agency promoting networking in Japan.

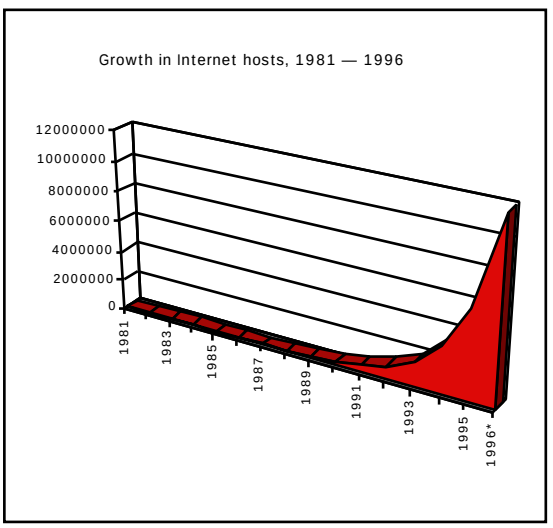

Figure 1: Growth in actual numbers of Internet hosts, 1981 – 1996 (1996 is estimated based on early figures for the year)

The growth of the Internet has led to an explosion in traffic that makes it difficult to distribute large and media-rich documents, and this problem has grown acute as the number of users online continues to expand. Data transfer rates are frustratingly low for all classes of users, from research institutions to private homes. It seems risky to base a major publishing enterprise on network distribution, especially if it involves interactive publications with high-resolution pictures and video which can take many minutes to transfer. However, it may be worthwhile to consider the Internet for technical publications where the user can take the time to download a full document to be read on a local system. The problems of network capacity will change, and there are leading-edge network experiments that are starting to develop the advanced networks that will be needed to this to hapen, but that will take some time to accomplish.

One feature of the Internet that publishers must consider is the culture of network users. There is a general attitude that information must be free to anyone who wants to access it, and it is easy to see how such an attractive concept has grown up. Yet part of the world of publishing must be developing income streams that support the costs and returns that must be available to keep publishing alive; so far it has been difficult to build a business around network publishing.

Digital libraries

As an example of the opportunities presented by electronic publishing, consider the digital library. The traditional library has served its patrons by building collections of materials for the reader, by providing the reader with catalogs and indexes to these collections, by preserving the historical record of the areas in its collections, and by guiding the reader who has specialized needs in the collections' areas.

Digital libraries, which are shared collections of digital materials that may be widely distributed geographically, have the same goals. But these new libraries have access to new document formats and new technologies to support access, catalogs, indexes, and expert assistance. They also allow access to collections all across the world. By supporting new kinds of materials, they also broaden the concept of what a collection can contain.

Examples of the new libraries include specialized collections of digitized documents from traditional libraries, formal and informal reports, multimedia documents, and even data sets that allow researchers and others to access and use them in unified ways with computational tools. This provides a much broader and deeper kind of collection than any but the wealthiest library could ever dream of building. These efforts will expand to support educational work and eventually many will be available to the general public.

Document Authoring and Development

The overall process of authoring and developing a publication is quite extensive and will not be described fully here; it is covered at length in the Cunningham & Rosebush book. To summarize, however, you should plan for the following steps, either formally or informally:

- Develop the publication's concept. You should be able to tell someone in a few well-chosen words

just why this publication must be done, and why it must be done in electronic media.
- Create a prototype of the publication so that you can show others your vision. This is especially important if you must get funding for the project.
- Assemble the design team who will work with you to work out the details of the publication and to develop the specifications for it.
- Create the functional specification for the publication—state exactly how the publication will work. Remember that an electronic publication is not just content, but is also presentation and interaction; these are the areas covered by the functional spec.
- Create the technical specification for the publication, outlining in detail how the functionality specified above will be created.
- Assemble the content team who will work with you to develop the contents for the publication. This team will include individuals who have skills in each of the media areas you want to use for the publication.
- Acquire or develop the content that will go into the publication. This can involve licensing existing material or creating new material, depending on the nature of the publication. This step is the one that is most like the usual concept of authoring and publishing, but it can be more complex for an electronic publication than for print.
- Integrate the contents you developed in the previous step into the publication This step also implements the graphic design for the publication and builds the interface as specified earlier.
- Test the material thoroughly, at the stage of individual content materials and at the integrated content stage. This testing must cover both correct functioning of the parts and the whole, and working correctly on the platforms and configurations that are intended for the publication. After each test, appropriate corrections must be made and the project returns to an earlier stage to continue.
- Prepare the publication for production by creating its final assembled form and getting it to the manufacturing or distribution source. If the publication is online, then this means moving it to the actual Web site; if the publication is on disc, then it means creating the disc image and getting it to your manufacturing plant.

Business Issues in Electronic Publishing

These notes are not intended to cover many of the areas of costs and income that are involved in a professional publishing program; these are described much more fully in the author's book on the subject. But there are two items that should be mentioned before we go on with other issues.

The first is the cost of developing your products. The activities of title development that are described above are not usually done by amateurs for the sheer love of the work. In the professional world, they are done by professionals who have serious skills and who expect to be paid for their work. The cost of developing a mass-market "edutainment" project can easily be several hundred thousand dollars; those of developing a technical title would be less because of the reduced need for content licensing and the lesser amount of media content that must be integrated into the project. But it can still be significant, because we have seen quotes as high as $100,000 to develop a single conference proceedings CD-ROM (those quotes were not turned into contracts, needless to say!)

The second is the question of marketing and distributing electronic publications. Those of us who work with professional societies have the opportunity to have our publishing funded by membership and conference revenues, so we do not have to deal with the mundane, but necessary, issues of finding buyers for our products. But if you plan to produce an electronic journal or other electronic products, you must find a way to make your potential audience aware of your publications and to get your products to them. These are specialized topics—people get degrees in marketing and distribution from business schools—so we will not try to cover them further.

Both development costs and issues of marketing and distribution are discussed in the book mentioned at the beginning of these notes, so we will not discuss them further here.

Producing Electronic Publications

Electronic publications have both similarities and differences compared to print publications. Print involves a relatively small production team: editors, proofreaders, separators, and printers. With the new kinds of media that you can have in electronic publications, you need editors and proofreaders, you need people skilled in creating or scanning digital images instead of separators, and you need either disc manufacturers or Web site managers instead of printers. But you need people with other skills as well: people to create or capture digital video, to create or capture digital sound, and to integrate all these media into the overall publication. This group needs to function as a team to create the publications efficiently and cost-effectively.

The technologies for producing electronic publications also differ from the technologies used to produce print publications. We must make a clear distinction between electronic prepress, where materials intended for print publication are manipulated and assembled using computers, and authoring systems for electronic publications.

Authoring systems resemble prepress systems in working with digital media components of publications, but they also define and create the interface between the publication and the user and define the functionality of the publication—how it manages interactions and displays its components. Navigation and media integration are the most important capabilities they offer. Authoring systems are thus much like programming environments because the publication's developer must define interface components, responses to events, and actions that the publication will perform.

Acrobat and HTML publications will both have the same kinds of components and will both need the same kinds of teams described above. However, these two document systems have different kinds of authoring systems and integrate their components differently, so

we will describe the details of producing documents in the two systems separately.

Properties of technical publications

Technical publications are somewhat different from other kinds of publications found in the electronic world. They are very strongly oriented to text and use other kinds of content, such as images, movies, sound, and simulations to amplify and illustrate what the text is talking about. This differs from strongly activity-driven products such as video games, picture-driven products such as multimedia entertainment, or video-driven products such as some educational discs. This has important effects on our choice of authoring and delivery systems, and makes traditional multimedia authoring systems such as MacroMedia Director less suitable than one might expect.

While we will not discuss the kind of multimedia publication that is usually created with Director-like authoring systems, we'll simply note that there are some kinds of training or educational materials that are based around animations or video where these are appropriate. The omission of Director does not mean that you should not consider such systems for authoring and delivering your documents.

There are a few fundamental features that the finished electronic versions of technical publications should have:

- they must have at least the same production quality as printed versions of the publications would have,
- they must be able to be produced quickly and with a reasonable cost,
- they must be readable on all the kinds of computers commonly used by members of their audience,
- they must be able to have at least as much information as would be in the printed versions of the publications,
- they should be able to include all the kinds of media content, such as color figures, movies, or simulations, that are important in their particular field, and
- they should be compatible with whatever delivery media we choose.

As we look at our options in choosing authoring systems and document formats, we need to keep these issues in mind.

The two best opportunities for technical publications seem to be provided by Adobe's Acrobat document production and delivery systems, and by HTML authoring and Web browser delivery systems. Here we will describe the basic characteristics of these systems, and later we will discuss actually producing and delivering publications based on each.

Adobe Acrobat

The Acrobat electronic document system from Adobe Systems was introduced in 1993. It has had an important impact on how publications with a primary text base can be produced and delivered electronically. Acrobat is built on PostScript, and because of this it supports very high-quality design for publications. This is as important for technical publications as it is for general publications; among its benefits is the easy integration of figures, diagrams, tables, and equations in a publication. In particular, the readers of technical publications seem to place a very high value on print and want to either have printed pieces be their primary source, with electronic versions as backup, or to be able to print high-quality copies of electronic originals, so it is important to keep print in mind as documents are developed.

PostScript has two fairly well-known problems: it produces very large files, and there are many obscure ways to present a PostScript file. The Acrobat system has its own file format called PDF (the Portable Document Format). This format contains some features that improves PostScript's architecture. In addition, PDF files offer a number of options for images (JPEG, CCITT Group III) and uses LZW text compression to reduce the size of PDF files. PDF files may be stored in either binary or ASCII formats, and the latter may be sent by ordinary email.

The Acrobat document reader provides a number of useful standard features. These features are provided by the reader, not the publication, so all Acrobat publications will work in the same way for all readers. An example of a Acrobat reader screen is shown in Figure 2; note that this has text over a background figure, something that is easy to do with a PostScript heritage. The standard page-oriented nagivation tool is the palette of buttons at the top of the screen.

The user can select an outline view or thumbnail view to accompany a document; can zoom into the text or figures; can navigate pages sequentially or retrace the sequence of pages that have been viewed; can execute a link or a plug-in to move around in the document or to view an accompanying document; can select text from the document to be pasted into another document; or can do a word search on the document. A user reading the document with the Acrobat Exchange can also use indexes to speed up word search. These searches are quite intelligent and can be based on word stems and homonyms besides the usual full word matching. Finally, because of the PostScript basis of Acrobat, the Acrobat reader can easily print out any Acrobat file to any PostScript printer, including full color if the printer is color-capable.

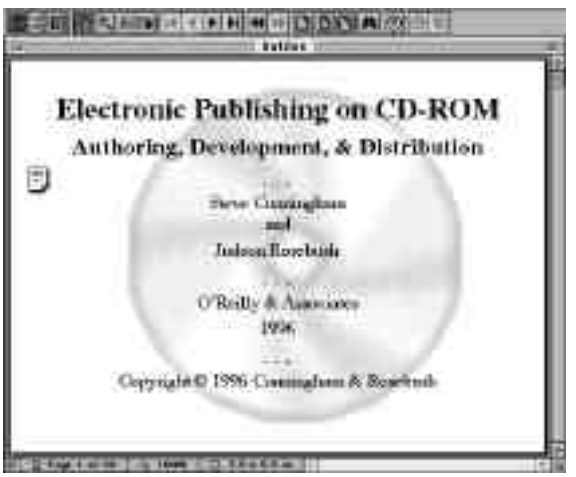

Figure 2: a sample Acrobat Reader screen

The original release of Acrobat provided faithful and high-quality on-screen versions of printed pages and supported internal hyperlinks, searching, and copying. Its capabilities have expanded to include in-line movies and sound clips, hyperlinks to other

Acrobat documents, and even Weblinks to the entire World Wide Web (WWW). The latest release of the Reader is compatible with Netscape Navigator and other Web browsers and allows Acrobat documents to be viewed within the browser window. Overall, Acrobat is proving to be a remarkably capable publishing system.

HTML and Web browsers

HTML, the HyperText Markup Language, is a tagged-text system that allows an author to insert tags into a document's text that describe how that part of the document is to be displayed. HTML files are displayed by a Web browser, an application that understands these tags and produces the display that is required. The tags available in HTML can include various levels of text, can specify fonts and sizes, and can include links to materials such as images that are stored outside the text. One of the issues with the Web environment is that the set of allowable tags is not consistent between different browsers and is constantly expanding for each separate browser. HTML standards are being developed, but the field is moving so quickly that it is difficult to keep up with it. In fact, at the risk of being slightly heretical, we would suggest that one of the problems of the Web is that people seem to be adding tags and functions to HTML more to make their browsers distinct from other products than to create the best possible tool for publications.

An example of an HTML file displayed with a Web browser, in this case Netscape Navigator, is shown in Figure 3. This shows the various kinds of tools available in a typical Web browser. Some of them support working with the particular document, but most of them are oriented towards the browsing aspect of the application. So we see tools to open a remote document, to use any of a set of user-defined hot links, and the like.

Figure 3: a document presented by Netscape Navigator

Web browsers are very capable in their ability to recognize kinds of files and open appropriate applications on the home system to play a downloaded file from across the Web. Most browsers allow a user to add to the number of media players the browser will recognize. More and more, browsers are integrating these players so that a file is played in-line in the browser window. Most browsers will play digital movies in-line now, and several have the ability to open an Acrobat window with their own window to play an Acrobat document. At least one will play VRML documents. In fact, integrating a growing set of capabilities across a wide set of host computers is one of the most important things about the Web; it is breaking down the differences between systems that has been a major barrier for years.

Although Web browsers offer solid capabilities, HTML publications have two main problems from the user's point of view: navigation and printing. Both of these are the result of the structure of an HTML document. This is made up of a number of separate files that are linked through tags in the various parts. These files may all reside in one directory system, such as on one CD-ROM or one server, or they may be distributed anywhere in the world. The structure of such a set of sections, particularly as one might find them for a publication such as a technical article, is shown in Figure 4. Here there are one or more overall index files, such as a table of contents or a keyword index, and each entry in the index is linked to a section of the document. The section may contain images and/or movies as well as subsections, and this structure is recursive because the each subsection may itself contain subsections, images, movies, or applets.

Navigation is a problem because HTML publications are made up of many different components, and a user can lose track of his or her location in the overall set of files. This problem is reduced by keeping a history of the user's file accesses, as most browsers do. However, each section of the publication should contain a short table of contents for the overall publication and an indication of which one the user is in; then there is no question of the user's location and there is a tool to allow the user to move to any section from the section where he or she currently is.

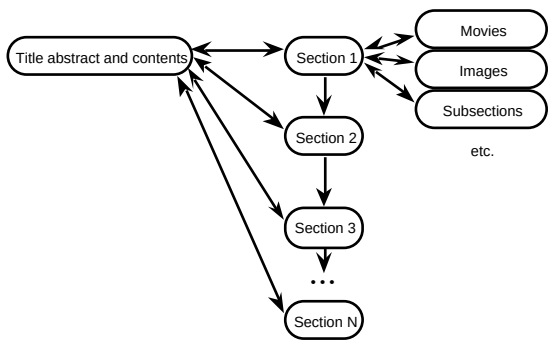

Figure 4: A possible structuring of the components of an HTML publication

Printing HTML documents can also be difficult. Even for electronic publications, there will be times when a user really needs a printed copy. Since HTML publications are developed as many individual files, it can be difficult to get a reasonable printed copy of the entire publication. A printable version of the publication should be available and the HTML file should include a link to that printable version to make it easy for a user to get a printed copy.

HTML documents are primarily text documents with additional components such as images or movies. These components are linked together to produce the screen presentation to the user. HTML components can be played in the browser window or, by launching an external player (a "helper application"), in separate windows.

Creating an HTML document involves creating the text part of the document and adding the linking information that displays all the other components, both text sections and the media components. This text may be created by an author or by another application, such as a database system that returns results of database queries as HTML documents. In these notes we will concentrate on author-written documents, but automatically-produced documents are important to consider for some uses.

HTML files use a tagged-text file format, in which tags tell a document presentation system how to interpret and present various parts of a file as well as other files as defined by different kinds of links. For example, a block of text that is to be italicized will be written

```
<i>Italic text goes here</i>
```

or a reference to include an image in the same directory would be written

```
<href IMG="filename.gif">
```

The key to producing HTML documents is in adding the tags that describe the text components, provide details of the layout of information on the screen, and create links to include components such as figures, movies, or applets in the document.

A number of different tools can create such tagged text. The most fundamental way is probably to write the HTML document in a word processor and add the tags manually. This is awkward and can take a great deal of time, especially for publications that involve a large number of individual files or a large number of style tags.

Another approach takes more of a WYSIWYG approach and allows the author to select the operations that are needed, shows the results of those operations, and automatically writes the HTML tags that perform those operations. Thus an author would select a block of text to be italicized and it would be shown in italics on the screen, while the pair of tags `<i>` and `</i>` would be inserted into the text invisibly to the author. Inserting figures would be done by making a menu selection that requests a figure; the figure would be shown as it would be displayed by a Web browser, and the `<href ...>` tag for the figure would be included automatically. This is the approach followed by the newest set of Web authoring tools, and is probably best for most authors; the only difficulty with using such a tool is that the number of new HTML tags is growing so quickly that these authoring tools can quickly become out of date.

A more advanced approach allows an author to write in a word processor or page layout program and to create links and media inclusions directly in the document as it is written. When the document is saved, one of the options can be to save it as an HTML file, which then has the text links, text formatting, and media links for the final document.

These tools do not solve all the problems, however. Of special interest in technical publishing is the ability to include formulas in a publication. This is not easy in HTML, because you cannot produce the special kind of formatting necessary to produce formulas in text form. The problem with formulas is solved, if one can call it that, by making formulas into graphics, usually as GIF files, and including these small graphics in the text. This can be reasonably effective for display, but it does not allow the reader to search for text in the equations.

Another interesting and relatively new component of an HTML publication is the applet. The name is the diminutive of app, meaning application, so an applet is primarily a small, embedded application that can be called through the HTML file. Applets allow you to include almost any kind of functionality you could imagine in an HTML file; early examples include interactive animations and scrolling text. In the context of technical publications, useful applets might be simulations that are presented to illustrate a point or demonstrate a theory, and the publication's reader can then interact with the simulation to experiment with the content. This opens immense opportunities for the author and reader.

Applets are typically written by programming in a language that has the necessary capabilities built in to link its result with a user across the networks. The two main languages for producing applets at this time are Java, from Sun, and Visual Basic, from Microsoft. Java compiles its applets into a device-independent byte code that is downloaded to the user's Web browser and is played from there, while Visual Basic's applets seem restricted to the Windows environment. Since both require programming, it may take programming assistance for some authors to use such applets in publications.

Finally, since we have just mentioned applets as a way of including simulations in HTML publications, we should mention the Virtual Reality Modeling Language (VRML; current level is VRML 2.0) as another way to include simulations involving virtual reality into a publication. This is a still-developing standard but Web browsers are starting to support it, making it possible to include VRML-based simulations with other content in a publication.

Designing for the Web

Publication design is an important issue these days because audiences are used to very highly-designed magazines, books, movies, and video. Our tastes are much more sophisticated and amateur-looking publications are likely to be poorly received.

Acrobat is clearly intended to address this design issue directly. All the design possibilities of print prepress translate directly to Acrobat publications because of the key use of PostScript in prepress and the link between PostScript and Acrobat. So if you need high-quality, flexible design for you publication, Acrobat is strongly indicated as one of your choices. However, you must realize that you are designing for

presentation on the screen, not on a printed page, so the overall shape of your design should be oriented to match the screen layout.

On the other hand, HTML documents on the Web are in many ways very primitive in their design. In spite of the continuing increase in browsers' display capabilities, indicated by the growing number of tags and tag types, many of the features of the displayed page are not under the control of the author. For example, if an author defines a section of text to be emphasized, he or she does not know what that will mean when the text is displayed; emphasis could be boldface, italic, a different font, or a different size.

Some of the differences are even more fundamental. Users will read the publication on systems with different sizes of screen, in different sizes of windows, with different kinds of screen fonts, with different numbers of colors and different treatments of color, and with browsers that do not support some of the features the author has chosen to use. Designing publications for this environment becomes a significant challenge. At this point, strongly-designed pages tend to use techniques such as large images with cgi scripts to provide clickable regions, but these are slow and somewhat awkward for many users.

There is work underway in the industry to reduce the amount of variation in page presentation and to create a more design-friendly environment for Web pages. This includes developing standardized fonts for Web browsers, so that everyone will know what the user's text displays will look like, and developing style sheets so that an author can create a style sheet that will accompany a publication and ensure that headers, emphasis, and other design features will be presented correctly. It is important to consult reliable sources in Web design so you will know what issues you face in creating HTML publications; these issues are subtle and can be easily missed if you assume that what you see on your own screen is the same thing that others will see on their screens across the world. Some books that discuss design issues for Web publications are listed in the References section.

Developing Documents in Acrobat

Adobe Systems' Acrobat was released in 1993 and has proven to be a valuable tool for developing and presenting electronic publications. The author first used it for a conference proceedings CD-ROM in 1993 and has used it several times since then. Acrobat is available for most platforms in common use: DOS, Windows, Macintosh, and Unix systems from Sun, HP, and SGI (with others probably also available).

Since a large part of the question of publication development is the production and delivery environment you will use, let us examine Acrobat from three points of view: using Acrobat's tools to produce electronic documents, how effective Acrobat documents are in supporting electronic document functions, and how Acrobat works as a delivery system.

Producing Acrobat Documents

Acrobat's foundation in PostScript technology has many advantages in delivering publications with the kind of high-quality design that PostScript allows for page layouts. The documents that Acrobat produces are in Portable Document Format (PDF) format, and PDF is supported by many kinds of pre-press systems.

Acrobat's production environment is quite solid; it includes the following main tools:

Distiller
: A program that translates PostScript files into PDF files in batch operations or across networks

Exchange
: A program that allows the user to assemble documents and add functionality to them

Catalog
: A program that creates indices into sets of documents

PDFWriter
: A pseudo-printer that creates Acrobat PDF files from other applications

The use of the Distiller and Exchange in creating a document are shown in Figure 5. The lowest level in the Acrobat system is the Reader, which displays the document to the user and supports the functionality in Acrobat documents. The Acrobat Reader is available for most computer systems. Adobe allows the Reader to be distributed at no cost, so it can easily be put onto CD-ROMs.

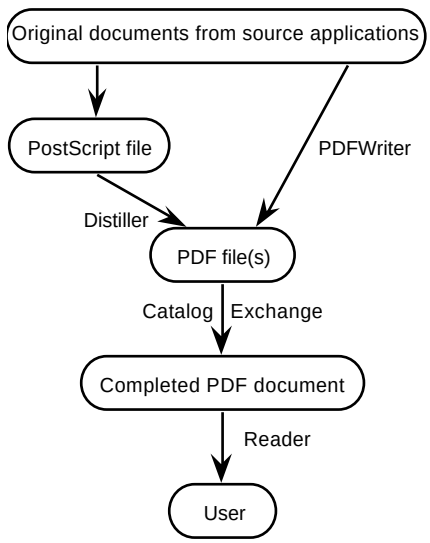

Figure 5: Processes in creating a finished Acrobat document

One way to create an Acrobat PDF file starts with a PostScript file, which most computer-based document creation systems produce easily. The Distiller translates these into PDF files through batch processing, either locally or across a network. A user sets various parameters controlling details of the process. For example, you can set the JPEG compression level for color images, you can choose to create your PDF file as an ASCII or binary file, or you can choose to create thumbnails for your pages automatically.

Another way to produce PDF documents is to use an application that can produce PDF directly. This capability is becoming relatively common, both in Adobe products and in other design tools, because of

the widespread use of PDF as a vehicle for sharing document drafts over the network or by e-mail.

An alternate method of creating PDF files uses the PDFWriter, a pseudo-printer that is part of the Exchange product. This "printer" can be invoked as though you are printing the document and it produces a PDF version of the document. This generally works quite well, though it has a few limitations.

The linking capability of the Exchange allows you to create hypermedia versions of several traditional kinds of reader assistance, such as a table of contents, an author index, and a subject index. We can link these to the main text so that a reader can click on a paper's title in the table of contents, or an author's name in the author index, and the author's contibution will move to the screen. These additions make a large difference in the usability of the document.

Once the original PDF files have been created for each piece, you can use the Catalog application to create an index for one or more documents and the Exchange application to put additional capabilities in the file. This is a critical piece of the process, because many things that make electronic documents special are created here. These include browsing support, notes, indexing, linking, document following, and setting an opening view on a publication. Exchange has the ability to include movies and sound clips within an Acrobat document. It also lets you merge documents or add and remove pages in documents, so you can create a document in segments and assemble the segments into the finished piece.

Acrobat's production process is very smooth and takes little special skill (no programming is needed, for example). The only special work needed is the work of creating links and adding video and sound clips. Although there are applications that support this process automatically, it is usually done manually and is fairly painstaking, but it is not difficult. Creating a finished publication from its various parts is both fast and straightforward.

Producing SIGGRAPH's electronic proceedings

The version of this production process that the author has used in producing the SIGGRAPH Proceedings CD-ROMs is shown in Figure 6. Authors download their files to SIGGRAPH's Internet site, usually by ftp, from where they are uploaded to the production editor's desktop system. The PostScript files are sent to the Acrobat Distiller for translation into PDF files in a batch process. As papers come in over two weeks, the PostScript files are translated and the other files are checked to be sure they work with a standard set of viewers, and then the files are stored on a large hard disc for later processing.

When all the individual papers are translated, and when the Proceedings' front and back materials are created and translated into PDF files, Acrobat Exchange is used to add links, notes, and movies to the files. This process assembles an electronic "book" from the authors' individual electronic "chapters." The additional content added to the book includes PostIt notes with copyright information and some overall directions, movie clips as they are provided by the authors, and navigation support through a large number of links between the Table of Contents and the Author Index and the actual papers in the electronic proceedings.

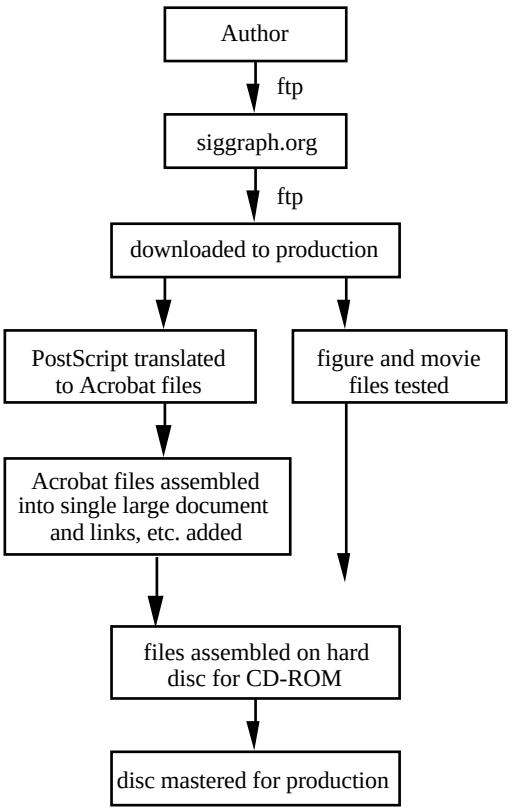

Figure 6: Information flow in SIGGRAPH's CD-ROM production process

The Acrobat reader displays the electronic proceedings in an "electronic paper" format, since the screen display duplicates the printed pages in the traditional printed Proceedings. This has both good and bad aspects; it allows the reader to print copies of each paper easily (on either color or monochrome printers), but the layout of the pages means that the reader must scroll to read the entire page.

But the electronic proceedings includes much more than just this document. The disc also includes software such as the Acrobat reader (for many platforms) and QuickTime players; image files for most of the color (and some of the line art and halftone) figures in the papers; movies in QuickTime or MPEG format from many of the papers; and additional material such as source code for graphics systems and datasets for some of the models or experimental results in the papers. Some authors have also included their text and images in a HTML file set for Web browser use.

Acrobat capabilities

Acrobat documents can have quite extensive electronic functionality, such as both local and Web links, text notes that can be opened by the user, and embedded movies and sound. For example, a user can:

- Select either an outline or a thumbnail view to accompany a document; outlines and thumbnails provide a good facility for browsing.
- Use article following to read a single article as it moves down separate columns or onto extra pages of the publication.

- Zoom into the text or figures (or anything else that has very fine detail).
- Navigate pages sequentially or retrace the sequence of pages (the history) that have been viewed.
- Execute a link to move around in the document or to view an accompanying document.
- Print any part of the publication to either a color or monochrome printer.
- Select text from the document to be copied and pasted into another document.
- Do an intelligent word search on the document or on a set of documents.
- Play a movie that is in-line in the document.
- Execute a link on the Web to an online document.
- Use indexes to speed up word search (if the user is reading the document with the Exchange application); such searches are intelligent and can be based on word stems and homonyms, as well as on full word matching

Acrobat does an excellent job of maintaining the graphic design of the original document, since it presents very high-quality text and diagrams. Its use of JPEG image compression for color images represents a good tradeoff between quality and storage needs.

Design opportunities

The simplest way to create Acrobat documents is to take documents that were originally designed for print, and translate them into PDF format to be played by the Acrobat reader. (This is sometimes called "reporposing" a publication.) The Acrobat reader supports this with page navigation buttons described below.

When documents are designed primarily to be used on the screen, however, the questions change from those of designing for print-oriented pages. Now the view is wider than it is high (landscape) instead of higher than it is wide (portrait) and the designer must wrestle with presenting the material on screens that vary in size, in resolution, and in color values and capabilities.

These design questions are well-known for Web designers, but they are also important for designers who are creating Acrobat documents. Fortunately, Acrobat's PostScript heritage eliminates many of the most difficult design problems in HTML authoring. For example, because of Acrobat's ability to use output from strong design tools, it is relatively easy to create documents that fit standard screens and move a page at a time. This is strongly recommended for screen-only documents but is more difficult to do for HTML documents.

In fact, Acrobat offers an important alternative to HTML for publishing on the Web. By setting an opening view that suppresses the standard Acrobat reader button bars, and by using link graphics and similar visual navigation vocabularies, a designer can create an Acrobat document that, when played via the Web, looks almost exactly like any other Web document except that it allows the sophisticated design made possible by PostScript-based design. People reading these notes are urged to try it out — it really works!

Navigation options and facilitation

Let's look at how the Acrobat Reader works and what tools it provides for the person using an Acrobat document.

When the user opens an Acrobat document, the document's first page is presented. This presentation includes the page's contents, along with whatever notes and links were provided for it. The presentation can also include either thumbnail or outline views of the document. The reader can move freely along the document's pages, as well as along the history of his or her use of the document, with a set of buttons (first, previous, next, last, previous visited, next visited) for that purpose.

Acrobat also has a linking ability within a single document or among documents. A region within a page can be designated a "hot area" and a click in that region will cause the reader to display a new page or a new document, which is specified when the link is created. The combination of including a button-like graphic in the layout, and making that graphic a hot area that jumps the user to a new place, gives an Acrobat document a full hyperlink capability. Because the link can include not only a target but a presentation, this goes well beyond the linking capability of HTML and Web browsers.

Acrobat's other navigation capability is its searching ability, available from the icon window bar. This takes advantage of sophisticated text matching (for homonyms, misspellings, word stems, etc.,) and is based on document indices that can be searched very quickly and can include many documents. It is possible to search quickly through large document libraries, including several discs, and then open the document that was found.

Acrobat distribution

Once Acrobat documents are created and fully prepared, they can be distributed in many ways. Acrobat is a cross-platform document system and its original release supported document production and reading on Macintosh, Windows, and a number of Unix systems. Acrobat documents are easy to e-mail or share by disk, and the Distiller provides enough compression that even complex documents are of modest size.

From the reader's point of view, Acrobat delivers a very clean representation of the document. It preserves the original document design very well, gives good color representation, and offers a set of controls for document viewing that are very easy to understand. The "look and feel" of the Acrobat Reader are identical on all the platforms it supports.

Issues with PostScript sources

As we noted above, it is important to have good, clean PostScript sources to produce good Acrobat files. We had two ongoing problems with SIGGRAPH papers that may be instructive to mention here.

First, we found that a fairly large number of people use tools that create bitmapped versions of fonts instead of higher-quality outline fonts (that is, Type 3 fonts instead of Type 1 fonts.) This typically come from using Computer Modern with a TeX-family system. This produces a bitmapped page for Acrobat, and such a

page is very slow to display and very difficult to read except at very high magnification, though it prints well. The solution is to force the tool to use a Type 1 font, and we were able to provide templates that showed how to accomplish this.

Second, we had the well-known problem of texts that were prepared for A4 paper but that we wanted to present on 8.5x11 templates. This was not a problem except when it was accompanied by a too-small bottom margin, typically caused by the tool forcing a larger than usual top margin. This is fixed by setting letter-size paper instead of A4 in the document. Presumably the opposite problem, of preparing a document for letter paper instead of A4 paper, could be fixed in the same way.

Delivering Your Publications to Your Audience

There are two primary technologies for delivering your publications to your audience: CD-ROM and computer networks. Each has its advantages and disadvantages, and one of the most interesting developments of the few months has been that both these delivery technologies can deliver both of the document technologies we discuss in these notes. A year ago, CD-ROM could deliver HTML publications but the networks could not deliver Acrobat publications; now this difference is about to disappear.

The nature of CD-ROM and networked publications are different in many ways. One of these is illustrated by the fundamental terminology of the two areas. Discs are *manufactured*, meaning that something fixed is created to be distributed. Publications on the networks are *put online*, meaning that a resource is made available. The distinction is that CD-ROM contents are fixed and these contents are sent out to persons who order them, while online publications are capable of being changed and the contents are available for persons who come and get them. The distinctions between fixed and changeable content and between active and passive distribution are the real differences between the two media.

Let us consider some of the implications of the fact that a CD-ROM is a physical artifact, while a document on the networks is more of an on-demand resource. With an artifact, you have something physical that can be ordered, paid for, kept, and re-used without having to use other resources such as hard disk drives to hold it. In fact, most commercial CD-ROMs cost around $50 to $100 per gigabyte, while a hard disk (as of June 1997) costs around $100 to $200 per gigabyte. The difference is that the contents of the CD-ROM are always available to you with no fear of losing it through a disk crash and without having to sacrifice anything else you would also want to keep available. If the content of a publication is seen as having long-term value, such as a reference, then the CD-ROM is the most appropriate for the publication.

The fixed content of CD-ROMs is familiar to publishers because it behaves just like the contents of a book or magazine. This allows an audience to have content that can be clearly referenced and that merits the time to be polished in its presentation. This is consistent with an editorial point of view that takes the time to consider the content of a publication and its relationship with past and current work in its field.

However, the fixed nature of a CD-ROM's contents may sometimes be seen as a disadvantage. Many areas, including technical fields, are constantly seeing new advances and you may want to have your publications incorporate these advances. If your publications are on-line, you may update them at any time and your audience will immediately see the new information when they next visit your online site. If it is important to your audience that they have the latest information, then online delivery would seem to be worthwhile for you.

There is also something of a cultural difference between people who want to have their information fully worked out in a polished form, and people who want to have the very latest information, even if it is rough and awkward. Sometimes this seems to be a difference between generations, although this is probably too casual a distinction. In any case, be careful to know your audience well enough to know if this difference could be important to them.

CD-ROM delivery

The capacity of a single CD-ROM is 654.7 megabytes (mb). This capacity may be shared among any blend of text, sound, pictures, video, and software needed for a title. The data transfer rate of the original CD-ROM drives was 150 kb/sec, but it is now standard to assume that a CD-ROM can deliver its contents at much higher rates. Conservative designers probably still design for a "2X" data rate of 300 kb/sec., but many people now believe that the "4X" rate of 600 kb/sec is a more appropriate standard. Higher rates are available, up to the DVD transfer rate of 11 MB/sec. This is important because it affects the kind of time-critical content, such as digital video, you can put on a disc.

Measuring only the capacity for raw information, such as megabytes of storage or data transfer rates, really does not convey a good feeling for the quantity of information a disc can hold, however. Consider an ordinary printed book such as one by the author, which was the basis for some experiments. A 300-page book with an average of 300 words per page and an average word length of 6 characters contains 1800 bytes per page or 540,000 bytes total. A single CD could store well over 1000 such books in raw text form. If we extend this book to include monochrome figures, the book can take about one MB in a standard word-processor format, so the disc could hold about 600 such books. The same book, fully formatted in Acrobat PDF form, takes about 2.4 MB, giving us a capacity of over such 250 books on a disc. Finally, the same book takes roughly 20 MB in full PostScript, allowing only about 30 books in PostScript. But this is still a remarkable number of books to hold in a small plastic disc!

To look at the capacity of a disc in another way, let us consider digital video. The amount of video that can be stored on a disc depends on the size of the video window, the number of frames per second, and the compression ratio. The single-speed CD-ROM standard for performance is flawless movies in 240x160

pixel frames. If the video is compressed so that it can run at 90 KB/second, then a minute of video consumes 5.4 MB of disc space so a disk can hold 120 minutes of video (and nothing else). A double-speed drive can play back a flawless 320x240 movie, but the video requires 18 MB of disk space per minute giving you a total of only 36 minutes of video. For higher-speed drives, the extra bandwidth for video continues on the same path of larger possible frames with shorter movies. Here, of course, is the driving force behind the higher-capacity DVD disc: storing the very large data sets represented by full-screen, full motion digital video requires truly massive storage.

The large capacity of a CD-ROM is accompanied by its being a very small and light physical object. It is easy to store many of these in a small space, and in fact one of the things SIGGRAPH's audience likes is that the SIGGRAPH conference proceedings discs are much smaller and lighter than the same proceedings as books. As we suggested above, CD-ROMs allow you to keep publications readily available without having to dedicate hard disk drives to holding them. Certainly maintaining a large electronic reference library on CD-ROM is more inexpensive and more durable than maintaining it on hard disks. In fact, one of the kinds of CD now available, the writeable CD-R disc, is widely used for backup storage.

One of the features of CD-ROM is its remarkable durability as a storage medium. A disc suffers no physical wear while it is being played and it is difficult to actually damage the reflective surface on which the information is stored. The polycarbonate plastic through which the laser passes is extraordinarily transparent and is not damaged by the laser or other light, including sunlight or ultraviolet light. In fact, the only normal way a disc is damaged is through scratching the disc surface; this is readily avoided by careful handling, and when this does happen we have heard of discs being repaired by polishing them with a jeweler's compound.

The costs of publishing on CD-ROM are quite favorable when compared with costs of traditional publishing. After all the various parts of the publications are prepared, they are placed into a single file system and all the linkages are tested to be sure they work correctly together. The files may then be sent to a plant that manufactures CD-ROMs to be mastered, or a pre-master may be created on a writeable CD-R disc for on-disc testing, which is strongly recommended. Testing can show up a number of errors, and there are standard methodologies for testing discs, just as there are for testing software. After the pre-master disc is fully tested and any corrections are made, a final CDR can be made and sent to the manufacturer. Costs for manufacturing may vary from country to country, but in the USA it costs under $2.00 per disc to manufacture a few thousand discs.

The recently-announced Digital Video Disc (DVD—now apparently called a "digital versatile disc" although that smacks too much of marketing and will not be used here) disc format is a truly challenging advance in CD-ROM technology. These discs will hold from 4.5 gigabytes up to some 17 gigabytes of data. While they are initially intended for digital movies for the entertainment market, there are publishing projects that will take advantage of this large capacity. It is too soon to say much about this because very few of the playback or manufacturing systems are yet out, but it will be interesting to see what impact this format has on CD-ROM; it will probably lead many people to buy DVD players and thus to add CD-ROM capability to their systems. It is also too soon to see just what the schedule will be for desktop writeable DVD systems (DVD-R) to reach the public, but it seems clear that we will see the first of these before the turn of the century.

Network delivery

Delivering publications through the networks is very attractive. This is a rather new option and many things about it are still uncertain, so it is difficult to make general statements about network delivery at this time.

Cost is one of the things about network distribution that makes it seem quite attractive. If both you and your audience are on networks that do not charge for the quantity of material sent or received, network delivery involves neither any cost nor any particular handling for either you or your audience to send or receive any particular document. If either of you are on what is called a "metered network service," then each document you send or receive could cost a small amount to send or receive; this amount is probably much less than postage for a single letter, but if you have no control over who receives your publications and if your publication is very popular, you could have fairly large costs for distribution.

Even if there is no cost for distributing an individual publication on the nets, though, there are other costs involved with using the networks for your distribution. One of these is the cost of setting up an online site, keeping it going, and keeping up a network connection. Many online publishing ventures are based in Universities where some part of the University takes care of maintaining the systems and the network connections, so it is not always clear just what these costs are. The actual costs will be different in different countries and at different times, but the fundamental costs can be readily defined:

- The purchase cost of an appropriate server,
- The purchase cost of additional equipment for the server, such as large hard discs and network connection hardware,
- The cost of maintaining, replacing, and updating the hardware, including the cost of a system manager,
- The initial cost of creating a network connection, and
- The ongoing cost of maintaining the network connection.

In addition to these costs, there is the cost of creating the publications you put online and the cost of maintaining current information on the site. Keeping online information current is critically important to having a successful Web site, particularly if you position your publishing enterprise around information that is time-dependent. This information maintenance will need one or more people with expertise in taking content from authors and preparing it for network delivery. In all, we are seeing surprisingly high costs for online publishing, so you should be careful to

investigate your own costs fully before undertaking such work.

With cost out of the way, we should look at the capabilities of the networks to support serious publishing. In what is being called the "good old days," about three or four years ago, it was possible to move fairly large amounts of information over the networks reasonably quickly. Transfer rates typically ran in the 50 kb/sec to 100 kb/sec range. Now, however, the popularity of the Internet has dramatically lowered the throughput we see on the networks, as we noted earlier in these notes, and a data transfer rate at my site rarely runs over 2 kb/sec and often is only a few hundred kb/sec. This certainly needs to be taken into consideration when you evaluate your delivery capabilities.

Hybrid delivery

Many times it may be productive to combine carefully-crafted archival information and up-to-the-minute timely information in a single publication. This kind of product takes advantage of the CD-ROM's large content and fast access, and of the ability of on-line publishing to stay fully current. In order to provide both features, publishers are beginning to look at the concept of a single product that has both disc and online parts. The technology for providing the off-disc links can be simple or complex, but in the technical publishing context of these notes it can be simple to include HTML publications on a CD-ROM that include links to any Web site in the world. When a user with Internet connections uses a Web browser to read the publications on the disc and then clicks on a link in a publication, that link will be executed in the same way whether the link is to a file on or off the disc. When a user without Internet connections uses the disc, they still have access to all the contents of the disc.

As an example, the author's recent book on electronic publishing includes a CD-ROM. This disc includes an outline of the book, linked to all the color figures in the book and to additional resources such as lists of resources and Frequently Asked Questions (FAQ) files. The disc has both on-disc and off-disc links, with the off-disc links going to the origins of the resource lists and to sources of updated FAQ files. The disc thus serves as both a primary resource and as a way to organize the access to updated primary resources as well as secondary resources for the book.

The hybrid approach is not without its hazards. Web sites quickly disappear if they are not maintained, and a CD-ROM which relies on a Web site that is not there is quite useless. CD-ROM is a stable technology, but the state of the Web is anything but stable; Web sites come and go with startling speed. If you add network connections to a distant Web site to a product then you create a real chance that your technology may fail. If you choose the hybrid approach, then you must be aware that you are committing yourself to provide a continuing service to your audience with all the costs of online services that we discussed above. Some publishers are committed to doing this, however. O'Reilly & Associates has created a Web site to accompany some of their books and understands that this site must be seen as a commitment to their audience that will support a growing number of publications. Note that the online site is not going to generate any direct revenue for the publisher; it must be seen as a part of the cost of document production.

CD-ROM Technology and Issues

Since compact discs were first introduced in 1980, the CD-ROM industry has grown at a rapid pace. One of the major reasons that CD-ROMs have so quickly become a low-cost and reliable vehicle for information exchange is the existence of stable standards and formats. CD-ROM standards define the format in which discs are written and organized, and determine the kinds of information that can be included on a disc of a particular format. This chapter describes the main standards and formats of interest to most CD-ROM authors and publishers. It discusses in some detail the major formats—ISO 9660 and HFS—in which most CD-ROMs are written today. It also describes the new DVD technology that is expected eventually to replace today's CD-ROMs.

Standards summary

The following list summarizes the types of CD-ROM formats and standards you'll hear about. The various books mentioned here (Red, Yellow, Orange, Green, White) are standards documents developed by the CD industry. Some have gained the status of International Standards Organization (ISO) documents. We don't describe the details of these books in this chapter (the details aren't usually relevant to CD-ROM publishers), but you can obtain them from ANSI at the address below if you are concerned about these details.

CD-DA: Compact Disc-Digital Audio. The standard used for encoding audio data onto a compact disc. (This standard is defined in the Red Book.)

CD-ROM: Compact Disc Read-Only Memory. The usual mode for data-only CD-ROM (that is, discs that contain only data and applications). (This standard is defined in the Yellow Book.)

CD-WO / CD-R: Compact Disc Write Once and Compact Disc Recordable. The standard for creating write-once compact discs that may be mastered on a PC. (This standard is defined in the Orange Book.)

CD-I: Compact Disc Interactive. A compact disc format in which computer data and compressed audio are interleaved on the same track. The format includes both a disc layout and an operating system, CD-RTOS, that can read the layout and play the disc contents. CD-I discs must be mastered on special proprietary systems. (This standard is defined in the Green Book.)

CD-Video: A disc format, also known as karaoke CD, that allows a combination of audio and full-motion video. It uses interleaved MPEG video and audio sectors to maximize the amount of information that can be stored on the disc. (This standard is defined in the White Book.)

The ultimate authority on CD-ROM standards are the standards documents themselves. The Red, Yellow, Orange, Green, and White Books, and the ISO 9660 standards documents, are available from:

American National Standards Institute
1430 Broadway
New York, NY 10018
212-642-4900

Several technical papers from Disc Manufacturing, Inc., (DMI) also provide information on disc standards and offer a deeper view of CD-ROM technology (particularly the ISO 9660 standard) than these notes can provide. See the DMI technical reports in the references for more details.

CD-ROM layout

Compact discs are written with one single spiral track of pits and lands and are read with constant linear velocity. That is, the disc spins at a higher speed when the track is read on the inner part of the disc than it does when the track is read on the outer part of the disc; the density of the pits along the track remains constant. Each sector along the track is 1/75 second long and contains a total of 2352 bytes of data, as defined by the Red Book standard. For the usual type of CD-ROM disc, Mode 1, this sector has the following information, in order:

12 bytes of sync
4 bytes of header
2048 bytes of data
4 bytes of EDC
8 bytes of blank
276 bytes of ECC

In addition, each sector contains 98 bytes of control information that contains several subcodes (although subcodes seem to be little used for standard CD-ROMs).

For some projects, primarily those involving a great deal of digital video, the needs of the projects far exceed the storage capacity provided by CD-ROMs. Recent work on both larger capacity discs and higher speed data transfer has led to the Digital Video Disc (DVD), announced in late 1995 but only beginning real production and distribution in 1997. We describe these discs at the end of this chapter, as well as the emerging ISO 13490 file system standard which will support them. Most people in the industry believe that, once players and manufacturing capability are available for the new DVD technology, the 13490 standard will probably become the standard for traditional CD-ROMs as well.

ISO 9660 standard

ISO 9660 (sometimes called simply "9660" in these notes) is the standard file system specification implemented on CD-ROMs. This specification was designed to be independent of any particular operating system. Thus, a disc formatted to the 9660 specification can be read on any system that supports 9660, including PC (DOS/Windows), Macintosh, and UNIX systems. Of all of the CD-ROM formats, 9660 is the most interoperable across platforms. You do pay a price for this interoperability, of course: a 9660 disc is the lowest common denominator for CD-ROM file systems, and may not have some functionality that you might wish to have.

If you have a PC running DOS and you want to allow access to a 9660 CD-ROM, your PC needs to run Microsoft Extensions. The Macintosh and most UNIX systems have special extensions built into their operating systems that allow them to access a 9660 disc.

ISO 9660 has three interchange levels; level 1 is the most strict and the one most often used, because it is supported by essentially all systems. These notes describe the requirements for ISO 9660 interchange level 1 CD-ROMs.

Actual data storage on CD-ROM

The digital data on a CD is not stored in the standard way we would expect. Instead, it is stored in a very specialized way called *eight-to-fourteen modulation* (EFM). The function of EFM is to spread the pits in a way that increases the accuracy of the disc reading process.

A CD-ROM is read by sampling the reflected laser beam in the player at a constant rate, and each sample provides a low or high reflectance signal. A 1 bit is signaled when the laser reflectance signal changes, either from a land to a pit or from a pit to a land. These changes cannot be too frequent; if they are, they can't be properly detected. The run length (the length of a sequence consisting of a 1 followed by zeros) must be at least 3 and cannot be more than 11. To accommodate this limitation for a single byte can take up to 14 bits, so that much space is allowed for each byte. In writing the disc master, or in playing back the disc, the 14-bit code is associated with its unique byte and the association between standard data formats and EFM formats are made.

File and directory names

There are a number of restrictions on the characters that may be used in file and directory names, and a number of issues arising from the ways that various platforms and applications deal with these names.

Valid names

File names and directory names on a CD-ROM must use only the characters A to Z (capitals only), the digits 0 to 9, and the underscore character '_'; these are called the *d-characters* in the ISO 9660 standard. You must also use the *8.3 format*—that is, file names must not exceed eight characters, and extensions must not exceed three characters. The file name and extension name must be separated by a period (i.e., the valid format is FFFFFFFF.EEE.

The following are all valid 8.3 names:

DOCUMENT.NEW
DOC.NEW
DOCUMENT.
DOCUMENT.1

Directory names must have no more than eight characters, and no extensions are allowed (i.e., the valid format is DDDDDDDD).

File name extensions

The extension on file names can be a source of difficulty with the ISO 9660 standard. Authoring

systems and operating systems differ in how they enforce restrictions on extensions:

- Some systems allow you to create 9660 discs in which files have no extension; other systems do not.
- Some systems that allow file names without extensions still require that the file name terminate with a period (.); other systems do not.
- Some operating systems' 9660 device drivers overlook a missing extension; other systems do not. For example, on some UNIX systems, such as the RS/6000 running AIX, the period must be included even if an extension is not specified.

The only truly safe approach seems to be to include the extension on *all* file names.

Version numbers

The standard ISO 9660 file naming convention adds a semicolon separator ';' and version number (e.g., '1') to each file name on the CD-ROM. Certain mastering software and platforms are more forgiving than others in enforcing this standard. This inconsistency can lead to problems in interpreting file names later on, especially when CD-ROMs are used on multiple platforms.

For example, suppose that you have a file on your PC named FILE.EXT. When you write that file to a 9660 CD-ROM, the file now appears with the name FILE.EXT;1 on the CD-ROM. However, when you view that CD-ROM on a PC running DOS or Windows, Microsoft Extensions strips off the version information, so your file name appears again as FILE.EXT. Programs needing to use the file will recognize it under the original name, FILE.EXT.

Things don't work quite as smoothly on the Macintosh and on many UNIX systems. On these platforms, the 9660 convention effectively changes the name of the file. When you view the same CD-ROM on those platforms, you will see the version information as an intrinsic part of the file name (for example, FILE.EXT;1). Programs needing to use the file will have to recognize the file name with its version information. If your program looks for the file named FILE.EXT, it won't be able to find it.

Many applications such as hypermedia systems require exact file names as link targets. You will find it very frustrating to carefully create 9660-compatible names and then have different systems change those names in different, and seemingly unpredictable, ways. What can you do about this problem? There are several approaches.

First, many disc mastering systems allow you to suppress file version numbers so you can make 9660 CD-ROMs without the ";1". In general, we recommend that you take advantage of this feature.

Second, if your system cannot suppress version numbers and you are able to target your application for a particular system, look carefully at how that system handles 9660 names and use that exact version of names in your application.

Third, if your application will work across different systems, you will have to keep the ";1" suffix in mind when you create a 9660 disc for in the Macintosh or UNIX world. One programming technique, which is awkward but effective, is to first try to open the file without the version number and, if that fails, try to open the file with the version number.

Unfortunately, the latter approach is not always possible to implement. It poses a significant problem if you want to put HTML documents on a 9660 CD-ROM because it's not possible to create HTML links that can make a "second try" if the first link fails. This would be the case if you searched for a name *with* a version number on a DOS system (since DOS strips off the version number before viewing it) or if you searched for a name *without* a version number on the Macintosh (since the Mac preserves the version number).

Your best overall approach for multiplatform CD-ROMs is to use only a disc mastering system that allows you to suppress 9660 version numbers. If you are developing a CD-ROM only for the Macintosh (which always shows version numbers for 9660 discs if they are present), you can avoid the problem by including an HFS directory on the disc. (We describe this approach later, in the discussion of "Hybrid discs.")

Upper- vs. lowercase

Case conversion is another interesting side effect of different implementations of the 9660 file system. Some systems automatically downshift all the d-characters (A-Z, 0-9, and _), converting all uppercase letters into lowercase, and some systems do not. As with the version number, case conversion has the effect of effectively changing the name of the file. Such a change will affect any program that needs to refer to the file.

MYFILE.DAT;1	MYFILE.DAT
myfile.dat;1	myfile.dat

Figure 7: File names produced by different 9660 implementations

Figure 7 shows the various names a file may have, depending upon how a particular disc mastering system handles version number and case conversion.

File name sorting problems

In some cases, you may be able to get away with specifying illegal characters in file names (e.g., lowercase letters or nonalphanumerics, such as z, a, b, !, #,$, etc.) . Certain software and platforms are more forgiving than others. But don't take advantage of such a situation. These illegal names will cause sorting problems later on your CD-ROM. When a sorting problem occurs, you will be able to see all of the files on your CD-ROM directory, but when you try to access them, the system will display the message, "file not found."

Directory depth

The depth of the directory structure must not exceed eight levels. This often presents a problem for the CD-ROM designer, because many electronic titles, particularly those with lots of content, end up having individual files organized into very deep directories.

Unfortunately, there doesn't seem to be any way around this restriction; if you are using the ISO 9660

standard, you will have to restructure your files so they have no more than eight levels of directories.

Volume labels

Volume labels or disc identifiers may contain as many as 32 characters, but only 11 characters are displayed in the DOS environment. The volume label is optional; you can ask the CD-ROM manufacturer to include one for you if you wish. (The volume identifier is displayed when a MS-DOS PC directory command is performed on a 9660 disc; it also shows up as the name of the generic CD-ROM icon on a disc mounted on a Macintosh.)

Number of files

In general, it's not a good idea to put more than 200 files per directory on a 9660 CD-ROM. Having too many files per directory can degrade disc performance.

To improve performance, you might consider putting files that are accessed continually or frequently at the beginning (inner) part of the CD-ROM.

Resource forks

Another issue arises when files from the Macintosh are written onto a CD-ROM for use on other platforms. On the Macintosh, every file has two components: a data fork and a resource fork. Resource forks do not exist on other platforms. When you process a Macintosh file on a Macintosh, there is, of course, no problem; the device driver expects to see a resource fork. But how do other platforms (or, more properly, these platforms' implementations of the 9660 device drivers) handle resource forks when they encounter them on a CD-ROM?

Most systems, such as those for DOS and Windows, and some versions of UNIX, correctly ignore the resource fork and simply use the data fork for the file. However, in the original BSD 4.2 version of UNIX and in other UNIX systems derived from it, the CD-ROM device driver ignores the data fork of a file and only provides the resource fork to a program. This isn't very helpful, because the resource fork typically does not hold the data that would be of use to non-Macintosh systems. As a consequence, such systems cannot use any file that contains a resource fork.

This situation basically means that you can't put fully-functional Macintosh files in the shared file area of a dual-directory disc. (We'll describe dual-directory discs later in this chapter.) Instead, you'll need to remove the resource fork for the version of the file being used for the 9660 file system. In addition, if you need to have full Macintosh capabilities for the file, you'll need to have a second copy of the file in the HFS directory with the resource fork intact.

Rock Ridge interchange protocol

The 9660 file format does not maintain the basic look and feel of the UNIX file system. The Rock Ridge Interchange Protocol was developed to help address this issue. This file system was designed to allow UNIX-like systems to retain much of the directory information in the native file system and maintain compatibility with the POSIX file system standards. At present, there does not seem to be much adoption of the Rock Ridge standard outside the UNIX world.

Compatibility with UNIX/POSIX standards is important because UNIX systems use directory entries for much more than just pointing to files. Directory entries can point to other entries (symbolic links) or to device drivers that are linked to peripheral devices such as hard disks, tape drives, and CD-ROM drives (device files). The entries also contain file permission information. The Rock Ridge Interchange Protocol retains these features. Of course, Rock Ridge discs also do not have the version number and character case problems with file names that we discussed for ISO 9660-standard CD-ROMs above.

Macintosh hierarchical file system (HFS)

The Macintosh file system is known as the Hierarchical File System (HFS). Most CD-ROM mastering systems allow you to create a CD-ROM in either ISO 9660 or HFS format. (Some offer additional choices, which we'll describe later in this chapter).

If you are developing a CD-ROM application that is going to be used only on the Macintosh, we recommend that you create an HFS CD-ROM. Doing so will allow you to create the typical Macintosh desktop look and feel on your CD-ROM, including the ability to launch an application or file by double-clicking. With an HFS disc, all of the icons and window layouts will show up on your disc just as they appeared on the disk from which your CD-ROM was created.

By contrast, if you create a 9660 CD-ROM and read it on a Macintosh, all you will see are generic folders and files, with names under them. Because the CD-ROM does not have a desktop database, generic folders and icons are placed on the desktop by the system. You have little control over how the icons are presented in the window.

Although you may be unhappy with the look and feel of the desktop displayed by a 9660 CD-ROM on the Macintosh, the Mac nevertheless will be able to read the CD-ROM correctly. You'll have to use the 9660 format unless your application calls for a Mac-only CD-ROM or unless you are able to create a dual-directory disc, as we describe later.

Displaying a CD-ROM icon

In general, when you mount a 9660 CD-ROM on the Macintosh, a generic CD-ROM icon will appear on the desktop. But there is a way that you can control the appearance of this icon. If you want a special icon to appear in place of the generic one, you can supply this icon to your CD manufacturer. When you send your product to the manufacturer, simply provide the icon with the disc image, as part of your backup media, or on a separate floppy. The name you give your CD-ROM can be any standard Macintosh name, containing up to 27 characters, with no limitations on the characters you choose.

Desktop and window layout

When you send your data to the CD manufacturer, be sure that you arrange all of your folders, icons, windows, and files exactly the way you them to appear on the CD-ROM. The exact layout is not important on a 9660 CD-ROM, but it is very important on an HFS CD-ROM because this layout determines the desktop database that controls what the user sees when he opens

the CD-ROM. (This is the case because the desktop database is saved to the disc.)

When you consider your layout, remember to look carefully at the location and sizes of all the windows that will ever appear on the desktop. If you don't, windows will open at random locations and their contents will be displayed randomly--neither of which instills confidence in your customers.

The usual procedure in building a CD-ROM is to have all of your windows open at the upper left-hand corner of the screen and to keep them modest in size, so none of the windows will disappear if the CD-ROM is used on a Macintosh with a small screen. It's also a good idea to offset each additional window by a tiny amount (perhaps one-eighth of an inch) when you cascade deeper into a hierarchy of displays. This way, as you open successive windows, you will continue to see a small portion of the previously opened window at the upper left corner.

On a Windows disc, you have to create the `setup.ini` file to handle icons, window locations, and installation management. Tools such as Install Shield can help with this process.

Hybrid Types of CD-ROMs

Suppose you need to create a CD-ROM that's going to be used on both a PC and a Macintosh. If you create the CD-ROM in ISO 9660 format, you'll be able to use it both under DOS/Windows and on the Mac; however, as we've mentioned, on the Mac you won't have the familiar look and feel that would be available with an HFS disc. Unfortunately, you can't create an HFS disc, or it won't be usable on the PC—or can you?

Actually, there is a way to get the best of both worlds, as we describe in the following sections.

Original hybrid CD-ROMs

Many mastering systems allow you to build a *hybrid CD-ROM*—one that contains both 9660 and HFS images on the same disc. Hybrid discs have two distinct sets of components:

- An Apple partition map for the HFS file system, the HFS disc directory, and the associated HFS image
- A 9660 disc directory and the associated 9660 image

You place the Apple partition map in the first sector of the disc; this works because the 9660 disc image doesn't begin until sector 16 on the CD-ROM. The Apple partition map specifies the address of the HFS disc directory, which in turn indicates the first sector of the HFS image--the image that begins after the 9660 image ends.

When you put a hybrid CD-ROM into a Macintosh, it reads the first sector on the disc, finds the Apple partition map, and then goes directly to the HFS image, completely skipping over the 9660 image. On the other hand, when you put this CD-ROM into a DOS PC, the CD-ROM drive doesn't start reading until sector 16; because it skips over the Apple partition map, the system sees only the 9660 image.

The original kind of hybrid CD-ROM described here totally separates the Macintosh files from the 9660 files, so each of the HFS and 9660 directories gives access to a distinct set of files. If a file is to be used by both systems, it must be included twice on the CD-ROM. The effect of this hybrid design is that each of the HFS and 9660 images can only be as large as *half* the maximum size of the CD-ROM. On most hybrid discs, HFS and 9660 images end up being about the same size, which may cause problems if you have a substantial amount of content.

Dual-directory CD-ROMs

A newer type of hybrid CD-ROM has been developed to get around the space limitation of the traditional hybrid discs. With dual-directory CD-ROMs, you can indicate that certain files on the disc are common to both the 9660 and the HFS images by allowing both directories to indicate these files. A dual-directory CD-ROM effectively allows each individual image (9660 and HFS) to be much larger than on the traditional hybrid disc. The available space for data is decreased only by the amount of the extra directory—about 15 megabytes.

Every file need not be shared. A dual-directory disc may still have individual partitions for 9660-only and HFS-only files. This is necessary because an individual file, such as executable code, may not be appropriate on both partitions. Most disc mastering systems allow you to specify that a file is to appear in only one of the HFS or 9660 sides.

Producing a dual-directory CD-ROM is not altogether straightforward, and may introduce some subtleties into your file structure. Here are some things you need to watch our for:

Keep files in proper directories

If some files are intended for use only in the HFS partition on the disc, or only in the 9660 partition, be sure to include those files only in the directory for that partition.

Resource forks

If files are to shared, be sure that they are readable by applications on both Macintosh and non-Macintosh systems. The main complication in this area is with the resource fork that is part of many Macintosh files. To ensure complete compatibility, you must remove this resource fork from any file that came from Macintosh sources. The resulting files will probably not have custom icons and may not support a standard double-click launch, but this is the price you must pay for the ability to share the file between the two directories.

End-of-line characters

The PC, the Macintosh, and UNIX systems all handle end-of-line in a slightly different way:

- The PC terminates each line with both a carriage return and a line feed.
- The Macintosh terminates each line with a carriage return.
- UNIX systems terminate each line with a line feed.

If text files, such as READ_ME files for the CD-ROM, are being shared by all of the platforms, you must take these rules into account. Often, developers include different versions of these files; sometimes a least common denominator approach is used, adopting the PC convention for all text files. Another approach is to

include a browser, editor, or script on the CD-ROM that performs the necessary translation of lines.

Digital Video Discs (DVDs)

The Digital Video Disc (DVD) is a new format created from at least two competing laboratory developments in the early 1990s. This format is expected to be very important to the future of digital video. Steven Spielberg has been quoted as saying that CD-ROM is the future of the moving image; it seems clear that he is really referring to DVD video. We anticipate that we'll see a full spectrum of DVD capabilities, just as we have for CD-ROM. This section looks briefly at the new DVD format.

DVD-ROM is the data equivalent of the DVD video disc. *DVD-R* is the writeable version of the DVD disc; that version is expected to have a smaller capability than molded discs.

Physical characteristics

The DVD standard allows for four different configurations, providing four different disc capacities. The configurations are based on single or double layers on a side, and either single or double-sided discs:

- The single-layer, single-sided disc is very much like the current CD-ROM, but it achieves 4.7 GB capacity (about eight times the capacity of a CD-ROM) by using higher pit densities on the disc.
- The double-layer disc uses a semi-transmissive reflection coating on the outside layer, with a fully reflective coating on the inside layer. (See Chapter 10 for details on the reflection coating on a disc.). This configuration achieves a capacity of 8.5 GB. The double-layer disc is read by using a laser that can switch between reading the two layers, taking advantage of the fact that the two layers will have reflections at two different positions.
- The double-sided, double-layer disc will have a capacity of 17 GB, which seems enormous to us now.

With some retooling, CD manufacturers are able to manufacture DVD discs. The manufacturing cost for DVD discs are likely to settle into something in line with current manufacturing costs (taking into account the extra work of multiple layers and lamination).

Use for video

When the DVD standard is used for pure video (for example, as a distribution medium for movies or broadcast television), they will use MPEG-2 video encoding and Dolby Surround Sound. These technologies will provide more than two hours of video on the initial single-sided, single-layer disc. As a bonus, the publisher will be able to provide multiple language tracks, multiple subtitle tracks, closed captioning, and the usual stop- and slow-motion, reverse, and search/scan options.

Use for multimedia

When the DVD standard is used for digital data and multimedia, the new discs will allow authors to use whatever type, size, and rate of media they feel necessary, up to the standard data transfer rate of about 11 MB per second. DVD recorders and blank media are expected to be on the market fairly soon (at least for the single-layer, single-sided version), but it's not clear when the technology's prices will drop to desktop levels.

Before you become too entranced by DVD opportunities, remember the implications of larger capacity. There are costs for content licensing and development, and increasing the content by a factor of eight (or more) is likely to lead to significantly increased costs in developing electronic titles. We also question the penetration rates that some developers have projected for DVD-capable players. Real market penetration has been delayed because of copyright issues, but it is starting to happen. It seems fairly safe to say that there will be enough DVD players by sometime in 1998 to start making large-market DVD products worthwhile, and it should be possible to build technical titles by the turn of the century. Mastering your own DVD-R discs may still be very expensive then, but creating a disc image on a large hard disc will work, and the manufacturing capability is now being created by the industry.

Of course, besides the pure multimedia kind of content we imply here, DVD discs will allow a publisher to build even more extensive hybrid projects as described above, where the fixed media content is significantly larger than is possible with CD-ROM capacities.

ISO 13490 Standard

We noted earlier in this chapter some of the inadequacies of the 9660 CD-ROM standard. These inadequacies became critical when the new DVD disc was first introduced. In response, a new standard from the ISO, known as the 13490 standard, has emerged.

We expect that the ISO 13490 standard will eventually replace the 9660 CD-ROM standard. This new standard allows such features (missing from 9660) as the following:

- Long file names
- POSIX attributes
- Unicode conventions to support file names in non-Roman characters
- CD-WO and multi-session discs.

CD-ROM Mastering and Manufacturing

Once you've finished developing the content for your CD-ROM, and have fully tested both your data and applications, you're ready to send the data to your CD-ROM manufacturing plant. The manufacturing process includes several main processes, primarily mastering and replication. *Mastering* refers to the recording of a glass master and the eventual creation of a metal stamper used to replicate your discs. *Replication* refers to the steps of manufacturing and packaging the discs for distribution to your customers.

This section provides an overview of the details of CD-ROM production. It discusses the need for an overall project timeline, includes details of many aspects of CD-ROM production (e.g., mastering,

replicating, packing, labeling, and shipping), and describes a variety of options available to you during this process.

You may wonder how the emerging Digital Video Discs (DVDs) relate to the CD-ROM discussion. DVDs are expected to use almost identical procedures for sending data to the manufacturer, for mastering, and for manufacturing the individual discs. For some time, it will be more difficult for developers to produce write-once masters, because DVD-density disc writers will initially be quite expensive; we expect that disc images will probably be put on tape. We also expect there to be some differences in metalizing the outer discs for laminated products, because this metalization will need to have half-silvered properties. This implies that there will be laminating steps in manufacturing, but this should not present a problem because disc manufacturers are already preparing for these discs.

Costs

The cost of manufacturing the finished CD-ROM artifact is relatively low. Earlier in the notes we gave an overall estimate of $2.00 per disc, but it is worthwhile to break this down so you can see its various parts.

The first part of the cost is the cost of packaging that accompanies the disc. This can be a jewel box, a cardboard sleeve, or any of many other choices, but almost all of these choices involve printing, either on inserts for the jewel box or whatever printing the package will take. In turn, this requires artwork, separations, and printing costs, and can usually be figured at around $.50 to $1.00 per package. The package itself may run a small additional cost, or it may be bundled with manufacturing costs.

The rest of the cost is in the actual manufacturing of the discs. These costs will vary a bit between manufacturers but are mostly determined by your volume, the turn you request, and the way you provide your data to the manufacturer. For a modest run of a few thousand discs with a leisurely turn of a week or more, with a fully laid out design for printing on the disc, and for data that is given to the manufacturer as a complete disc image, you should count on costs of $1.00 to $1.30 per disc.

So the actual manufacturing costs of discs is quite modest for standard jobs. Of course, you can go off the deep end on your packaging or your disc printing and drive these costs up. We recommend very strongly getting current quotes from disc manufacturers because the figures above are very general and can only serve as guidelines.

CD-ROM production timelines

The overall process of producing a CD-ROM breaks down into a number of distinct steps. Table W lists these steps, in last-to-first order and provides estimated times for each; the result is a timeline for producing a complete disc title (one that has already been developed).

There are some variations in the time required for some of the steps shown in the table, and there are tradeoffs. A good example of such a tradeoff is the disc turnaround time at the duplication facility; if you really need to have your finished disc quickly, you can usually pay more to get quicker turnaround.

Activity	Time Required
Ship discs from duplication	1 week
Master and duplicate discs	1-2 weeks
Print disc inserts	1-2 weeks
Test premastered discs	2 weeks
Assemble documents for disc	3 weeks or more

Figure 8: Estimated CD-ROM production schedule

The “time required” around which we constructed Figure 8 is the time normally expected for each activity in the production process. Some activities, such as testing time and document preparation time, are based on straightforward technical publication projects, and may not exactly fit your own projects. However, your experience with your kind of content should help you know how much time to allocate for your own tasks. From this information, adapted to your own circumstances, you can readily calculate how long before your required disc delivery date you must begin each activity.

CD-ROM manufacturers can provide you with their own checklists for disc production, and you should obtain one from your own manufacturer early in the process; you’ll find that most of these checklists look a good deal like the one we’ve provided.

Liaison

Most CD-ROM manufacturers prefer that you designate one person to act as project manager, or liaison, for disc production. This person will be the contact point for your disc development and duplication contractor, and will deal with tasks such as these:

- Providing the disc inserts and disc labels to the manufacturer
- Getting the input data to the duplicator (fully tested and in plenty of time)
- Dealing with any logistics, such as paperwork, scheduling your time for duplication, arranging credit or payment, and arranging for shipping the finished discs to you

The “turn”

As we describe later in this chapter, there are a number of steps involved in creating the disc image, making the disc stampers, and actually pressing the discs. Because it takes extra effort to perform such steps on a rush basis, most manufacturers charge somewhat more for a rush job than for one that can be scheduled as part of their normal work.

The term for the amount of time allowed for the job is the *turn* on the project; the turn is measured in working days, and throughout this book we generally assume a relatively leisurely 10-day turn. Thus, a CD-ROM with a 10-day turn is shipped from the manufacturer two weeks from the day the master gets to the manufacturer. The turn on a project can be less than 10 days, of course, but the shorter the turn, the higher the cost.

The turn isn’t the only scheduling factor. It’s not always possible to schedule a disc at your convenience. At peak times, the manufacturer may be booked fully and may not be able to schedule a disc at all--no matter

what the turn. For example, in the fall when discs intended for Christmas sales are being manufactured, or when there is a major software release such as Windows 95 being manufactured, most manufacturers' schedules are completely full. Be sure to give your manufacturer all the notice you can so that you can get a place in their production schedule.

Sending Data to the Manufacturer

In this section, we assume that you have created your disc contents and have tested them for correctness and functionality. At this point in the process, you will usually have a hard drive with data from your DOS/Windows, Macintosh, and/or UNIX system on it.

What happens next depends on whether you have access to CD-ROM premastering software.

If you have premastering software...

If you have access to premastering software, along with a tape drive or a CD recorder, you can generate your own CD image that can be sent directly to the CD-ROM manufacturing plant for replication. Preparing your own premastered disc is the safest way to send data to a CD-ROM manufacturing plant, and is highly recommended. There are two reasons:

- Assurance. Having a CD recorder gives you the security of being able to test your application on your own disc before replicating hundreds or thousands more of them.
- Economy. The CD plant can master directly from the CD-R or 8mm tape you send them. This saves you the manufacturer's charges for generating a disc image or CD-R for you.

Premastering has become fairly straightforward, thanks to the availability of software tools that do most of the work for you. These tools allow you to specify:

- What files are to go onto the disc
- How these files are to be organized into directories
- Whether these files are to be visible to the user
- If you are creating a hybrid disc (described in Chapter 8), which file systems are to include each of the files.

The tools then create a disc image that can be written to CD-R or DAT media for testing and for delivery to your manufacturer.

An example of premastering

This section contains a brief example of the premastering process. We describe how you specify the contents of a CD-ROM to create the disc image for disc premastering.

Among your options when selecting a file, you can specify whether the file name should have a version number, whether the file should appear on the HFS and/or 9660 partition, what creation date should be shown for the file, and whether the file should be invisible to the user. (You might make the file invisible to protect some of the contents of your CD-ROM from being copied to a user's own system. If you make a file invisible, it's much more difficult for a user to recognize what's on the disc and copy it.)

Once the files have all been specified, you can create a disc image, and then mount and execute the disc image as if it were an actual disc. Doing so allows you to test performance and general disc characteristics before creating the actual CD-R disc. In fact, all testing can be done from the image, though it is more convenient to work from a CD-R disc and you must create a CD-R if others are to help test the disc. Once the image is finally created, it can be written to 8mm tape, DAT, or a CD-R disc to undergo additional testing and, eventually, to be sent to the manufacturer.

How a CD recorder works in premastering

The CD recorder works with a special kind of disc that contains a gold reflective surface and a green photosensitive dye layer. The laser writing light is driven by the premastering software and is pulsed according to the data to be recorded. When the laser light hits the dye layer, it erodes the dye and creates a surface whose reflection property matches that of a pit. The pits simulate the surface of a standard CD-ROM, and the resulting CD-R disc can be played on standard CD-ROM players.

If you don't have premastering software...

If you don't have CD premastering software, you will need to send the data from your PC, Macintosh, and/or UNIX system to the CD-ROM manufacturing plant on backup media--generally, some type of tape, magneto-optical (MO) disk, hard drive, or the like. Once the backup is received, the manufacturer will restore your original data and generate a CD image for you.

Although it isn't strictly necessary, at this point we strongly recommend that you have the manufacturer generate a CD-R or "proof disc" (gold disc) for you to check. (This is part of the testing process we've described above, aimed at helping you ensure that your application is performing as desired, with no bugs.) Once you've approved the proof disc, the manufacturer can master and then replicate your CD-ROMs. A proof disc usually costs around two hundred dollars, and is a very small price to pay to make sure that your application works correctly from disc before you spend several hundreds or thousands of dollars to master and replicate your image. A CD-ROM plant is like a large copy machine. If you send in a bad copy, you will not only get bad copies back, but you will have to pay for them. That is why it is so important to see a proof disc before the final copies are made.

Data input formats

As we've mentioned, there are two methods for sending data to a CD manufacturing plant for replication:

- If you have premastering software, you'll send the data as a CD image, ready for mastering. You'll normally put the image file on one of the following media: CD-R, MO (magneto-optical) disc, DAT tape, or backup tape (8mm, 4mm, or 9-track).
- If you don't have premastering software, you'll send the data in a nonimage format, putting the data on a hard disk (SCSI, IDE, etc.) backup, MO disc, Syquest or Bernoulli, cartridge, DAT tape, or backup tape (8mm, 4mm tape, or 9-track). The data then needs to be formatted or converted to an image format.

Image format

A CD data image is a copy of the exact data that will eventually be placed on the CD-ROM, including all the directory information and all the extra information in each data sector. The best way to send your input to the CD plant to be mastered and replicated is on a CD-R disc; the plant can master directly from a CD-R.

If you aren't able to put the data on a CD-R, the next best choice is probably an MO disc. A magnetic tape is also acceptable; in general, DAT or 8mm is preferred. If you send a disc image on 9-track tape, it will have to be converted to 8mm for mastering. This conversion is safe and easy, but takes time to do and there will usually be a conversion charge for the work.

When you send a tape to a CD-ROM manufacturer, follow these rules:

- Make sure that tape labels conform to the ANSI X3.27-1978 standard (recommended) or to IBM standard practice. If the labels do not conform, the tape may be unlabeled.
- Make sure that your tape has an adequate block size. For 9-track tapes, you should have a record size of 2048 bytes and a block size of 8192 bytes. A recording density of 1600 or 6250 bpi is acceptable for 9-track tapes. For 8mm tapes, you should have a record size of 2048 bytes, and your block size should be between 8192 and 16384 bytes; 16384 is recommended because it results in faster processing. These block sizes are compatible with the 2048-byte data blocks of the standard Mode 1 CD-ROMs.
- When you send a CD image on a tape, indicate what type of label is on the tape, the record and block size of the tape data, and the size of the image (number of sectors) on the tape. It is also a good idea to indicate the type of data or application on the tape. In addition, always indicate the number of 2048-byte blocks (2K sectors) that make up the image.
- When you send any type of media to a CD plant, send two copies. Having a backup copy is useful if the original copy is bad or is damaged in transit.

Nonimage format

If you don't have premastering software and aren't able to create a CD-ROM image, most CD manufacturers will accept your data in nonimage format. You can put this data on almost any media, from DOS backups on hard disks or streaming tape to PKZip files on floppies. The non-image data will require formatting and media conversion to create a correct CD-ROM image format. The CD-ROM manufacturer usually charges a fee for formatting and conversion.

Sending data on various devices

There are several ways that customers can send nonimage data, in the form of a set of directories or files, to a CD manufacturer. The following sections summarize the more common methods of sending such a backup for each platform, along with other issues, recommendations, and comments that are associated with each method.

Hard disk drive (SCSI, IDE, or other)

Sending your data on a hard drive is not a good way to get your data to a CD-ROM manufacturer. There are several reasons:

- There can be problems getting your drive configured to match the plant's PC systems.
- Your drive might get damaged during transit.
- You will not have the use of your drive while it is at the manufacturer.

But if a hard drive is the only way you have to transport your data, it will get the job done.

It's more common for a disc duplicator to have problems with the configuration of a disc from a PC or UNIX system than from a Macintosh; there are rarely any configuration problems when you send in a SCSI drive from a Mac. On the Mac, when the drive comes up on the system, you'll see one partition that has all the data in the exact layout you intend to use on the CD-ROM.

Sending in a hard drive from a DOS or Windows system can be more complicated. The most common types of DOS drives are SCSI or IDE. If you send in a SCSI drive, be sure to let the CD manufacturer know what controller you were using to run the drive (Adaptec, Future Domain, etc.). If you are not using a common controller or drivers for your drive, you might need to send your controller and/or drivers in along with the drive.

Always call the manufacturer before sending in your drive to see what types of drives, controllers, and drivers they support; doing so will avoid unnecessary delays in processing your data.

If you send in an IDE drive, be sure to send the drive's documentation (setup information) and any special drivers the CD-ROM manufacturer will need to set up the drive.

Magneto-optical (MO) disc

MO discs are more problematic for PCs than for Macintosh systems. Macintosh MO discs are usually interchangeable among different Macs, and appear as ordinary hard drives to the computer. However, when you send an MO disc from a DOS or Windows system, you must let the CD manufacturer know what type of magneto-optical drive, SCSI controller, and drivers you used. If the CD-ROM manufacturer does not have compatible drives, controllers, or drivers, you might need to send yours in along with your cartridge. Be sure to call the CD manufacturer ahead of time to find out what kinds of drives they support.

Syquest and Bernoulli cartridges

Syquest and Bernoulli cartridges are usually interchangeable between different Macintosh systems and even between most DOS and Windows systems. However, if you are sending in DOS data, be sure to let the CD manufacturer know what type of controller and drivers were used with the disks.

Backup tapes

Backup tapes are a good solution if you need to send a large amount of data to the CD-ROM manufacturer. If you are sending Macintosh data to a manufacturer, we recommend the use of the Retrospect tape backup

software writing to 4mm or 8mm tape. For DOS or Windows systems, there are several different backup programs available, including Colorado, Sytos, Novaback, and Secure.

As with other types of devices, it's a good idea to call the CD-ROM manufacturer ahead of time to find out which programs they support, and what type of tape drives they have. If the manufacturer doesn't have tape systems compatible with the one you are using, you might need to send in your software, controller, or drive so the manufacturer can successfully read and restore your data.

In the UNIX world, tar (tape archive) backups are the most popular method. Most tar backups are made on 4mm or 8mm tapes. When you send a tar tape to a CD-ROM manufacturer, be sure to indicate the block size used to create the tape; 2048-byte blocks are preferred since they are compatible with CD-ROM block sizes.

Floppy diskettes

Use floppy diskettes only to send small amounts of data to the CD manufacturer or to update files for existing CDs. Sending a large amount of data on many floppy diskettes will cost you a lot of time, and probably a lot of money, for the additional work required by the manufacturer.

Packaging and Printing

So far in this chapter, we've discussed only the CD-ROM disc itself. But a disc is not the only component of a CD-ROM product. A complete CD-ROM package includes several components:

- The disc itself
- The jewel box or other package that houses the CD
- The label that is printed on the CD (which must be prepared as graphical artwork)
- The insert--the printed card stock or brochure that goes in the front of the case cover
- The tray card--the printed card stock that goes in the back side of the case and includes the spine
- Shrinkwrap or polywrap
- Possibly, a box that houses the CD and displays the product on the store shelf; the box may also contain a more extensive user manual, sometimes a warranty card, and sometimes advertising for other products

All of these pieces present rather specialized design issues. Remember that your CD-ROM packaging and associated printed pieces are graphic products. You'll need a professional graphic designer to do the design and make the mechanicals for these various pieces.

Your packaging may be the main way a potential customer first encounters your product. You should view your packaging as a major marketing tool, and be willing to consider part of your packaging cost as a marketing expense. Jewel boxes—the familiar packaging for music CDs—are very common and make a nice package. For some markets, though, you may want to consider cardboard sleeves, packaging your disc in a sleeve inside a printed piece, creating a shelf box to display the publication, using a self-mailing package, or using a magazine-style package. Your CD-ROM manufacturer can help you explore your options and their prices. Leave enough time--and enough money in your budget--to develop and produce appropriate designs to match your product.

Disc inserts

The printed disc inserts for jewel box packaging--or any other printing required by your choice of packaging--are standard printed pieces. As such, they can use any traditional printing process. Your CD-ROM manufacturer will have detailed requirements for the dimensions and paper of these pieces. Be sure to get the exact disc insert specifications from your own CD-ROM manufacturer before you do any production. Tolerances for packaging pieces will vary between manufacturers.

Be aware that the print pieces must be produced and delivered to the production site before disc replication can begin. You may find it difficult to believe, but the printed inserts may cost more to manufacture than the discs themselves!

Disc printing

The image on the surface of the disc itself can be produced by either a silk-screen or offset process Because the image is printed on plastic, not paper, this process presents special issues.

You must prepare the contents of this disc graphic yourself, using guidelines provided by the replicator. Typically, you'll provide the replicator with a camera-ready film negative that is exactly what goes on the disk. The cost of replication depends heavily on the number of colors in the image. Most replicators provide a two-color disc image as part of the standard stamping price, and then charge extra for each additional color (where four colors provide a full color capability). If you want two colors, you will have to prepare two-color artwork and deliver two film separations. Remember that because the background is already a metallic silver color, you actually have silver plus two colors to work with.

The following sections summarize characteristics of silk-screen and offset printing. More details on requirements and specifications for these methods of printing can be obtained from your CD manufacturer.

Silk-screen printing

Most CD-ROM manufacturers print the labels on discs using a silk-screen process. Screen printing can include any number of colors, but most discs use no more than five different colors. More colors, of course, means more cost. Different images are printed onto the disc surface, one color at a time, by placing a screen over the disc surface and pressing ink through the openings. (This process is shown in Figure 21 below, in the discussion of disc manufacturing.) The ink is cured or dried using ultraviolet light.

You will need to send the artwork used to create the screens used in printing to the CD-ROM manufacturer in the form of film positives, emulsion side up. Be sure to make the film from original art, not multigeneration copies, and check that the film is free from scratches. As with offset printing (described below), a white background is recommended for discs that use several colors; that background enhances the colors that are used.

Offset printing

A few CD-ROM manufacturing plants offer the service of offset, as well as silk-screen, printing on a disc. Offset printing allows a photographic image to be printed on the CD. It is done from a metal plate of the image you would like placed on your CD, just as it would be for offset printing on paper. The CMYK primary colors can create almost any color needed for the desired image.

It's advisable to include a film positive matching the separation diameters to create a solid white background that is screen-printed onto the disc. The solid white background is highly recommended because it allows clearer color definition in the image that will be printed over the background. In addition, include a film negative, emulsion side down, for each of the separations that will be printed by this offset process on top of the white background.

Disc Packaging

What kind of packaging do you want for your CD-ROM? You'll need to figure out how you want to present your disc to the customer. The so-called "jewel box" packaging it is familiar from music CDs and is very common for CD-ROMs. There are several other packaging options, however, as we discuss in this section. Talk with your CD-ROM manufacturer ahead of time about all of these options.

Jewel boxes

Many CD-ROM titles come in jewel box packaging, accompanied by shrink wrapping. This is certainly a popular option for CD-ROMs, as it is for music CDs, and is usually easy to arrange with your disc manufacturer. There are a number of types of jewel box packaging:

- The traditional jewel box
- The traditional jewel box, with shrink wrapping
- Double jewel boxes (for titles that need two discs)
- A double jewel box that is the same thickness as the usual single box but whose central piece is hinged so it holds one disc on each side.

Jewel boxes can also be inserted in other boxes or in plastic display pieces to make them more attractive for point-of-sale display.

Shrink-wrapping (sometimes called polywrapping) is included at no extra cost (beyond jewel box packaging) by some manufacturers; others charge a separate fee for it. Look carefully at the bids submitted by prospective manufacturers; be sure to tell them whether you want shrink-wrapping, and ask if it will add more time to the process.

Other packaging

Other choices for CD-ROM packaging range from no packaging at all to very elaborate disc presentations. For example:

- Discs can be shipped to you in a bulk package with no individual disc packaging (this package is called a spindle) so that you can add your own packaging, put the disc into a blister pack in a book, or package it with other materials.
- Discs can be inserted into several different types of envelope packages to reduce packaging volume or mailing costs.
- Some types of full packages have less plastic than jewel boxes and use cardboard enclosures; these produce less packaging waste and can include self-mailing packages. (Note, though, that the printing issues are quite different from those of jewel boxes.)

If you are going to use jewel box packaging—or any other kind of packaging for which the discs are put into the packaging at the manufacturing plant—make sure that the inserts are received at the plant before the disc duplication is scheduled to begin. If the CD-ROM will be put into your packaging at another site (as is usually the case when CD-ROMs are inserted into books), realize that the timing of disc and packaging manufacturing may well be different. Check with your vendors and/or assembly contractor to make sure that your work meets their schedules.

What Happens During Manufacturing?

When your data or premaster gets to the CD manufacturing facility, it goes through a number of steps before it becomes an actual CD-ROM for your audience. As we've mentioned, if your data is not in the form of a disc image, the CD-ROM manufacturer must first create the disc image from your sources. In this section, we'll assume that the disc image has already been created.

The CD-ROM production cycle consists of these major steps:

- Mastering
- Developing
- Creating stampers
- Injection molding
- Metalizing
- Printing
- Inspection and packaging
- Shipping

Mastering

The first step in the actual CD-ROM replication is *mastering*. During mastering, the manufacturer transfers the data from the customer's input media to a piece of optically flat glass (shown in Figure 9) which is coated with a photo-sensitive material called *photo-resist*. The photo-resist material is spincoated (as shown in Figure 10) onto the glass master to a uniform thickness of 150 nm. The photo-resist process must be performed in a clean room that has constantly circulated air with careful temperature and humidity control; this tight control is required to get the uniform etching necessary to produce consistent pit shapes.

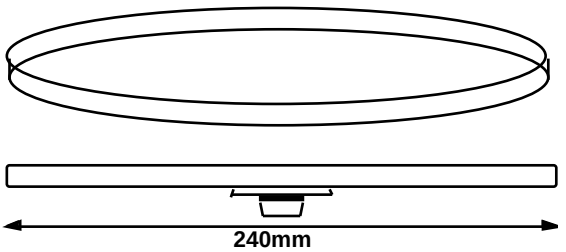

Figure 9: Glass substrate on mounting hub

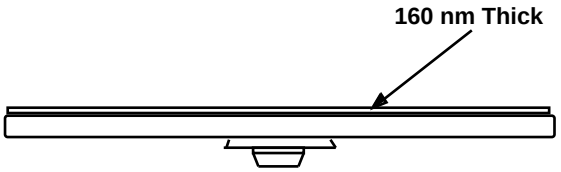

Figure 10: Photoresist layer

The photo-resist is sensitive to blue and ultraviolet (UV) light, and is chemically altered when it is exposed to the light from a blue argon laser.

During the mastering process, the digital data from the customer's input disc image is converted in real time into a CD format through *eight-to-fourteen modulation* (EFM), as described earlier. During EFM processing, *error detection codes* (EDC), *error correction codes* (ECC), and necessary *sub-codes* are added to the data stream. The data stream, including data, EDC, ECC, and subcodes, is interleaved to minimize the effects of surface defects on the CD. The output of the EFM encoder is used to modulate (turn on and off) the cutting laser, exposing the areas of the photo-resist where the pits will reside.

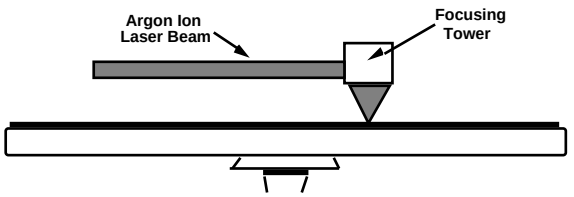

Figure 11: Laser writing on the blank

Once the laser writing (shown in Figure 11) is complete, the pits are present only as chemical alterations of the photo-resist. The physical pits are not created until the glass master is developed.

Developing

A CD master is developed in much the same way that film is developed. First, a prerinse of deionized water washes off any particulate contamination prior to the actual developing process. After the prerinse, the manufacturer rinses the glass master with a sodium hydroxide developing solution, as shown in Figure 12; this solution washes away the photo-resist that was altered by the cutting laser. The developing period is carefully timed because the precise size, shape and depth of the CD pits are critical.

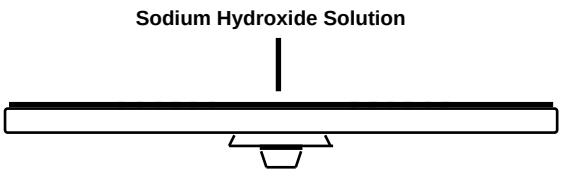

Figure 12: Glass master developing

The developing period is followed by a postrinse, which dilutes any remaining developer solution, thus preventing any further developing. The glass master is then dried, and for the first time the image of the CD can be seen. Finally, the glass master is then visually inspected for any defects which could affect the playability of the disc. If a defect is found, the master will be rejected and a recut will be ordered. If the master has no visual defects, it moves on to the next step.

Creating stampers

The next step in the CD-ROM manufacturing process creates the disc stampers. First, the glass master is placed in a vacuum deposition chamber, and a thin layer of vaporized silver is deposited onto the photo-resist side of the glass, as shown in Figure 13. As for certain other steps in manufacturing, this operation must be performed in a clean room. When this process is completed, the stamper (known as the glass master at this point) is inspected once again for visual defects.
There are three reasons for coating the glass master with silver:

- The silver makes the surface of the glass master conduct electricity for the electroplating process.
- The electrical conductivity of the disc allows electronic testing of the plating results.
- The silver makes the CD reflective, allowing it to be played on a special master player.

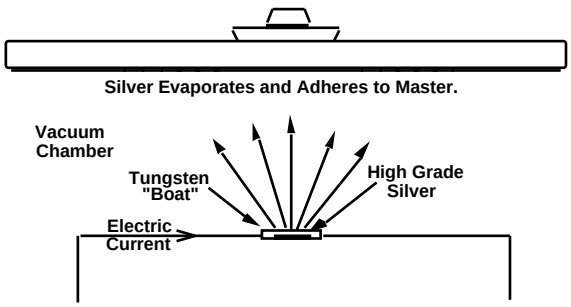

Figure 13: Silver deposition

At this point, the master player runs an extensive inspection of the entire surface area of the CD-ROM (shown in Figure 14), allowing the operator to ensure the quality of the master. The master player also verifies that all the signals on the disc are within tolerance. Masters that do not meet the quality requirements are rejected, and the master is recut.

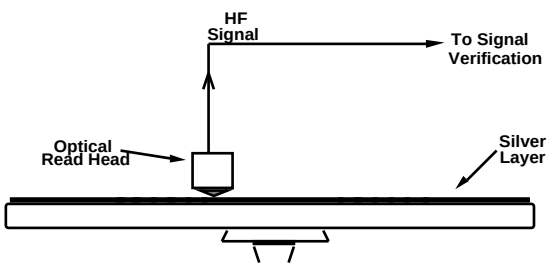

Figure 14: Master player inspection

Electroplating

Once the master has successfully passed the master player, it is sent on to electroplating.

During the master electroplating process, the glass master is immersed in a plating solution, and an electrical field is created around the master. This process causes nickel particles in the plating solution to adhere to the metalized surface of the glass master, as shown in Figure 15.

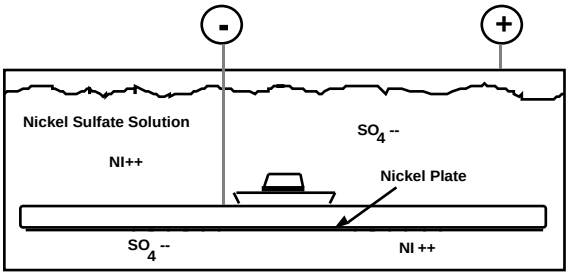

Figure 15: Electroplating

After about two hours, the nickel layer is thick enough so that it can be peeled off of the glass master, forming a negative image of the glass master called a metal master or *father*. The structure of the metal master is shown in Figure 16.

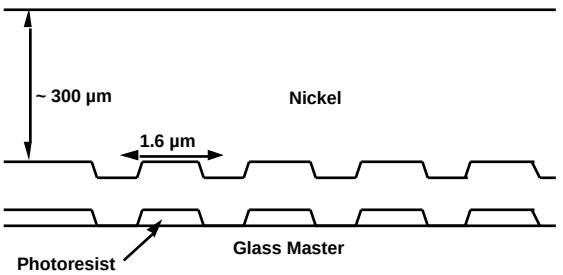

Figure 16: Metal master

Because the father is a negative image of the glass master, it has bumps instead of the pits that were on the glass master. For short runs, the father can be used as the CD stamper. For larger runs, however, the manufacturer electroplates the father, and creates a negative image of the father, called the *mother*. Because the mother has pits, it is electroplated again to make stampers for the main part of the production.

After electroplating, stampers must be finished by sanding the back side and trimming them to the finished size, as shown in Figure 17. Upon completion, the stamper(s) are inspected for visual defects.

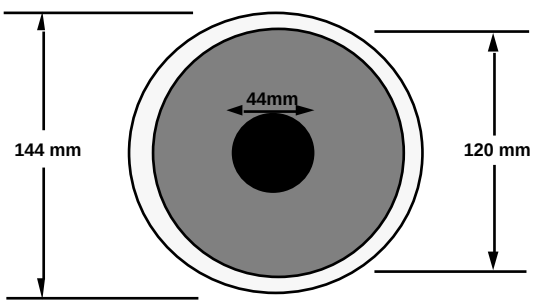

Figure 17: Stamper finishing

Stampers wear out after a substantial amount of use; a stamper is normally good for 10,000 to 15,000 discs. By making several stampers, the manufacturer can produce discs on several production lines at once—an important issue on long jobs with short turnaround times.

Once again, because the mother has pits and the stampers are a negative image of the mother, the stampers have bumps which will result in pits on the finished CD.

Once the stamper passes inspection, it is sent to injection molding.

Injection molding

During injection molding, the stamper is loaded into the molding machine and becomes one side of the mold. The other side of the mold (called the mirror block) is flat and forms the readable side of the CD-ROM. During the molding process, the mold closes, and molten polycarbonate is injected into the mold (as shown in Figure 18). The polycarbonate is injected as a liquid and fills the space between the stamper and the mirror block, forming a negative image of the stamper.

At this point, the mold is cooled to solidify the polycarbonate and the disc is removed by a robotic arm. The CD-ROMs are clear at this point, so they cannot yet be played.

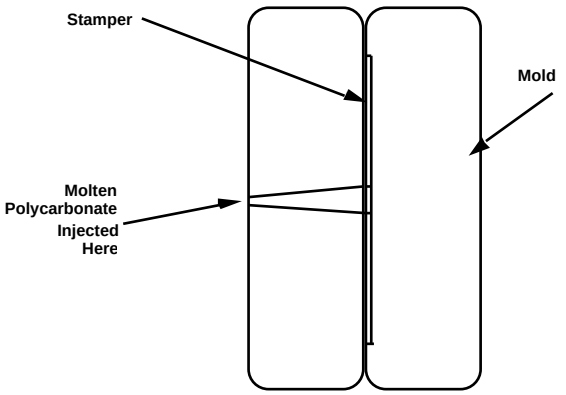

Figure 18: Molding

Metalizing

To make the just-pressed disc readable, it must be made reflective. This occurs during a metalization step, shown in Figure 19.

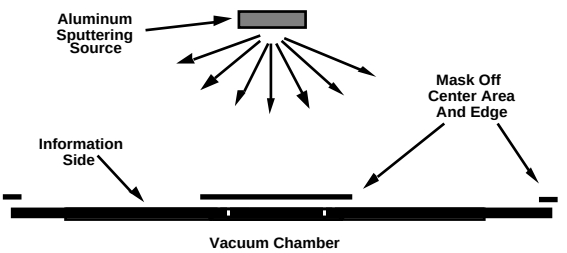

Figure 19: Metalization

Here, the transparent CDs are coated with a thin aluminum layer in a process similar to an electrostatic discharge; this layer makes the discs reflective. The manufacturer checks the first discs to be metalized, both visually and electronically, for defects in the stamper qualification process. Discs are verified, byte-for-byte, against the original customer-supplied source material, and this same testing is applied to samples from the production as it proceeds.

Next, the discs are spin-coated with an acrylic lacquer material, shown in Figure 20, that protects the delicate pits and prevents the aluminum from oxidizing. Some discs are made with gold metalization; since gold is not susceptible to oxidation. Although these gold discs are claimed to last longer, the gold doesn't make

the data any better. Also, because gold is less reflective than aluminum, these discs are a little harder to read.

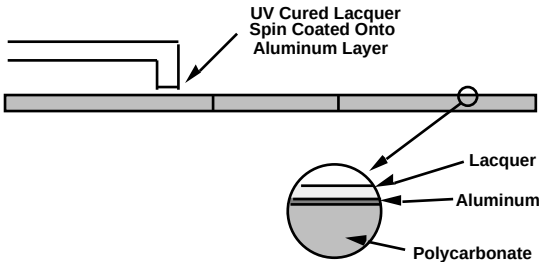

Figure 20: Protective coating with cross-section of disc layers

Printing

In the final step of the manufacturing process, the manufacturer prints the discs on the lacquer side, using customer-provided artwork, either by a silk-screen or an offset process, as discussed earlier in this chapter. The printing is then ultraviolet-cured to fix the image. Figure 21 illustrates the process of silk-screen printing.

When disc printing is complete, the CD-ROMs can safely venture outside of the clean room for the first time.

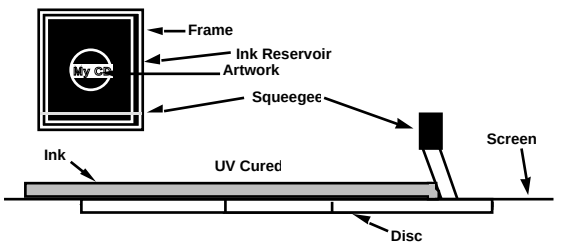

Figure 21: Silk-screen printing

Inspection and packaging

During the final inspection step, the CD-ROM manufacturer inspects every disc that comes out of the clean room, visually as well as electronically, and rejects, and scraps, any disc that does not meet stringent quality standards.

The discs that do pass final inspection are packaged according to the customer's requirements.

Shipping

Once the CD-ROMs have been packaged, they're ready for shipping. The manufacturer boxes and ships the discs in whatever way you specify, usually by a traditional land carrier, though they can be shipped by higher-cost courier services if your schedule demands it.

Because CD-ROMs are small and light in weight, they are usually relatively inexpensive to ship, so shipping costs will not significantly impact your budget. However, be sure to include the shipping time for the discs in your production schedule.

Shipping and scheduling details vary from one CD-ROM manufacturer and shipper to another, so contact your manufacturer early to make whatever shipping arrangements you need. And remember that, if you are taking delivery of the discs yourself, you'll need some place to store them.

List of Disc Manufacturers

The CD-ROM manufacturer information is courtesy of the Optical Publishing Association and The CD-Info Company. For more information on the Association, contact Richard Bowers, the Executive Director, at 614-442-8805, fax 614-442-8815, at Compuserve 71333,1114 (Internet 71333.1114@compuserve.com), or through their Web site at

`http://www.meta-media.com/opa/home.html`

The CD-Info Company is a company in Huntsville, Alabama, that publishes information through The CD Information Center at

`http://www.cd-info.com/`

The information in this list and the next section is current as of mid-1996 and the reader is encouraged to consult the Web sites above for up-to-date listings. These sites also contain other information related to CD-ROMs and the reader is encouraged to check them whenever fresh information is needed.

3M Company
3M Optical Recording Department
3M Center
Bldg 223-5N-01
St. Paul MN 55144
800-336-3636, 612-733-2142; 612-733-0158 fax

3M California
2933 Bayview Dr.
Fremont CA 94538
510-440-8161; 510-440-8162 fax

Allied Manufacturing (WEA)
6110 Peachtree St.
City of Commerce CA 90040
213-725-6900; 213-725-8767 fax

Americ Disc
Sales Office
7575 Trans-Canada, Suite 500
St-Laurent, Quebec H4T 1V6 Canada
514-745-2244; (+1) 514-337-3989 fax

Plant
2525 Canadian St.
Drummondville QB J2B 8A9 CANADA
800-263-0419, 819-474-2655; 819-474-2870 fax

Americ Disc
2 Sheppard Ave. E., Suite 900
Willowdale ON M2N 5Y7 CANADA
416-512-7001

AmericUK
London Service Center
London, UK
(+44) 181 600 3900; (+44) 181 749 7057 fax

Miami Plant
8455 N.W. 30th Terrace
Miami, FL 33122
305-599-3828; 305-599-1107 fax

California Plant
4701 Stoddard Road
Modesto, CA 95356
209-467-2400; 209-467-2417 fax

San Jose (California) Service Center
Triptych CD Corporation
1250 Aviation Avenue, Suite 155
San Jose, CA 95110-1121
408-271-7373; 408-271-7370 fax

Minneapolis Service Center
2100 West 96th Street
Bloomington, MN 55431 USA
612-881-6446; 612-881-6476 fax

American Multimedia
2609 Tucker Street Extended
Burlington NC 27215
910-229-5554; 910-228-1409 fax

ASR
8960 Eton Ave.
Canoga Park CA 91304
818-341-1124; 818-341-9131 fax

Astral Tech Americas, Inc.
5400 Broken Sound Blvd.
Boca Raton FL 33487
407-995-7000; 407-995-7001 fax

Australian Compact Disc Manufacturers
37 Orsmond Street
Hindmarsh, SA 5007 Australia
(+08) 3 46-2333; (+08) 3 40-9040 fax

BQC
146 West Olentangy Street
Powell, Ohio 43065
614-799-0884; 614-799-0885 fax
http://www.concourse.com/bqc

Cassette Productions
4910 W. Amelia Earhart Drive,
Salt Lake City UT 84116
801-531-7555; 801-531-0740 fax

Cinram Canada
2255 Markham Rd.
Scarborough ON M1B 2W3 CANADA
416-298-8190; 416-2298-0612 fax

CINRAM New York
660 White Plains Road
Tarrytown, NY 10591 USA
914-631-2800; 914-631-9281 fax

CINRAM SouthWest
5307 La Viva Lane
Arlington, TX 76017
817-465-2210; 817-465-1074 fax

CIS Technology USA, Inc.
Creative Data Products, Inc
1005 Montague Expressway
Milpitas, CA 95035 USA
408-934-2455; 408-934-2450 fax

(DADC) Digital Audio Disc Corporation (Sony)
1800 N. Fruitridge Ave.
Terre Haute IN 47804
800-323-9741, 812-462-8100; 812-462-8866 fax

Denon Digital Industries Inc.
1380 Monticello Rd.
Madison GA 30650
706-342-3425; 706-342-0637 fax

Disc Factory
6525 Sunset Boulevard, Suite 205
Hollywood, CA 90028
213-465-7522; 213-465-2457 fax

Disc Manufacturing Inc.
1409 Foulk Rd., Suite 202
PO Box 7469
Wilmington DE 19803-0469
800-433-DISC, 302-479-2500; 302-479-2527 fax
http://www.discmfg.com/

DMI Anaheim
1120 Cosby Way
Anaheim CA 92805
714-238-7156/7158; 714-630-0303 fax

DMI Huntsville
4905 Moores Mill Rd.
Huntsville AL 35811
205-851-0373/0236; 205-852-8354 fax

Disctronics Texas Inc.
2800 Sumitt Ave.
Plano TX 75074
214-881-8800; 214-881-8500 fax

EMI Manufacturing
1 Capitol Way
Jacksonville IL 62650
803-522-9893; 803-522-3242 fax

Europa Disk Ltd.
75 Varick S.
New York NY 10013
212-226-4401; 212-966-0456 fax

Eva-Tone
4801 Ulmerton Rd
Clearwater FL 34622
800-EVA-TONE, 813-572-7000; 813-572-6214 fax

Future Media Productions, Inc.
25136 Anza Drive
Valencia, CA 91355
800-360-2728, 805-294-5575; 805-294-5583 fax

Harmonic Hall Limited
19th Floor, Wharf Cable Tower
9 Hoi Shing Road
Tsuen Wan, N.T., Hong Kong
(+852) 2412-1388; (+852) 2414-2333 fax
http://www.hk.super.net/~harmonic

HMG
15 Gilpin Ave.
Hauppauge NY 11788
516-234-0200; 516-234-0346 fax

IBM
1001 W.T. Harris Blvd.
Charlotte NC 28257
704-924-6300; 704-924-7191 fax

IPC
9400 Jeronimo
Irvine CA 92718
714-588-7765; 714-588-7763 fax

JVC Disc America
9255 Sunset Blvd., Suite 717
Los Angeles CA 90069
800-677-5518, 310-274-2221; 310-274-4392 fax

JVC Disc America
2 JVC Road
Tuscaloosa AL 35405-3598
205-556-7111; 205-554-5535 fax

KAO Infosystems - Optical Division
1857 Colonial Village Lane
Lancaster PA 17601
800-525-6575, 717-392-7840; 717-392-7897 fax

Kao Infosystems
41444 Christy Rd
Fremont CA 94338
510-657-8425; 510-657-8427 fax

Matrics Corp.
300 Main Stree
East Rochester, NY 14445 USA
800-747-0583, 716-383-4185; 715-383-4188 fax
http://www.matrics.com

Matrics Corp. Georgia
6425 Lawrenceville Hwy
Tucker, GA 30084 USA
800-908-9920, 770-908-0426; 770-934-5145 fax

Matrics Corp. North Carolina
8101 Tower Point Drive
Charlotte, NC 28227 USA
800-747-0583, 704-845-2143; 704-845-3595 fax

Matrics Corp. Maryland
8639 Cobbscook Harbor
Pasadena, MD 21122
800-875-7288, 410-360-2772; 410-360-2866 fax

Matrics Corp. Florida
2117 Dekle Avenue, Suite J1
Tampa, FL 33606
800-878-4997, 813-251-2966; 813-251-2846 fax

Metatec-Discovery Systems
7001 Discovery Blvd.
Dublin OH 43017
614-761-2000; 614-766-3140 fax
http://www.metatec.com/

MPO
MPO
Paris, France
(+33) 1 41 10 51 51; (+33) 1 41 10 51 52 fax

MPO GmbH
Cologne, Germany
(+49) 221 92 16 700; (+49) 221 92 16 703 fax

MPO UK Ltd.
London, UK
(+44) 181 600 3900; (+44) 181 749 7057 fax

MPO Replitech
Barcelona, Spain
(+34) 3 638 34 45; (+34) 3 638 15 72fax

Madrid, Spain
(+34) 1 643 12 38; (+34) 1 643 02 38 fax

MPO-Siam
Bangkok, Thailand
(+662) 651 91 51; (+662) 651 91 54 fax

Nimbus Information Systems, Inc.
Nimbus East Coast
P.O. Box 7427
Charlottesville, VA 22906
800-782-0778, 201-379-2890/4883; 804-985-4692 fax

Nimbus West Coast
4524 Tobias Avenue
Sherman Oaks, CA 91403
800-292-0932; 818-783-7475 fax

Optical Disc Corporation
12150 Mora Drive
Santa Fe Springs, CA 90670
310-946-3050; 310-946-6030 fax

P&O Compact Disc GmbH
Auf dem Esch 8<
49356 Diepholz, Germany<
+49 (0) 5441 977 0, 977 175, 977 180;
+49(0) 5441 977 177 fax
100043.764@compuserve.com

P & Q CD
5460 N. Peck Rd. #E
Arcadia CA 91006
818-357-4088; 818-359-4229 fax

Pioneer Video Manufacturing, Inc.
1041 E. 230th Street
Carson, CA 90745
310-518-0710; 310-522-8698 fax
http://www.pioneerusa.com/replication.html

Pilz Compact Disc Inc.
54 Conchester Rd.
Concordville PA 19331
610-459-5035; 610-459-5958 fax

PMDC PolyGram Manufacturing
1251 Avenue of the Americas, 22nd Floor
New York, NY 10020 USA
212-512-9350; 212-512-9358 fax

PMDC East
P.O. Box 400
Grover, NC 28073 USA
704-734-4100; 704-734-4180 fax

PMDC West
3815 West Olive Avenue, Suite 202
Burbank, CA 91505
818-848-2442; 818-848-3090 fax

ProtoSound
14 School Street
Briston, VT 05443
802-453-3334; 802-453-3343 fax

Quebecor Integrated Media
4918 20th Street East
Fife, WA 98424
800-451-5742, 206-922-9393; 206-926-0953 fax

Rainbo Records
1738 Berkeley St.
Santa Monica CA 90404
310-829-3476; 310-828-8765 fax

Sanyo-Verbatim CD Company
1767 Sheridan Street
Richmond, IN 47374
800-704-7648, 317-935-7574; 317-935-7570 fax

Six Sigma-Print NW
(Mailing) PO Box 1418
Tacoma WA 98401
(Shipping) 4101-D Industry Dr. East
Fife WA 98424
800-451-5742; 206-922-3383

Sonopress
108 Monticello Rd.
Weaverville NC 28787
704-658-2000; 704-658-6206 fax

Technicolor
3233 East Mission Oaks Blvd.
Camarillo CA 93012
805-445-3000; 805-445-4340 fax

Technidisc
2250 Meijer Dr.
Troy MI 48404
800-777-DISC, 810-435-7430; 810-435-8540 fax

Triptych CD
1604 Tillie Lewis Dr.
Stockton CA 95206
408-271-7373; 408-271-7370 fax

Ultra Media
2048 Corporate Ct.
San Jose CA 95131
408-383-9470; 408-383-0806 fax

Uni Manufacturing (MCA)
Highway 154
Pinckneyville IL 62274
618-357-2167; 618-357-6340 fax

US Optical Disc Inc.
1 Eagle Drive
Sanford ME 04073
207-324-1124; 207-490-1707 fax

VU Video
1420 Blake Street
Denver, CO 80202
800-637-4336, 303-534-5503; 303-595-4630 fax
http://thebridge.com/ibb/vuvideo/home.html

WEA Manufacturing, Inc.
New York Office
375 Hudson Street
New York, NY 10014
212-741-1404; 212-243-8255 fax

WEA Manufacturing East
1444 E. Lackawanna Avenue
Olyphant, PA 18447
717-383-2471; 717-383-1493 fax

WEA Manufacturing West
3601 West Olive Avenue
Burbank, CA 91505
818-953-2941

Zomax Optical Media
5353 Nathan Lane
Plymouth MN 55442
612-553-9300; 612-553-0826 fax

List of Disc Packaging Sources

Like the previous section, this information comes from the Optical Publishing Association and from The CD-INFO Company. The companies listed provide printing and manufacturing services for the physcial packaging of discs. These include bubble-paks, software-style boxes, jewel boxes, and other alternatives that allow you to package your product shrink-wrapped and ready for shipment

Advantage Plus Distribution Inc.
14202 Carlson Circle
Tri-County Business Park
Tampa FL 33626

AGI
5513 W. Bay Ct.
Midlothian VA 23112
804-739-1174

Blackbourn Inc.
10150 Crosstown Circle
Eden Prairie MN 55344
612-949-2155

C-Case Corp.
822 Hill Grove
Western Springs IL 60558
708-887-7000; 708-887-7010 fax

Calumet Carton Co.
16920 State St.
PO Box 405
South Holland IL 60473
708-333-6521; 708-333-8540 fax

Crawford Custom Packaging
1414 Crawford Drive
P.O.Box 191
Crawfordsville, IN 47933
800-428-0840; 800-962-3343 fax

Digipress
2016 Bainbridge Row Dr.
Louisville KY 40207
502-895-0565; 502-893-9589 fax

Golden Rule Printing
1864 Sparkman Dr.
Huntsville AL 35816
800-239-9060, 205-895-9060; 205-722-9806 fax

Ivy Hill
Eastern Sales Office
375 Hudson St.
New York NY 10014
212-741-1404
Midwestern Sales Office
1481 Countryside Dr.
Buffalo Grove IL 60089
708-506-9650
Western Sales Office
4800 Santa Fe Ave.
Los Angeles CA 90058
213-587-3131

Optima Precision
231 Industrial Park
Fitchburg MA 01420
508-342-9626; 508-345-6153 fax

Reliance Plastics & Pkg Division
217 Brook Ave.
Passaic NJ 07055
201-473-7200; 201-473-1023 fax

Reynard CVC
550 Sylvan Ave.
Englewood Cliffs NJ 07632
201-567-8998

Software Packaging Associates Inc.
11431 Williamson Rd.
Cincinnati OH 45241-2234
800-837-4399, 513-489-2118; 513-489-2126 fax

Tecval Memories SA
c-o PTS PO Box 2471
Acton MA 01720
508-635-9863; 508-263-9350 fax

Univenture Inc.
PO Box 570
Dublin OH 43017
800-992-8262, 614-761-2669; 614-793-0202 fax

Glossary

This glossary covers much of the terminology that you will encounter when you work with CD-ROM discs and disc duplicators. Many of these terms are used in this book, but some are not; these are included to help you when you read further in the field. These terms are taken from a number of sources, including the publication "A Glossary of CD and CD-ROM Terms" by Disc Manufacturing, Inc. listed in the references.

Acrobat: An electronic document technology from Adobe Systems that includes the PDF file format, translators for creating PDF files, and readers for displaying PDF files on the screen. Acrobat is built on PostScript and provides faithful electronic displays of fully formatted pages, along with access to other media.

ADPCM: Adaptive Differential Pulse Code Modulation, a method of storing digital samples based on correlations between successive values. This is a compression technique as well.

AIFF: A sound file format primarily associated with Macintosh and Silicon Graphics computers.

Alpha: The first version of a product or title that is made available for testing or for external examination. This version will usually be incomplete or have known weaknesses, but it will have enough functionality that reviews can help shape the product.

Alt keys: Keystrokes that are formed on the PC by pressing the Alt key together with another key; these keystrokes are a primary means to execute commands from the keyboard. See *Command Keys*.

AMSTeX: A particular version of a TeX macro set that is used by the American Mathematical Society to typeset their publications. See *TeX*.

Animation: A moving image created by a sequence of individually computed frames, each of which is crafted separately.

Applet: An application designed to be downloaded to a user across the networks, with the implication that it is a modest-sized application intended for a special purpose.

ASCII (American Standard Code for Information Interchange): A standardized way of representing text and control characters in one byte each.

Assets: The resources an author or developer has to bring to a title. This can include things such as interface objects, images, movies, or text.

Attribute: A property of an object, such as an icon for a document, a color or shape for a button, or font or size for text. Attributes can be set in an authoring language and may be changed by a user.

Author: The person whose creative ideas shape a title and who may create significant pieces of the title's contents.

Authoring language: A system or application used to create an electronic document and define its contents, functionality, and interface.

AVI: Audio Video Interleaved, a file format defined by Microsoft to be used by Video for Windows. This format interleaves sectors of video data and sectors of audio data so that the video player can maintain both audio and video data streams with only minimal data buffering.

Bandwidth: The amount of information that can be transferred in a given time over a communications line, a network, or a device connection.

Beta: The second level of release of a product or title before the final release to manufacturing. This level is essentially complete and has the most glaring bugs removed, but may still have minor problems, and is the final level at which testing can occur before release.

Birefringence: Double refraction (of light). Plastic materials, such as a CD substrate, demonstrate this double refringence ability due to residual stresses remaining in the plastic from the molding process. High birefringence can interfere with the laser beam of the reader and cause reading errors.

Bitmap: A two-dimensional array of pixels that is the computer's internal representation of an image.

Block: A segment of data; on a CD-ROM, same as a sector. On CDs, data is arranged in blocks that contain header and sync, user data, error detection and correction, and control information.

BMP: A device-independent graphics file format developed by Microsoft that supports images in as few as one bit or as many as 24 bits and uses RLE compression.

Booklet: The printed material inside the front of a jewel box that includes the cover artwork and other information on the disc. The usual booklet sizes are one, two, four, or eight pages.

Bookmark: An indication of a place in a document to which the user may return directly. A bookmark can be created by the authoring system (such as a chapter title) or be created and saved by the user.

Box: A packaging technique for CD-ROMs in which a box of lightweight cardboard, printed with product identification, is used to hold the disc and additional material for an electronic title.

Blister pack: A packaging technique for CD-ROM or diskettes in which a disc is placed in a transparent envelope that is attached to the binding of a book.

Browse: To look through a document without a predetermined goal in order to see what it might contain that would be of interest.

Bundling: Including more than one title in a package that is marketed as a unit.

Button: An area on the desktop which will respond to a mouse click. These are often represented by icons that represent a choice or an action. See *Hot Region*.

Button palette: A collection of buttons, each with a descriptive icon or other artwork, that acts as a vocabulary unit for user interaction.

Byte code: A device-independent, compressed version of a program produced by a compiler that can be interpreted on any of several platforms. The original version was probably the Pascal P-code, but byte codes are now used by Java.

Capacity (Data Capacity): The amount of data that can be recorded and replicated on a CD. Normal capacity of a disc is 654.7 megabytes, which is 335350 2K sectors. This is equivalent to a music playing time of 74 minutes 30 seconds. The amount of data on a disc is controlled by several factors: track pitch, speed or rotation. It is possible to record and replicate somewhat more data on the disc than the "normal" capacity and still remain within Yellow Book specifications, but some drives have trouble reading from these discs.

Catalog number: A number printed on the disc label that gives the publisher's catalog number for ordering or inventory.

CAV: Constant Angular Velocity. Refers to the speed of the information track with relation to the reading head (laser). Video discs, most magnetic discs, and traditional phonograph records rotate at constant angular velocity, i.e. RPM is constant so that the tracks on the outside radii move past the reading head much faster than tracks on the inside radii. See *CLV*.

CD-I: Compact Disc Interactive. A compact disc format in which computer data and compressed audio are interleaved on the same track. The format includes both a disc layout and an operating system, CD-RTOS, that can read the layout and play the disc contents. CD-I discs must be mastered on special proprietary systems.

CD-MO: Compact Disc Magneto-Optical. A standard for discs that can be recorded and played by magneto-optical techniques (Orange Book, Part I).

CD-Plus: A mixed-mode disc for the music industry that allows both computer data and music to be placed on a single disc.

CD-R: Compact Disc Recordable disc (Orange Book, Part II). See *CD-WO*.

CD-ROM: Compact Disc Read Only Memory. A compact disc format that is used to hold text, graphics, and hi-fi stereo sound. The disc is built on the same technology as the music CD, but uses different tracks for data. The music CD player cannot play CD-ROM discs, but CD-ROM players may be able to play music CD discs and have jacks for connection to an amplifier and/or earphones.

CD-ROM Mode 1: The usual mode for data-only CD-ROM (that is, discs that contain only data and applications); it has three layers of error detection and correction for computer data.

CD-ROM Mode 2: Another mode for CD-ROM use that has two layers of error detection and correction, for audio or compressed audio/video.

CD-Video: A disc format, also known as karaoke CD or the White Book standard, that allows a combination of audio and full-motion video. It uses interleaved MPEG video and audio sectors to maximize the amount of information that can be stored on the disc.

CD-WO: Compact Disc Write Once. A CD-ROM version of the WORM (Write Once Read Many) technology. For companies wishing to do in-house preparation through premastering, this format is useful for creating test discs (one off) before sending data for mastering and replication. CD-WO discs conform to ISO 9660 standards and can be played in CD-ROM drives.

Cinepak: An asymmetric codec that gives good results for compressed digital video. Cinepak was developed by SuperMac Technology, Inc., and is now a Radius product as a result of the merger of SuperMac and Radius in August 1994.

CIRC: Cross-Interleaved Reed-Solomon Code, the error detection and correction technique used by audio CDs.

Click: A pair of events, mouse-down and mouse-up, that are used to make a selection on the screen.

Clock: An object that represents time in a document and is driven from the computer system clock. This allows you to order events, such as a sequence of movie frames, to the user in real time.

CLV: Constant Linear Velocity. Refers to the speed of the information track with relation to the reading or recording head (laser). CD tracks pass the laser head at a constant linear velocity (1.2 to 1.4 meters per second), meaning that the speed of disc rotation when reading the inner radii must be faster than when reading the outer radii. See *CAV*.

Codec: A system for encoding (co) and decoding (dec) information to provide compression of a file or document.

Color depth: The number of bits used to determine the color of a pixel on the screen. The more bits used for a color, the more colors are available. Eight-bit color allows you to have 256 colors, for example; 24-bit color gives you over 16 million colors.

Color lookup table: The system used in an indexed color system to define the color used for each color index. See *Palette*.

Compound document: An electronic document that includes more than one computer medium in an integrated presentation, frequently with a user interface that allows the user to manipulate the individual document components.

Compression ratio: The ratio of the original size of a file to the size of the compressed file. Thus a 3:1 compression ratio means that the compressed file is one-third the size of the original.

Content: All the assets of an electronic title, together with the user interface, as assembled into the final title.

Control panel: A system extension that allows the user to set parameters for system functions. See *DLL* and *Extension*.

Copyright: The legal rights to reproduce, publish, and sell an electronic title or the contents of a title.

Copyright notice: A notice in the electronic title that states the appropriate copyrights for the title and, if appropriate, for individual content items.

Credits screen: The screen of an electronic title that lists the credits for the title's development, production, and support.

Cursors: On-screen graphic items that follow the motions of the mouse or trackball. Their shapes and actions can indicate to the user the capabilities of the area where the cursor is located or can tell the user something of the status of the system.

Cyberspace: A large and rich virtual reality in which a user can experience a range of virtual worlds.

d-characters: The characters that are legal to use in file names in the ISO 9660 disc standard. These are the upper-case alphabetic characters A...Z, the digits 0...9, and the underbar character _ .

DAT: Digital Audio Tape, a digital tape format that can be used to store or communicate data, or as disk backup.

Data fork: The part of a Macintosh file that contains the data or executable code for the file. See *Resource Fork*.

Data transfer rate: The speed with which data can be read from a CD-ROM drive. The standard "single-speed" rate is 150 KB/sec. Double-speed drives read at 300 KB/sec, and drives are now available that read at 800 to 1,000 KB/sec, which seems to be the top speed that is supported by the physical properties of discs.

Deliverable: A component of an electronic title, such as a function or an asset, or the title itself, that is to be completed and working at a specific time in the development schedule for the title.

Derived work: Generally, a work that is created by modifying or adapting a previous work; as provided in copyright law, the right to create derived works is limited to the owner of the copyright on the original work.

Designer: The person who takes the assets of an electronic title and, using an authoring system, designs the user interface and the presentation that define the overall look and feel of the title.

Developer: The person or group who takes the assets and design for an electronic title, adds additional assets as needed, and assembles them all into a working document using one or more authoring systems.

Dial: An interface component that displays a value, magnitude, or position of something and possibly allows the user to change that value. Scroll bars are an example of dials.

Dialog box: A box on the screen that requires a response from the user in order to control the application or document that presented the box. Dialog boxes can be modal or nonmodal; see both terms.

Digital library: A collection of material, gathered much as a traditional library would gather material, which is all in digital form and which can be accessed across the networks. Such libraries may be distributed anywhere in the world.

Digital Video Disc: A new optical disc technology that uses denser recording techniques along with layering and two-sided manufacturing to achieve very large disc capacities. Digital Video Disc readers are able to read CD-ROMs as well. This is apparently now sometimes called “Digital Versatile Disc.”

Direct manipulation: A user interface style in which the user first picks the object, then indicates what is to be done with it; e.g., first selects a file, then drags the file icon to another disc icon to indicate that the file is to be copied to the other disc. This is a "subject-verb" kind of command, as opposed to a "verb-subject" command style such as "COPY MYFILE.DAT C:\DATA"

Directory: The list of all the files on a computer medium.

Directory path: The full sequence of directory names needed to specify a file relative to the medium on which it resides. For example, this section’s data file on one of the authors' discs has the directory path BigDisk/EPBook/Chapters/Glossary .

Disc real estate: Disc storage, so called because there is a limited amount of storage available on the disc, and it must be developed carefully to get the most value from the resource.

Distributor: The person who takes an electronic title and sees that it is placed in retail stores and other places where it can be noticed by the potential audience.

DLL: A small piece of software that adds functionality to the Windows operating system or to an application. See *Control panel*.

Documentation: Information for the user of a disc, often provided on the disc's printed pieces or in the disc's contents. This often includes fundamental program operating instructions or information on required configurations for computers to use the disc.

Drag: An action sequence of a mouse-down to select an object, moving the mouse while holding the button to indicate that the object should follow the mouse, and a mouse-up to indicate an ending position, used to move a graphical object on the screen.

Drop: An ending behavior for a drag, in which the final position of the drag is on an icon for an action; this indicates that the action should be taken on the object of the drag.

DTD: Document Type Descriptor, a detailed description of the way each tag in an SGML document affects the formatting of the text covered by the tag.

Dual-directory disc: A disc with both HFS and ISO 9660 directories that may share the same files. See *Hybrid disc*.

DVD: The acronym for Digital Video Disc; see that entry.

ECC: Error Correcting Code. A code construction that facilitates reconstruction of part or all of a message received with errors. The error correction scheme for compact discs is the same CIRC that is used for music compact discs. See *CIRC*.

EDC: Error Detecting Code. A code construction that makes it possible to detect when a message is received with errors.

EFM: Eight-to-Fourteen Modulation. The operation of converting 8 bits of data to 14 bits for storage on the disc, to facilitate reading data from the disc.

Electroforming: A means of creating a metal master (father) disc by electroplating nickel onto the glass master until a sheet of nickel has been built up to a usable thickness. The father can then be used in the same system to create a mother, and from the mother, stampers or metal parts are made that are used in the injection molding machine to manufacture the CD.

Encapsulation: Protecting data by only allowing access to the data through a fixed set of operations whose action on the data is limited. One of the key advantages of object-oriented programming languages is that their objects encapsulate their data.

EPS: Encapsulated PostScript. A file format for importing and exporting PostScript language files in a variety of heterogeneous environments. EPS files use a subset of the full PostScript language and may include a screen preview.

Error concealment: Techniques for recovery from disc data error with minimal notice by the user. For example, in music CDs errors are concealed by interpolating linearly between good data values.

Event: A logical entity created by a hardware condition, such as a keystroke or mouse action, or by a software action. Events are asynchronous, that is, an event may happen at any time, and a program or title must respond to each event and perform an appropriate action for each.

Event handling: Responding to an event by passing the event entity to the appropriate part of a program so it may be processed correctly.

Event queue: The collection of events that have been generated but have not yet been handled, managed as a FIFO (first-in, first-out) queue.

Fair use: Generally, the personal, noncommercial, limited use of copyrighted factual material in a way that does not damage its market potential, as provided by copyright law in the United States and many other countries.

FAQ: Frequently Asked Questions, a posting on an Internet newsgroup that is intended to orient a user to the issues in the newsgroup and to provide fundamental background information for the reader.

Father: The first electroformed part made from a glass master. See *Metal master*.

File: A collection of information, either program or data, stored on a computer's secondary storage systems such as disk or CD-ROM.

File header: Information at the beginning of a file that describes the content of the remainder of the file. This is fairly common in image files, where the file header can describe the dimensions, color depth, and palette of the image.

Find: To be able to move to the point in an electronic document where a user-specified item is located; this is usually a word or phrase, but it could be another kind of content.

Firewall: A facility that screens network access to a site to protect the site's contents from undesirable access.

FOB: Freight on board, a term indicating the point from which the customer is required to pay shipping.

Font: A typeface for displaying text on the screen or for printing text on the screen.

Frame rate: The number of frames of a (digital) movie displayed per second.

FTP: File Transfer Protocol, a network technology for transferring files on the Internet.

Full color: A computer color system where each pixel holds its own RGB values, typically with 8 bits of precision for each color for a total of 24 bits per pixel. See *Indexed color*.

Functionality: A behavior or activity in an electronic title such as playing a digital movie, allowing keyword searches, having hypertext links, or being able to print the document.

Ghost site: A site on the World Wide Web that may once have been available but that no longer can be reached.

GIF: A digital image format developed by CompuServe that is widely implemented on many kinds of computers. It describes images that use lookup tables of up to 8 bit depth.

Gigabyte (GB): Roughly a billion, or thousand million, (actually 1,073,741,824) bytes, used as a measure of the capacity of a computer's disk system or of a medium.

Glass master: The medium on which manufacturers record data as the first step leading to the replication process. Consists of a glass disc larger than replicated discs, coated with a photosensitive material in which the data are recorded by a laser beam recorder (laser light).

Gold master: The completed electronic title ready for release to manufacturing (RTM) or candidate RTM. The name reflects the fact that this release is usually done on a CD-WO disc, and these discs are gold-colored. See *RTM*.

Green Book standard: The standard for Digital Video Discs.

HDTV: High-definition television, an emerging standard for commercial television that is expected to have a large impact on computing and electronic publishing.

High Sierra standard: The first draft proposal submitted to the International Standards Organization for common file structure for CD-ROM. When it became adopted, it was changed in minor ways and became known as ISO 9660.

Highlighting: Indicating a selection or possible selection by changing the color or presentation of the selected object to indicate that the user has chosen or could choose the object.

History: A record of the sequence of actions taken by the user in working with an electronic title, often used to allow the user to retrace steps or move to anything that was previously seen.

HFS: Hierarchical file system; the file system used by the Apple Macintosh system.

Home page: The hypermedia document that is first loaded when you start a World Wide Web browser, or the page that an individual creates to represent his or her presence on the World Wide Web.

Hot text: A piece of text on the screen that can respond to a mouse event.

HTML: HyperText Markup Language, an SGML document type that is generally used as the authoring language for creating documents on the World Wide Web. This language allows a document author to apply appropriate text styles and to create links to hypermedia components or to additional documents by interpreting URLs.

HTTP: HyperText Transfer Protocol, an Internet document transfer protocol that is becoming the dominant protocol for sharing documents across networks based on URI references.

Huffman coding: A way of encoding data in variable-length symbols so that data that occurs more frequently is represented by shorter symbols.

Hybrid disc: A disc that contains data in both HFS and ISO 9660 directories. This can be done by having separate data for each directory or by having the two directories share part or all of the data. This latter is also called a *dual-directory disc*.

Hypermedia: An extension of the concept of hypertext, in which the user chooses the sequence in which to view not only the text but also the other media components of the document. See *Hypertext*.

Hypertext: Text in which some words are hot text that provide links to other text in the same or in a different document, allowing the reader to choose the sequence in which the text will be read as he or she is reading.

Icon: A graphical representation of a system component such as a document, disc, or application.

Image: In CD-ROM, the data assembled in the exact form wanted on the replicated CDs, i.e., completely premastered or image ready. In graphics, a picture to be displayed on the computer screen.

Indemnification: To promise to secure another from penalties or liabilities resulting from one's actions.

Indeo: A digital video codec developed by Intel that is primarily used for Video for Windows, although it is also available for QuickTime.

Indeo Video Interactive: A wavelet-based video coded developed by Intel.

Indexed color: Color that is specified by giving a color index for each pixel instead of giving the actual color components of the pixel. The color index is used to specify a color from a color lookup table where the actual colors are maintained. See *Full Color*.

Injection molding: A process for replication of CDs wherein molten plastic is injected into the cavity of a mold under pressure, cooled and removed as a solid, clear plastic disc. The data information is transferred to the plastic in this process from the "stamper."

Intellectual property: Generally, the property created when an author creates an original work; protecting the author's rights in this property is the purpose of copyright law and a copyright transfer gives the rights in this property to the copyright transferee.

Interframe compression: Compression of digital movies that takes advantage of frame-to-frame coherence to provide information for a given frame in terms of other frames that precede or succeed it.

Internet: An interconnected set of networks around the world that allows computers attached to any of these networks to exchange data communications freely. One of the chief functions of the Internet is to tie together these separate networks' diverse set of underlying technologies.

Intranet: A network operating entirely within an organization such as a business, allowing persons in that organization to work with network tools without exposing their work to outside access.

IP address: A number, composed of four octets (eight-bit numbers), that identifies any computer on the Internet. This is being expanded to six octets as the number of computers on the Internet grows.

ISO 13490: A file format standard for CD-ROM and DVD discs that extends ISO 9660 and corrects many of its shortcomings.

ISO 9660: A widely used file format for CD-ROM. The ISO 9660 (formerly High Sierra) standard defines a directory structure that has been accepted by the International Standards Organization. This standard allows a CD-ROM disc to be read like a write-protected hard disk. Formatting a CD-ROM to this standard will allow CD-ROM interchange on any platform that supports the standard.

IVI: The acronym for Indeo Video Interactive; see that entry.

Java: An object-oriented programming language from Sun Microsystems, somewhat similar to C++, that provides the capability of creating applets for network use.

Jewel box: The thin, clear plastic box in which audio CDs are usually sold and in which many CD-ROMs are distributed. Besides the usual jewel box, there are thinner jewel boxes as well as double-disc jewel boxes.

JPEG: A standard for encoding digital images using discrete cosine transformations and entropy coding. This standard can achieve part of its efficiency by data quantization, so it is usually a lossy encoding technique.

Kilobyte (KB): Roughly a thousand (actually 1024) bytes, used as a measure of the size of a file or the capacity of a computer's memory or of a medium.

Label: The printed area on a disc; also a brand label or imprint from a publisher.

Land (Lands): The space between the pits on a CD disc where the photoresist on the glass master was not exposed to laser light. This space is more reflective than a pit. See *Pit*.

Laser: A device that creates a beam of coherent light. In a CD-ROM setting, this beam of light is reflected off the disc and the resulting reflection is measured to detect pits and lands on the disc.

LaTeX: A version of the TeX document formatting system developed by Leslie Lamport (hence the "La") and widely used in preparing technical documents. See *TeX*.

Lead in: On a CD, the area at the beginning of a disc where the Table of Contents (TOC) is recorded.

Link: A button whose function is to move the user to a different place in an electronic title.

Localization: Adapting the text of a program or title to a language other than the one in which the original was developed so it may be sold in additional markets.

Lossless: A technique for storing or compressing files that maintains all the information in the original file.

Lossy: A technique for storing or compressing files that does not maintain all the information in the original file. Typically, lossy techniques for video or images focus on removing information that is not perceived by the viewer. This kind of technique can be used for images or movies when it is deemed that the user will have an adequate result without needing all the original information.

LZW: Lempel-Ziv-Welch coding, a dictionary-based coding technique that encodes variable-length data in a constant-length code word. The dictionary is not transmitted but is reconstructed from the data that is received. This technique is patented by Unisys and commercial applications that use it must do so under license from them.

Manual: A printed (or less often, electronic) set of instructions on how to use a program, title, or piece of equipment.

Manufacturing: The process of creating the physical discs from the gold master, consisting of creating the glass master, making stampers, molding the polycarbonate discs, electroplating the discs, printing the labels, and packaging the discs for the customer.

Master: (verb) In compact disc manufacturing, the recording of the original media (glass) in preparation for making replicates (copies). (noun) In CD-ROM, the final recording of the desired CD-ROM image to be used as a source for mastering; this may be on tape, magnetic disc, optical disc (M-O or W-O), etc.

Mastering: The process of encoding input data, created during premastering, to the compact disc standards and recording this information as a series of pits in a light-sensitive layer on a glass substrate.

Megabyte (MB): Roughly a million (actually 1,048,576) bytes, used as a measure of the capacity of a computer's memory or of a medium.

Metal master: The first electroformed part from a glass master. See *Father*.

Metalizing: A process by which a thin metal coat (usually aluminum) is deposited on the clear plastic disc after it has been injection molded. The usual process is by sputtering, although vacuum vapor deposition or wet silvering can be used.

Microdollars: Very small amounts, often less than standard currency amounts, which may be charged for accessing documents online or through digital libraries. With enough uses, these amounts build up so that actual charges can be made.

MIDI: A standard for representing and communicating instructions for a music synthesizer that can be used for including music in an electronic document.

Modem: A device that connects a computer to a telephone line and allows the computer to communicate with other computers.

Mosaic: An application developed at the National Center for Supercomputing Applications (NCSA) that supports viewing hypertext documents, both locally and on the World Wide Web, with embedded graphics, sound, and movies. Mosaic was the first successful Web browser.

Mother: A metal part electroformed from the father, used for making stampers.

MPEG: A standard for encoding digital movies that includes some original frames, prediction of intermediate frames from the original frames, and interpolation of additional frames between predicted and original frames.

Multisession disc: A disc format for CD-WO discs that allows a user to write beyond the section already written.

Navigable movie: A QuickTime VR movie that includes frames of a single subject shot from a number of different points of view, and that gives the user the capability to select these individual points of view, thus navigating around the subject.

Navigation: The process of selecting and viewing various parts of an electronic title by using the user interface tools of the title to select the parts to be viewed.

Netscape Navigator: A very widely used Web browser that has led many advances in the HTML language.

Network: A connection among a set of computers that allows them to share data communication.

Object: A software entity that has properties and can respond to messages and actions.

Object-oriented: Software that is constructed on the basis of loosely coupled objects that communicate by sending messages to each other.

OLE: Object Linking and Embedding, a Microsoft architecture for including components from several applications in a single electronic document and linking the components to the applications that created them. This allows the components to be manipulated by their creating application, not by the application associated with the main document. See *OpenDoc*.

OpenDoc: Apple's architecture for allowing live links from a document's components to applications that can manipulate the components directly. See *OLE*.

Opening sequence: The sequence of screens or information that is displayed as a title is opened.

Orange Book standard: The Philips/Sony standard for CD-WO and CD-MO discs, named for the color of the cover of the book in which the standard was published.

Package: The container in which a CD-ROM is stored, shipped, and sold, and which identifies the disc to the consumer.

Packet: A fixed-size sequence of few hundred bytes of data that is created when a message is sent on a network and is given identification that lets the network deliver this data along appropriate paths so the receiving computer can reassemble the larger message.

Palette: The set of colors used in an image when that image uses an indexed-color scheme; typically the palette can be adjusted to fit a particular image.

PDF: Portable Document Format, a file format developed by Adobe Systems for their Acrobat system that is becoming fairly widely used for electronic document interchange.

Pit (Pits): Information spots on a CD (or optical disc). Pits are formed in a photosensitive layer on a glass master by exposure to laser light. Exposed material washes away in the developing process to form a pit. Pits are less reflective than lands, the space between pits. See *Land*.

Pixel: The smallest spot on the screen that can take on a discrete color; in an indexed color system, each pixel is associated with a color index, while in a full color system, each pixel has a full set of color information.

Platform: A computer system on which an electronic title can be played.

Player: The software that displays the contents of an electronic title for the user. Often this is a small application that interprets a master source file for a specific platform.

Plug-in: Software that extends the capability of an application.

Points of contact: The persons at a publisher, distributor, or wholesaler with whom an author or developer will work directly.

POSIX: A set of standards for the UNIX operating system that were established to ensure portability and security.

PostScript: A standard page-description language developed by Adobe Systems that has become one of the key standards for printing and prepress technology and for document transportation.

Premastering: Preparing the digital data to send to the CD manufacturer for mastering and replication. The data is assembled as a contiguous image the way it should appear on the CD-ROM, including the file structure (such as ISO 9660). Disc manufacturers usually have hardware and software to premaster for customers, at an additional price.

Print: To create a copy of part or all of an electronic title on paper or another hardcopy medium.

Proof disc: A CD intended to be used for testing. This usually refers to a one-off, CD-WO, or CD-R disc, but can also refer to one or more discs from a replicated group submitted for testing. Can be used as input for disc manufacturing.

Protective coating: A coating of lacquer or polymer deposited over the metal coating on a CD to protect and seal the metal layer. The most common method is spin-coating of a UV curable polymer over the surface of the metalized disc and then passing it under ultra-violet light to polymerize or cure it.

Publisher: The person or organization who funds the development of an electronic title and arranges for its marketing.

Pull-through: Sales of a title that are created when the title is associated with a larger series of publications.

Quantization: Replacing higher-accuracy data, either digital or analog, by lower-accuracy digital data. This can lead to aliasing effects and the loss of information.

QuickTime: An architecture for time-based presentation that was developed by Apple Computer. It is the basis for the QuickTime digital movie system, as well as for sound or any other time-based control. It is said that one could automate a factory based on QuickTime.

QuickTime VR: A system for creating immersive environments from photographic panoramas. It runs under QuickTime 2.0 for Macintosh and Windows. The VR stands for Virtual Reality. See *Virtual reality*.

Rack jobber: Someone who takes products from a wholesaler and sees that they are placed in shelf space, preferably good shelf space, in retail outlets.

Reader: Software that displays an electronic title on the screen and that manages the user interaction in the title.

READ_ME: The name of a text file that tells a user important information about a disc, often including last-minute information that is not found in the disc's printed pieces or other documentation.

Red Book standard: The original CD standards, set by Philips and Sony, were published in a book with a red cover. These standards are the basis for later standards for other kinds of CD standards, such as those for CD-ROMs.

Registration card: A card that a purchaser of a title sends to the publisher to register himself or herself as a legal owner of that title. The registration card is often a printed piece that accompanies a CD-ROM, but it may be an online form that is filled out an printed or filled out and delivered to the publisher electronically.

Resource fork: The part of a Macintosh file that contains information such as icons or text that supports the file's presentation or operations. See *Data Fork*.

Retail: Selling an item one unit at a time to individual customers.

RGB: A color model in which a color is determined by giving its red, green, and blue components. This color model is directly associated with the way a standard computer monitor creates color with a red, green, and blue color mask, so it is the most common color model for computer images.

RLE: Run-length encoding, a system for compressing data, particularly image data.

Rock Ridge standard: A standard for disc file systems and directories that is designed to allow users to maintain much of the directory information in a UNIX operating system. This goes beyond ISO 9660 to include longer file names with richer character sets, symbolic links for files, and the like.

ROI: Return on investment; the percentage of profits in relation to the total cost of a project.

Royalty: The portion of the publisher's income from a title that is passed to the title's author as payment for the author's intellectual property.

RTF: Rich Text Format, a set of conventions developed by Microsoft to include formatting information in a text document independently of the computer system or word processing system used to generate the text.

RTM (Release to Manufacturing): To sign off on a gold master and certify that it is correct and ready to manufacture.

Schedule: A set of dates by which the individual tasks associated with a larger job are to be completed.

Screen real estate: The area of the screen, so called because there is a limited amount of screen space available, and it must be developed carefully to get the most value from the resource.

SCSI (Small Computer System Interface): A standard for connecting computers to peripheral devices such as CD-ROM players or external hard discs. This is the standard interface for peripherals on the Macintosh and is sometimes used on other computers.

SCSI-2: A newer version of the SCSI interface that is up to 5 to 10 times as fast as the original SCSI standard.

Search: The instruction to examine the contents of an electronic title in order to find an item in the title that matches the user's request.

Search engine: online software that helps the user locate information on the Web or, more broadly, on the Internet by specifying words to be matched.

Sector: The smallest unit of a CD-ROM's file structure that may be accessed. This corresponds to 1/75 of a second of audio and contains 2352 bytes of digital data.

SGML: Standard Generalized Markup Language, a system for marking a text document based on its structure so it may be presented according to a format defined external to the document.

Shelf space: Space in a retail outlet in which products are displayed. Since retail sales are directly linked to exposure, it is very desirable to get shelf space in outlets for your product.

SKU: A unit of piece goods sold through retail channels.

Specifications: A description of the contents or functionality that are to be present in an electronic title.

Spin coating: Creating an even layer of photoresist on the glass master by using centrifugal force to spread the material.

Spindle: A set of finished CD-ROMs that are not packaged or wrapped, but are delivered by the manufacturer for the customer to package. This is named for the metal rod on which the discs are stacked during manufacturing.

Stamper: A metal part electroformed from the mother. The stamper is inserted into the mold cavity to become one side of the cavity. “Stamper” is a misnomer inherited from the phonograph record industry. CDs are not stamped, but are injection molded.

Subcodes: Codes used in the CD format to hold various kinds of information, depending on the disc type.

Substrate: The main physical body of a disc, on which other coatings or layers may be added. Compact discs are made of polycarbonate plastic, coated with metal, then coated with a UV curable polymer. A label is then printed. The polycarbonate is the substrate.

Synchronization: The process of ensuring that simultaneous events, such as events in the video and audio parts of a digital movie, are presented to the user simultaneously. This is one of the major tasks of digital movie playback architectures.

Tag: A notation in a text file that specifies how that text is to be treated or displayed.

Tagged text: Text that includes tags for formatting or linkage purposes, such as SGML or HTML text.

Terabyte (TB): Roughly a million million (actually 1.0995×10^{12}) bytes, used as a measure of the capacity of a computer's disk system.

Term: The duration of a contract or agreement, after which time the agreement is no longer valid. Term may be defined as a certain period of time (e.g. five years), as perpetual, or until a condition occurs (e.g. the work goes out of print).

TeX: A text formatting system invented by Donald Knuth in order to provide proper formatting for technical documents, particularly for mathematical expressions. It uses tagged text and often custom macros to interpret the tags.

TGA: A file format originally designed by Truevision to accompany their Targa graphics boards.

TIFF: Tagged Image File Format, a popular format widely used to store and transmit graphic images. The TIFF standard supports full-color images as well as lookup table images.

Title: An electronic document or publication that is identified and sold as a unit.

TOC: Table of contents. This is information located in the lead-in area. The TOC contains a listing of where tracks start on the disc, as well as indications to the player as to what kind of disc it is: ROM, audio, etc.

Track: The sequence of pits that are read by the reading laser comprises a track. On a CD, a track is a spiral beginning at the inside of the disc and spiraling outward, and is about three miles long. Also, a contiguous portion of the spiral of pits and lands on a compact disc; on an audio CD each track ordinarily corresponds to a single song or piece. The tracks on a disc are identified in the disc table of contents.

Track pitch: The physical distance between two rows of information pits, center to center. In CD the specification of track pitch is 1.5 to 1.7 microns. Most discs are recorded with a track pitch of 1.6 microns.

Trademark: A name indicating ownership of origin of a product that is legally reserved to the owner or creator of the product.

Tray drop-in piece: The printed piece that may be inserted in a jewel box below the plastic insert in which the disc hub holder is found. Also known as the tray insert or tray card.

Truevision-S: A video codec developed by The Duck Co. for both AVI and QuickTime movies.

Turn: The time required for CDs to be mastered, made, and shipped, measured from the time premastered data, artwork, and other materials are in the hands of the manufacturer. Also known as the turnaround time.

Unicode: A technique for storing characters from most languages using two bytes per character.

UNIX: An operating system widely used for workstations. There are many variations on this operating system, depending on individual vendors, but they all function in much the same way. See *POSIX*.

UPC: Universal Product Code, the bar code that is placed on a product so it may be scanned by point-of-sale devices. Bookstores and other retail merchants want products to use UPCs to streamline the sales process. UPCs contain a company code and a product code, and a product manufacturer must apply for a company code.

URI: Uniform Resource Identifier, a set of conventions for identifying any file on any system accessible on the networks so that the file can be transferred to the host and the appropriate application can be called upon to display it.

URL: Uniform Resource Locator, a URI convention based on the network address of the file to be transferred.

URN: Uniform Resource Name, a URI convention intended to allow access to a document by a name instead of by location.

User interface: The components of an electronic title that enable the user to navigate through a document and to execute the functionality of the document.

Video for windows: A Microsoft design for creating and playing digital video on the Windows computing platform.

Virtual reality: Real-time display of direct-manipulation interactive realistic computer graphics and digital sound with three-dimensional models, presented by a display technology that allows the user to have an experience of immersion in the model presented.

Virtual Reality Modeling Language: A specification for 3-dimensional models and interactivity across the networks, allowing for the creation and networked sharing of virtual worlds.

Virtual world: The synthetic model space in which a user is immersed in a virtual reality environment.

Volume descriptor: An area at the beginning of a CD-ROM reserved for the recording of information about the origination, originator, copyright, etc.

Volume name: The name of a file volume, such as a CD-ROM, that is mounted on a system.

Volume table of contents (VTOC): The list of tracks, along with their position and duration, on a CD-ROM; the list of all files in the volume (the directory) on a magnetic disc.

VRML: The acronym for Virtual Reality Modeling Language; see that entry.

WAV: A digital audio format for Microsoft Windows.

Web browser: Software that allows a user to access and view documents across the computer networks and supports hypertext links through the HTML language.

Web, the: See World-Wide Web.

White Book standard: A standard developed by Philips and JVC for compact discs that supports a combination of audio and full-motion video. The video is based on MPEG video and audio that are interleaved to achieve the proper data flow rates. The standard is published in a book with a white cover.

Wholesale: Selling a product to an individual or organization that will redistribute it to individuals through one of several possible distribution channels including direct mail, catalogs, or retail outlets.

W-O: Write Once. Recordable optical disc. Can be recorded on, but not changed or erased. See *WORM*.

Work for Hire: Creating a piece of work under contract to another, with the rights to that work going to the person who hired the creator. This is a legal relationship that needs to be defined by a contract between the creator and the person hiring her or him.

World-Wide Web (WWW): A system developed at the European Particle Physics Laboratory (CERN) in Switzerland for creating and browsing distributed hypertext documents.

WORM: Write Once, Read Many. See *W-O*

Yellow Book standard: When the CD-ROM standards were set by Philips and Sony, they were published in a book with a yellow cover. Thus the standard for CD-ROM is sometimes called the yellow book standard.

YIQ: A color model defined for NTSC color television that takes luminance (Y) as its primary component and uses two other components, a range from orange to cyan and a range from magenta to green, to carry chrominance information. It is relatively easy to convert back and forth between YIQ and the RGB color model that is most commonly used in computing.

Readings

Alber, Antone F., *Multimedia: A Management Perspective*, Integrated Media Group (Belmont, CA), 1996

Adobe Systems, *PostScript Language Reference Manual*, 2nd edition, Addison- Wesley, 1990

Ames, Patrick, *Beyond Paper*, Adobe Press, 1993

Bienz, Tim and Richard Cohn, *Portable Document Format Reference Manual*, Addison-Wesley, 1993

Bowers, Richard A., list of resources for CD-ROM publishing, Optical Publishing Association, 1996

Bush, Vannevar, "As We May Think," *Atlantic Monthly*, July 1945, 101-108.

Busk, J. Philip, Clayton Summers, C. Steven Langer, and James R. Fricks, *Compact Disc Terminology*, 2nd edition, Disc Manufacturing Inc.

Cunningham, Steve, and Judson Rosebush, *Electronic Publishing on CD-ROM*, O'Reilly & Associates, 1996

Deep, John and Peter Holfelder, *Designing Interactive Documents with Adobe Acrobat Pro*, Wiley, 1996

Disc Manufacturing, Inc., *A Glossary of CD and CD-ROM Terms*

Disc Manufacturing Inc., *Overview of DVD*

Disc Manufacturing Inc., *An Overview of Multimedia CD-ROM Production*

FAQ, alt.cd-rom, available by anonymous ftp from ftp.cdrom.com as /pub/cdrom/faq

Fricks, James R., *Compact Disc Terminology*, Disc Manufacturing Inc., 1992

ISO 9660 Standard, American National Standards Institure, 1988

ISO 10149 Standard, American National Standards Institute, 1989

Lopuck, Lisa, *Designing Multimedia*, Peachpit Press, 1996

Niederst, Jennifer, *Designing for the Web*, O'Reilly & Associates, 1996

Nielsen, Jakob, *HyperText and HyperMedia*, Academic Press, 1990

Phillips, Richard L., "Opportunities for Multimedia in Education," in *Interactive Learning Through Visualization*, S. Cunningham and R. Hubbold, eds., Springer-Verlag, 1992, 25-35.

Pohlman, Kenneth, *The Compact Disc Handbook*, 2nd edition, A-R Editions, Inc.

Samuelson, Pamela, "Copyright's Fair Use Doctrine and Digital Data," *Communications of the ACM*, 37(1), January 1994, 21-27

Samuelson, Pamela, "Copyright and Digital Libraries," *Communications of the ACM*, 38(4), April 1995, 15-21, 110

Schriver, Karen A., *Dynamics in Document Design*, John Wiley & Sons, 1997

Smedlinghoff, Thomas J., *The Software Publishers Association Legal Guide to Multimedia*, Addison-Wesley, 1994

Sosinsky, Barry, and Elisabeth Parker, *Acrobat Quick Tour*, Ventana Press, 1995

Summers, Clayton, *Introduction to ISO 9660*, Disc Manufacturing Inc., 1993

Wiederhold, Gio, "Digital Libraries, Value, and Productivity," Communications of the ACM, 38(4), April 1995, 85-96

Williams, E.W., *The CD-ROM and Optical Disc Recording Systems*, Oxford University Press, 1994