PDF/A

Both government and private industry produce vast and rapidly growing volumes of electronic records that, because of their historical value, need to be managed, preserved, and made accessible for future generations. Identifying and implementing the tools to manage and maintain these billions of pages of records is a daunting task and requires a solution that recognizes the wide range of information systems, technologies, and formats in which government records are generated.

Q: **What is the magnitude of the challenge?**

A: The volume of electronic information is truly staggering. A recent study by the School of Information Management and Systems at the University of California, Berkeley, estimates that the world's total production of information amounts to about 250MB—some 100,000 pages—for each man, woman, and child on Earth. For a copy of the study, go to *www.sims.berkeley.edu/ research/projects/how-much-info/.*

Q: **What about U.S. government records?**

A: The number of electronic records in the U.S. government required by law to be archived is also enormous. For example, the National Archives and Records Administration (NARA) estimates that there are 36.5 billion e-mail messages each year alone in the federal government—each of which needs to be reviewed and considered for preservation. Among other candidates for archiving are military intelligence records comprising more than 1 billion electronic messages, reports, cables, and memorandums, as well as more than 50 million electronic court case files. The Census Bureau has accumulated 600 million pages of information from the 2000 census that it will be transferring to NARA—equaling 10TB of data. That is more than five times the amount of data that NARA has captured and fully processed in its entire 30-year history. For a discussion of the problems facing U.S. government agencies and NARA, go to *www.gao.gov/new.items/d02586.pdf.*

Q: **What is PDF/A?**

A: PDF/A is a standard being established to set guidelines for archiving and preserving digital documents in Portable Document Format (PDF). PDF/A—a joint initiative of NPES the Association for Suppliers of Printing, Publishing, and Converting Technologies and the Association for Information and Image Management International (AIIM)—will address the growing need to electronically archive documents in a way that will ensure the preservation of their contents over an extended period of time and that will ensure that those documents can be retrieved and rendered with a consistent and predictable result in the future. For more information about the NPES/AIIM project, go to *www.aiim.org/pdf_a.*

Q: **What organizations are involved in establishing the PDF/A standard?**

A: NPES and AIIM have organized a working group to examine the business and technology needs for the PDF/A standard and to develop the criteria. The working group consists of representatives of the technology industry, government agencies, and private sector companies. Private sector participants across a wide range of industries have a vested interest in the standards issue because they represent highly regulated industries, or serve as consultants for these industries, and regularly submit documentation to the government. These include pharmaceutical companies, document management companies, and global management consulting firms.

Adobe

Q: **What are the advantages of PDF for archiving?**

A: PDF is a broadly accepted standard for the delivery of final-format documents. PDF reading applications, such as Adobe Acrobat® Reader® software, are available to the public for free and are already widely used. There are more than 20 million PDF documents publicly available on the Internet, and almost half a billion copies of the free Acrobat Reader software have been downloaded. PDF/A offers a cost effective and efficient solution that assures that electronic records are preserved far into the future with full fidelity and integrity no matter what application was used to create the original. PDF retains the content, look, and feel of the document exactly as it was created, ensuring document integrity and security while allowing documents to be searched. For a broader discussion of PDF as a standard for archiving, go to *www.adobe.com/ products/acrobat/pdfs/pdf_as_std_4_archiving.pdf*.

Q: **How were records archived before the rise of electronic documents?**

A: In the past, paper-based records were the only acceptable method for archiving important documents. With the enormous increase in electronic records, that paper-based system has been unable to meet the needs of governments and private industry. Yet, as digital technologies have become more pervasive, there has not been a clear guide for digital archiving standards, leaving agencies and industries unsure of which formats will be accepted in generations to come. PDF/A provides the solution to that challenge.

Q: **Has this been a serious problem?**

A: The lack of a recognized and accepted electronic standard for records preservation—particularly as new generations of hardware and software have made previous digital technology obsolete— has led to the loss of significant amounts of valuable information over the past several decades. For example, military files from the Vietnam War, records from the Viking Mars mission, data from the Census Bureau's 1960 census, and land use records have been lost because of an inability to read obsolete data formats and because of the deterioration of magnetic tapes used to store that data. For more information, see *Into the Future, on the Preservation of Knowledge in the Electronic Age*, a film by the Council on Library and Information Resources and the American Council of Learned Societies, at *www.clir.org/pubs/film/future/discussion.html*.

Q: **What about NARA? Does it accept PDF/A as a standard?**

A: On October 1, 2002, NARA, the federal agency charged with overseeing records management and archiving, announced that it would begin accepting e-mails and their attachments in their original electronic format. However, NARA did not address electronic agency records that were not e-mails. It also announced that it would develop transfer requirements for PDF by March 31, 2003, as part of its Electronic Records Management initiative.

Q: **How will PDF/A benefit outside organizations and companies faced with the challenge of archiving digital documents?**

A: While federal law requires that all government agencies must archive documents that will be of historical significance, many organizations and private companies voluntarily archive documents for legal and historical purposes. A broadly recognized archiving standard would allow the preservation of electronic records in a format consistent with the government's standard. Already, pharmaceutical companies can submit new drug applications to the Food and Drug Administration in PDF, and federal courts require legal submissions from attorneys in PDF. A common standard that is compatible with the standard recognized by government agencies would give records managers the confidence that their records could be readily accessed far into the future.

Q: **Is Adobe the only manufacturer of PDF products?**

A: PDF is a publicly available specification. In other words, although Adobe created it and advances the specification through subsequent releases, all the information about the nature of PDF is made available publicly. Because Adobe chose to make the PDF specification public, there are hundreds of companies making PDF creation, viewing, and manipulation tools. For information regarding these vendors, go to *www.pdfzone.com* and *www.planetpdf.com.*

Adobe

Tools for the New Work™