

Portable Document Format (PDF)—Finally, a Universal Document Exchange Technology

Wan-Lee Cheng

Electronic File Exchange

As you well know, electronic and computer technologies have revolutionized the graphic/visual communications industry over the past 24 years. Word processing, computer graphics, desktop publishing, digital media, and the Internet have completely changed the communication flow in every environment. Now people can enjoy producing and receiving high quality, color realistic, and information rich visual images in affordable forms. However, despite the rapid technological advancement, there have always been information exchange problems between users because of incompatibility among different computer platforms and software programs.

Many approaches in both hardware and software development have attempted to solve electronic file exchange problems, but none have proved promising until the development of the Portable Document Format (PDF) in 1993. Hamilton (1999), in his article “PDF Output,” pointed out the value of PDF:

After wandering in its infancy, PDF is now entering the commercially viable stage of its life. Having been initially pitched to the corporate/office communications and online publishing markets as a stable, cross-platform tool for document distribution, it is finding a home with Acrobat’s core audience in publishing, prepress, and commercial printing. (p. 26)

One of the keys to making information exchange work well is to have a universal vehicle to deliver electronic data without losing its fidelity and integrity. Therefore, developing an exchangeable file format has been quite diligently carried out by the computer industry. Some of the resulting accomplishments such as *Text Only*, *Rich Text Format*, *Initial Graphics Exchange Specification (IGES)*, *Drawing Interchange Format (DXF)*, *Tagged Image File Format (TIFF)*, *Joint Photographic Experts Group (JPEG)*, and *Graphics Interchange Format (GIF)* did make file transfer possible, but they only worked for particular kinds of image files and did not solve all problems.

For example, Text Only and Rich Text Format are useful for word processing files. Even though problems about exchanging data among users of different word processing

software mostly get solved, type fonts, styles, and text-formatting requirements may not all be converted properly. Other solutions—such as IGES and DXF—cater to computer-aided drawing files that exchange vector-based data, while TIFF, JPEG, and GIF are designed for pixel-based image conversion.

PDF, however, brings new promise. Finally, a software technology provides a common file format for computer users of Macintosh, PC Windows, and UNIX platforms, allowing them to communicate regardless of operating system, hardware configurations, or even native application software. Kessling (1998) clearly summarized the purpose behind the development of PDF:

PDF is the ground-breaking format of the Adobe Acrobat product line, which celebrated its market debut in 1993. Its intended purpose was the effortless exchange of electronic documents without having to worry about platforms, applications, versions, or fonts. (p. 213)

Features of PDF

PDF allows information containing text, graphics, sound, animations, and movies to be faithfully delivered via digital means. Adobe Acrobat—actually a suite of software—converts documents to PDF and allows users to view the file contents with their original visual richness across computer platforms.

Furthermore, a PDF document can later be converted back to the PostScript format, then go through the Raster Image Process (RIP) for printed media reproduction. (RIP is an essential process to transform digital images to printable visual images in any printing process.) In fact, the latest development of PostScript Level 3 allows a PDF file to be RIPped directly without going through the extra PostScript converting process. To RIP a PDF file directly will not only increase productivity but also reduce potential errors by eliminating the PDF to PostScript conversion process.

Many large advertising, book, magazine, and commercial printing businesses have adopted PDF for their digital workflow in both prepress and printing production. TC Advertising, an insert printer; McNughton and Gunn, a book printer based in Salina,

Michigan; Exped Printing in New York City; and R.R. Donnelley are just some of the companies across the United States that have already implemented PDF standards in their business operations (Hamilton, 1999).

Another great benefit of using PDF is the much reduced file sizes. For example, the file size of a typical eight-page, two-color newsletter can be reduced from 1.7 MB in its native Adobe PageMaker format to only 117 KB after it is distilled to PDF. Intranet and Internet users have recognized the small file size advantage for effective online communications for a long time. Because of its simplicity and portability, printing and publishing industries along with other businesses have already relied on PDF as a principal solution for archiving documents and storing them (Messenger, 1999).

Although PDF is still growing, there is no doubt that it will become the mainstream technology for electronic communications, document distribution, and printing/publishing workflow. Furthermore, its editability, portability, accessibility, and flexibility apply to both electronic media and print media for personal and professional use. As John Deubert (1999), president of Acquired Knowledge suggested, PDF will displace PostScript and TIFF/IT as the primary file format for document distribution for printing/publishing, CD-ROM, and Internet web-based applications. In fact, the Committee for Graphic Arts Technologies Standards

(CGATS), accredited by the American National Standards Institute (ANSI), is developing the standard for using PDF files in composite data exchange (Witkowski & Kew, 1999). Figure 1 illustrates the sources, processes, and typical applications of PDF.

Adobe Acrobat—PDF Software

Several software developers were engaged in the development of PDFs in the early 1990s. Adobe Systems developed Acrobat and defined its PDF as cross-platform and independent of native application software. Through its distilling process, a PostScript file can be converted to a PDF file containing embedded type fonts, compressed image elements, and many original graphic characteristics. Anyone who receives the document can then view it in its original state, regardless of the operating system of the sender (Mac OS, PC DOS, Windows, or UNIX), the availability of native application software, or the type of font files involved.

Acrobat was indeed designed for multiple applications. According to Stoy (1999),

Adobe has historically recognized four major applications for PDF files: first are files to be downloaded from online sources, second are files to be distributed on CDs, third are files directed toward output devices such as inkjet and laser printers and digital copiers, and fourth are files directed toward conventional printing prepress or computer-to-paper devices such as Xeikon DCP-1, Agfa Chromapress, and Indigo e-Print. (p. 27)

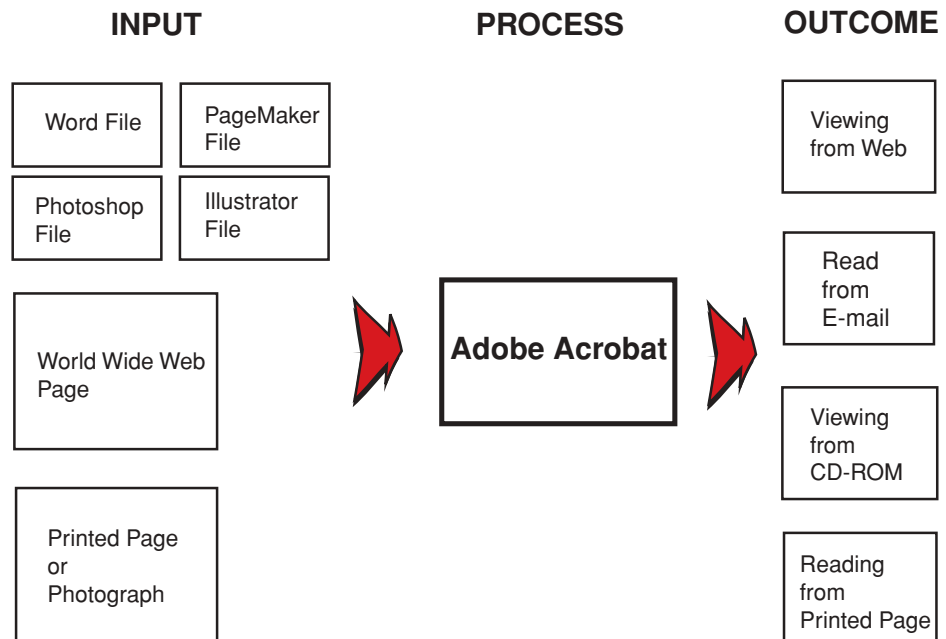


Figure 1. Typical PDF process and applications.

In 1996, Adobe Acrobat 3.0 was released, which was a breakthrough from previous Acrobat versions. It not only solved many problems, such as errors during distilling complicated vector-based entities and incapability of integrating into the prepress workflow, but it included functions to handle high-end printing needs. With the Acrobat 4.0 version upgrade in 1999, it further refined many of the weaknesses that 3.0 could not address, such as TrueType font support, editing capability, and color management. Acrobat's latest version (Acrobat 5.0) again enhanced its functions of editing, large format accommodation, and Asian font support, which not only increased the output power but also has made it a truly universal documentation exchange tool.

Adobe Acrobat consists of the following six parts:

- *Acrobat Reader* is a free software download from the Adobe web site (<http://www.adobe.com>). It is also distributed with other Adobe application software. It is a viewing program, so users can open a PDF file and navigate through pages of the document and even print out hard copies.
- *Acrobat Distiller* is a program that converts a PostScript file to a PDF file. After a document is created with a native application software, then saved as a PostScript file, Distiller converts it into PDF file format. Many of the latest Adobe applications, such as Photoshop, Illustrator, and PageMaker, are equipped with a built-in distilling function to handle the one-step direct conversion from the document itself to PDF. However, many non-Adobe programs still need to go through the PostScript conversion process.
- *Acrobat* is both a viewing and editing program. Acrobat provides many useful editing features to an already existing PDF document. For example, inserting pages, replacing a page with another page, cropping a page, deleting pages, creating a note, editing text, and adding annotations are all available for users to customize PDF documents. There are also many third-party plug-ins, which can be used with Acrobat to enhance the editing functions.
- *PDF Writer* is a printer driver that

converts native documents into PDF files. Instead of sending the document to a computer printer for a hard copy, selecting PDF Writer as a printing device creates a soft copy of the PDF file. This method of creating a PDF file is the simplest method. Even so, it must be noted that PDF Writer has fewer options than Acrobat Distiller.

- *Acrobat Capture* is a plug-in item that can be executed within Acrobat. The function of Capture is to allow scanned text images to be converted to a character-based PDF document. Therefore, text editing can be applied to the document.
- *Acrobat Catalog* is a program that indexes large volumes of PDF files for easy access at a later time. Similar to the card catalog system in a library, it lets the user easily locate a particular PDF document via the search engine in Acrobat Reader or Acrobat. Catalog is an essential tool for archiving PDF files in a structured manner.

Creating a PDF File

Figure 2 explains how PDF files can be created.

There are three methods to convert a native application document to a PDF file.

1. To use PDF Writer, you must have the PDF Writer driver installed and select it from the Chooser. Then, choose PDF to print your document. This is a convenient one-step process. However, because of some limitations, use this method only when the document contains mainly text and a small amount of graphics.
2. To use Acrobat Distiller, save the file in PostScript. The PostScript file will then be distilled into the final PDF document. Although this method requires extra steps for the conversion, the resulting PDF document will be most accurate and reliable. In general, if a document contains a sophisticated design or rich visuals, especially with Encapsulated PostScript images, use this method.
3. Some application software (Photoshop, Illustrator, PageMaker) have a built-in feature to create a PDF version of the document directly within the application. Use this method when it is available.

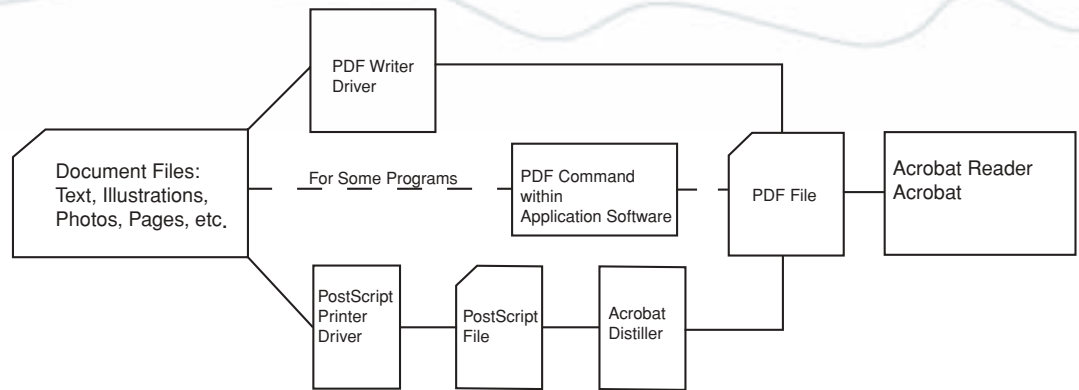


Figure 2. Typical workflow of creating a PDF file.

4. A scanned document can, likewise, be converted. A built-in Capture function, which is similar to OCR, can convert picture mode of scanned-in type-matter to a character mode text document.

Maximizing PDF's Final Outcomes

In order to preserve the accuracy of the graphic elements and the integrity of the original design, the following key control factors must be set up properly:

1. *Optimization:* Screen Optimized, Print Optimized, and Press Optimized options are available to enhance the user's outcomes.
2. *Compression and Quality:* While the size of a file is an important factor for storage and data transport, the image's integrity is essential for high-quality print reproduction. Select the best possible level of compression according to the types of images and the intended end use of the document. In general, for continuous tone, with color bitmap images used in printed matter, choose High or Maximum Quality. Medium or Low Quality is adequate for Web- and screen-based display.
3. *Font Embedding:* Retaining the exact font information of all original design is critical. The Font Embedding function in both PDF Writer and Acrobat Distiller enable PDF documents to maintain the exact fonts and format information in the original file. Therefore, the PDF document will reproduce the same type characteristics of the original, regardless of the computer platform used and/or the availability of such fonts in the system.

(To make sure that all type will be accurately reproduced and displayed, select the Embed All Fonts option.)

Applications of PDF Documents

Applications of a PDF file can be as simple as exchanging files between individuals or as sophisticated as serving a major role in the automated printing/publishing workflow. Following are some examples of PDF applications:

1. *File exchange:* One can send a PDF to another person for viewing/reproducing the document in its original form.
2. *Proof or hard copy:* To order a commercial-quality proof or hard copies from a service bureau, PDF is the most trouble-free file format for the process.
3. *Presentations:* PDF can be used for instructional or marketing presentations. The presentation can be kept in its original design and form even if the hardware and/or software that created the document are not available.
4. *Electronic publications:* Newsletters, magazines, instructional manuals, and even books can be published on CD-ROMs or via the Internet with PDF files. Interactive user interfaces and multimedia features can be created in the publications to make readings more interesting and enjoyable.
5. *Archives:* Because of the smaller file size of PDF files, images or text documents can be archived with the use of PDF to save considerable storage space.
6. *Designer-client interaction:* Graphic designers can send PDF comprehensives to their clients via modem/ISDN for proofing and approval.

7. *Printing/publishing workflow*: Designers/clients can send PDF files to printers electronically for reproduction without worrying about missing fonts, native application software incompatibility, and/or PostScript errors.

and application software are no longer barriers to communicating and exchanging information electronically, even if the contents are visually very rich. Applications of PDF cover all areas of visual communications, from printed media to displayed media to networking media.

Conclusion

Adobe Acrobat's PDF has become the information exchange standard because of its simplicity, flexibility, and universal ability. Differences in computer systems, platforms,

Dr. Cheng is an associate dean in the College of Creative Arts at San Francisco State University. He is a member of Beta Beta Chapter and received the Distinguished Service Citation in 2000.

References

- Deubert, J. (Interviewed by Keith Hevenor). (1999). Perspective. *Electronic Publishing*, 23(4), 20–22, 24.
- Hamilton, A. (1999). PDF for output. *Electronic Publishing*, 23(3), 26–28, 30, 34, 36.
- Kessling, M. (1998, May). Adobe Systems, digital media technology keeps all options open. *Pira International-Profit Through Innovation*, p. 213.
- Messenger, J. C. (1999). From asset management to asset utilization. *Electronic Publishing*, 23(8), 41–42.
- Stoy, J. W. (1999). Color workflow with PDF. *Electronic Publishing*, 23(4), 26–28, 30, 32, 34.
- Witkowski, M., & Kew, D. (1999). Composite workflow: Getting it together in 1999. *GATF World*, 11(1), 37.

