

IS IT TIME FOR DIGITAL FILE STANDARDS?

By Frederic G. Antoun, Jr.

Why is file format standardization important? Somewhere in every organization there is an individual or group that recognizes the primary reason to produce information of any kind is to communicate to others. While the simple use of a “common” language worked well to communicate when dealing with the spoken or printed word (the primary means of communication just a few years ago), it is only one aspect of facilitating communication in the digital age. The other, which has now become critically important, is a universal ability to open, view and print information from its native digital format.

Policy versus technology. For those of us who have to work with information to be published in print and on a LAN and on the Internet, hundreds of technical issues relating to platforms, applications and file sizes provide a constant source for confusion and frustration. Policymakers would say that the trees are obscuring the forest. They would explain that the “trees” are the hundreds of technology issues; while the “forest” is the need to communicate information to the broadest range of users, whether in print, via the Internet, or all three.

Policymakers on the Congressional Joint Committee on Printing long ago determined that their role was not to constantly make technology decisions, but rather to insure that government information produced with taxpayer funds was made readily available to as many citizens as possible.

Given this goal, it is not surprising that the new Staff Director of the Joint Committee on Printing, Linda Kemp, and a number of the members of the Committee have determined that the only way to guarantee continued access to government information is to provide it in a standard file format that the intended end users can open and view exactly as the author intended, and then store and/or print.

Standardization may address common tech problems. Government contractors who produce print and information dissemination products would welcome standardization. Currently, government authors produce publications and graphics using a wide range of applications, some of which are definitely out of the mainstream. In addition, digital files often include a variety of line art and halftones created in many different applications. I recently saw a publication which had art with the following extensions: jpeg; bmp; xls; ppt; and ppd. Worse yet, similar mix and match nightmares have been presented to printers not in Quark or PageMaker, but Microsoft Word or WordPerfect!

Clearly, there is a need for a standardization effort from both the technology point of view driven by contractors and service providers and the policy view, driven by the need of the government to efficiently communicate information.

Application versus file format. Even within the industry, there seems to be substantial confusion between an application and a final file format. For purposes of the standardization discussion, an application is a software program that enables the creation and easy revision of a publication. The two most common applications for this purpose are Quark and PageMaker. For text publications, the two most common word processing applications are Microsoft Word and Corel WordPerfect. On the other hand, the final file format for such a publication could be the de facto industry standard, Adobe PDF, or html. In discussions with my clients, I often hear the following: “There he goes again, oversimplifying complex technical issues.” There is some validity to this complaint, but I did not start the oversimplification. It was created by the convergence of traditional print technology and digital technologies. The hard facts of life are that while publishers, designers and printers need to have a convenient and widely used application to create and revise a publication for print and digital storage and distribution, it is very likely that the end user does not have Quark or PageMaker.

The government recognizes this, and the corresponding need to encourage standardization of the end user file format. (It would also be nice if agencies stuck to widely used application software to design their publications.)

The risk of oversimplification. Since there is a problem with the government providing digital files to printers which were created in strange or uncommon software, and there is a problem with a wide variety of file formats that are not end user friendly, why not find a file format that both printers and the end users find acceptable? This question has led to the suggestion by many that Adobe PDF is the logical solution to both the application and file format problems. While I am a big PDF fan, Adobe’s software has not yet developed to the point where it can meet the needs of all publishers, graphic designers and printers, for all publications. This is especially true for high quality 4-color publications and those that (you know the ones) involve changes and re-proofing. While providing PDF files as digital copy to printers can and has worked very well for a wide variety of government publications, making complex changes to simple publications or simple changes to complex publications is difficult to impossible when working on a PDF file. Thus, the government’s likely preferred policy solution may not provide the complete technical solution required by the printing industry.

PDF’s advantages to the government. Remembering that the government need for standardization is driven by policy—as opposed to production consideration—PDF offers the following distinct advantages:

1. Everyone can read a PDF file. PDF Reader is free, and easily downloadable from the Web. It is also bundled with the two major Internet browsers and many other software offerings.

2. The PDF publication is either right or wrong. Once distilled to PDF, there are no issues concerning missing fonts, missing or damaged illustration files, formatting, reflow, etc. In short, what you see is what you get.
3. A PDF publication is contained in one file.
4. A PDF publication always looks exactly as the government intended it to. It looks the same on Windows as it does on Mac, Unix, and Linux.
5. It is difficult for an end user to alter a PDF file, a definite advantage for most government publications.
6. PDF is already the de facto industry standard.
7. Many government publications and forms are already archived in PDF, and made available in that format on the Internet.
8. Print production of most government publications and forms from PDF is simple and effective.
9. PDF has recently made substantial strides in making the format more usable for the print production of 4-color publications, and promises continued advancements and enhancements.

The complete policy and technology solution. Because PDF clearly meets the government's policy needs for a file format which fulfills its primary goals of creating information which can be easily and quickly disseminated and used, it can not be abandoned as a standard file format simply because it creates problems for printers in certain instances. Instead, the government should move towards adoption of a standard file format, while at the same time requiring that both the PDF file and the original application files are furnished to the printer. If no changes are necessary, and the publication is of a type that permits PDF usage, the printer could simply use the PDF file. If, however, changes are required, or for technical reasons, the printer needs (or prefers) to use the original application, he would have it.

In order to facilitate the use of the government's end user standard, all contracts for the production of government publications—whether for printing, CD-ROM or Web publishing—should require the contractor to furnish the government with an “as printed or published” PDF file. This would provide the government with an exact copy of the approved final publication in a standard format for archiving or publishing on a LAN or the Web. It would also provide a file for reprinting or other republication.

A final word about “proprietary” formats. There has always been a great deal of confusion about the government's ability to specify a particular brand or product. In fact, there is no prohibition against specifying a brand product or proprietary format as long as it has been determined that there is no acceptable equivalent in the marketplace.

Postscript is the de facto format used by most government agencies now. Contrary to popular belief, however, postscript is not in the public domain, but rather, is owned by Adobe.

Typically the government attempts to use the “or equal” phraseology, but that simply does not work with digital file formats. If there is to be a universal format, there has to be certainty. Arguing about whether 20 other formats are “equal” defeats the entire purpose of establishing a standard.

Finally, it is important to note that the universal user driven file format would be a back-end conversion. In other words, publishers and graphic designers would use whatever program they desired to set up the product, and then convert it to the standard format. Despite some claims to the contrary, everything, even a tif file, can be converted to PDF via a number of methodologies, including a simple “print to Adobe PDF” option. If I can regularly convert both simple and complex documents to PDF, then anyone can do it!

Input and consensus needed! In order for all important factors to be considered, and to establish a system which will benefit the government, its taxpayers and other end users, and the printing and information dissemination industries, input from stakeholders and interested parties is needed. I would very much like to know what you think about this issue. Please email me at antoun@printlaw.com.

All comments will be summarized in a PIA (Printing Industries of America) report being prepared for the JCP. The source of comments will not be disclosed unless we obtain specific authorization . . . so feel free to editorialize.

Published in the GPIC Newsletter, April 1999.