Ubiquitous PDF: The Evolution and Revolution of Digital Document Distribution

Chris Cade
August 9, 2000

Professor Burkhard, Bus 119B
San Jose State University

1.1 IT Overview

The Portable Document Format (PDF) from Adobe Systems is a file format that enables documents to be viewed and printed cross-platform, regardless of what application the file was originally created with. Using Adobe's Acrobat software, any application that can print can also be used to create PDFs. Although the full Acrobat product must be purchased, the Acrobat Reader can view and print PDFs and can be downloaded for free from Adobe's website.

PDF is primarily used within organizations to post important business documents on corporate intranets; however, it has also found a place in almost all printing and publishing workflows to facilitate the efficient and effective transfer of documents between clients and publishers. It also continues to be used by organizations via online forms, including the Internal Revenue Service. PDF has already become the de facto standard for electronic document distribution, and in the future, is likely to become an important part of every organization's cost-effectiveness strategies.

First released in 1993, Acrobat was believed to be the future of document transfer. Resistance and skepticism to PDF has been very minimal over the years, especially because each new version of Acrobat included new features that addressed the concerns of any skeptics.

The development of PDF has gone hand-in-hand with the development and release of each new Acrobat version. Important developments over time include the addition of plug-ins to enable third-party development of Acrobat tools, multimedia capabilities, forms, and the ability to edit PDFs. The single-most important development was a very difficult business decision to give the Acrobat Reader away for free; this occurred in conjunction with the release of Acrobat 3.0.

PDF is now in version 1.4, and in general, has remained relatively unchanged since its inception. Acrobat 4.05c, the most current version, utilizes many tools to maximize the effectiveness of PDF. Further, PDF has become omnipresent and non-industry specific, finding an appropriate place in just about every company or organization in both private and government sectors

The future of PDF is very bright due to new applications of the file format, including Adobe's eBook software, PDF Merchant and Web Buy, as well as Acrobat Messenger. EBooks are one of the evolutionary applications of PDF and while there are other competitors in the emerging market such as Microsoft, it is likely that Adobe and PDF will persevere due to the already existing foundation that PDF has established. Perhaps the most revolutionary use of PDF is with the concept of 'reflow,' in which a normal PDF (currently unable to change width) can, in effect, be reshaped to fit any size page or device.

1.2 History.

1.2.1 Origin.

With the introduction of Acrobat 1.0 in June of 1993,[1] people first began to believe that PDF might have a significant impact on businesses. The file format had evolved from its PostScript ancestor that had been rapidly gaining popularity, and the introduction of Acrobat Distiller (software to convert PostScript to PDF), "in effect made it possible for any PostScript-generating application also to produce PDF,"[2] according to Ed Taft, Principal Scientist for Adobe's Advanced Technology Group.

Several facets of PDF provided the necessary justification for people to consider it as a potentially important business influence. The following four specific features of PDF exemplify why people originally believed it would become important:[3]

- Files can be viewed across all platforms, regardless of what operating system, software, or fonts were used to create the file.
- Navigation within the document is made easy through bookmarks and links.
- Regardless of platform or printer, files print exactly as desired.
- Optimization can reduce file size to less than that of its source.

The aforementioned features became the foundation for how proponents originally expected PDF would revolutionize businesses. Still, not everyone was quick to adopt PDF.

Even though high-end printing was not originally supported, print and prepress industries were first to try and incorporate PDF into their workflows, weighing heavily upon its cross-platform functionality for consistency assurance between clients and vendors.[4] Still, some people in those industries were skeptical of PDF's resourcefulness for their specific workflows—requiring color separations, duotones, and spot color blends—none of which were supported by Acrobat at that time.[5]

Further, with the advent of the web, the two most prospective new uses were the idea of a 'digital master' and the creation of all-PDF websites. Again, skeptics doubted that PDF would fill either role. There was a definite uncertainty whether PDF would succeed in the emerging Internet cyberspace because original versions of Acrobat did not have plug-ins to enable PDFs to be viewed on the Web.[6] Further, Anne-Marie Concepción mentions in her article, "Where are all the PDF Web Pages?" that as new versions of Acrobat Reader came out that did have the required browser plug-ins, any PDF-only websites would still not be viewable by older versions. In addition, there were no plug-in checkers to see if the viewer's Reader version was up-to-date.[7]

Concepción further points out that the inability to use a browser's back button (doing so will unload the PDF) while viewing a PDF has discouraged web designers from creating PDF-only websites. Lastly, opponents acknowledge that a large part of the Web's success at the time was based upon the searchability of its documents, and the unfortunate truth was that PDFs were not searchable by those search engines.[8]

Despite the skepticism, Acrobat pushed forward and as the product developed and new versions were released, individuals began to see more potential uses for PDF. In late 1995, PDF was first considered to possibly be the emerging "de facto standard for exchanging graphically sophisticated documents between platforms."[9]

Each new version of Acrobat was designed to meet the furthering needs of those individuals and organizations using it, and whatever skeptics there were quickly dwindled in numbers. This was especially true as each new version of Acrobat strived to address the concerns of those skeptics, as well as further facilitate the needs of current PDF users.

1.2.2 Development

PDF's development has been simultaneous with that of Acrobat, according to Taft. He says, "A PDF file is an external representation of a data structure representing a document." The data structure of PDF was based on Carousel Object Store (COS); Carousel being Acrobat's early code-name. Additions to PDF since its 1.0 inception were focused on adding new capabilities, including:[10]

- Adobe imaging model extensions to give PDF parity with PostScript
- Additional navigational aids
- Various kinds of annotations to enable the creation of interactive documents
- Multimedia capabilities such as sounds and movies
- Logical structure and other facilities to enable PDF to carry application-level information, in addition to the final-form appearance for viewing and printing

In general, PDF's development, successes, and challenges were closely intertwined with releases of the Acrobat product. According to Nabeel Al-Shamma, Engineering Director for Adobe's Core Technology Group, low sales proved to be very discouraging after the release of Acrobat 1.0.[11]

Around Acrobat 2.0, developers realized that for Acrobat to be successful, it needed to be usable by third-party developers. This is the point at which plug-ins were first incorporated into PDF.[12] Netscape's Mosaic browser was one of the first to do this, but not without a great deal of inter-company effort.

This development took place between Acrobat versions 2.0 and 3.0, according to Ken Anderson, Director of Engineering for Adobe's Web Hosted Applications Group, originally the Director of Acrobat Engineering responsible for managing the relationship between Adobe and Netscape. During that time, Adobe teamed very closely with Netscape to work on plug-ins enabling the rendering of Acrobat within the browser. In fact, Adobe engineers worked at Netscape's site with the Navigator team to help troubleshoot any technical issues. This joint venture was made easier by an executive mandate from both companies (at the time, Adobe was investing in Netscape). Even with that mandate, however, there were challenges according to Anderson:

> The Navigator team was pushing for a Windows-only OLE integration, but we wanted a cross platform solution that would work on the Mac and Unix platforms

> as well. We ended up convincing them to do a plug-in interface and worked with them to define the requirements for the interface.
>
> Acrobat had two primary requirements, first we wanted our plug-in to control the entire browser window rather than just be a component of an HTML page and second, we wanted to be able to support page at a time viewing of PDF files over the web with only a single HTTP connection per page. For the second requirement we worked with Netscape engineers to extend the HTTP 2.0 protocol to support the specification of multiple ranges of bytes that could be requested in a single HTTP header (sometimes referred to as byteserving). We then extended the PDF file format so the viewer could make these requests through the Netscape plug-in interface.
>
> The two major obstacles we faced were performance over the web and our desire to have Acrobat work both as a Netscape plug-in and as a stand-alone application without having to support too many software configurations. The fact that PDF files were multiple pages, that objects in the PDF file were shared across pages, and that some objects such as fonts and images were very large, required that we completely re-architect how the viewer accessed and displayed objects in the PDF files. The HTTP byteserving extension allowed us to minimize HTTP connect time, but our display architecture had to become asynchronous and we did a lot of work to draw objects as soon as they were available. We also implemented several techniques to improve performance such as optimizing the order in which the bytes were downloaded, drawing text with fauxed fonts first, then displaying images, then downloading the fonts and displaying the page again with the true fonts for the document. The other obstacle was our desire to install an Acrobat Reader for desktop use as well as a Netscape plug-in. We ended up building a cross-platform inter-application communication interface that allowed Acrobat to render into the browser window controlled by the very small Netscape plug-in, but this turned out to be hard and a lot of work.[13]

Al-Shamma adds that extensive lobbying of the W3C was required to get the byteserving protocol accepted, in addition to the evangelizing of web server managers. Further, he mentions that there was some reengineering needed to make Acrobat better suited for the web. Optimal organization for reading data over the web is as serial data; however, "PDF is organized as a set of objects that can be randomly accessed." Acrobat 3.0 introduced optimization, thus allowing a PDF to be read with limited random access.[14]

Acrobat didn't really take off until its 3.0 release, however. This release was the first to integrate Acrobat with both Netscape and Internet Explorer browsers, but more importantly, it was the first time that an Acrobat Reader was offered for free.[15] This was an especially controversial issue according to Greg Hansen, Senior Manager for Adobe's eBooks and Digital Rights Management group, because at that time Acrobat had very little revenue.[16]

It was believed by many that customers would not pay for the Acrobat product if the Reader could be downloaded for free. Further, even though the Reader is a component subset of

Acrobat, it still cost a lot of money to develop. The final decision was to allow the free Reader to only be able to view PDF documents without any abilities to modify or save PDFs, ensuring that the full Acrobat product would still create revenue.[17] This has proven to be the single-most effective business decision Adobe has made with regards to Acrobat, thus being the key step in enabling PDF to become omnipresent.

Acrobat 4.0 provided several features essential to the printing and publishing industries, and has continued the trend of success that its predecessor versions have paved the path for.

1.3 Status and Future

1.3.1 Current Status

PDF has remained relatively unchanged since its version 1.0 inception in 1993, says Taft.[18] Still, each new version of Acrobat utilizes PDF in evolutionary and revolutionary ways. Today, PDF is the definitive de facto standard for electronic file distribution.[19] Further, its resourcefulness is nearly impossible to gauge due to its ubiquity. In his article, "The Basics of PDF," Chris Heric likens PDF's possibilities to that of a box of Legos when he says, "Legitimate uses for the PDF file format are so varied that it is difficult to focus the attention of this article on all of the possibilities."[20]

The power of PDF, currently in version 1.4, is best exemplified by the impressive list of features—some of which were grandfathered in from previous versions of Acrobat or were previously available via third-party plug-ins[21]—which Acrobat 4.0 currently touts. Such features include, but are not limited to the following:[22]

- Forms: Acrobat may create forms to be filled in by users later. Further, the FDF (Forms Data Format) enables users to import previously entered and saved data into a blank form.
- Web Capture: Users may download websites and html content and convert them to PDF.
- Catalog and Search: These features create full-text indexes of PDFs that are fully searchable.
- Security protection: Passwords can be assigned to PDFs to prevent editing, text selection, printing, and more.
- Navigational features: Bookmarks and dynamic links make PDFs easy to navigate and use.
- Paper Capture: Scanned documents can be converted to PDF and be made fully searchable.
- Document editing: Images can be edited in conjunction with Photoshop 5.0 or greater, and text can be edited directly within Acrobat.
- Digital Signatures: SelfSign enables documents to be signed digitally. Further, signed documents can be compared to previous versions of those documents.
- Annotations: Files can be marked with text, image, and audio metadata to enable collaboration.

- Plug-in support: PDF is open-source so any developers can write plug-ins to further enable Acrobat to perform previously unavailable functions to meet the customizable needs of themselves and their organizations.

While some of the aforementioned features are new, others are from previous versions of Acrobat. Regardless, each new version of Acrobat has managed to impress users by not only addressing their concerns and meeting their needs, but also by adding new features to give users even more opportunities to utilize Acrobat in new ways.[23]

PDF has become pervasive throughout the entire world. Though it started out as a godsend for the print and prepress industries, it has immersed itself into just about every industry. While the print and prepress industries are still the heaviest users of PDF, others have also followed suit, especially the financial, pharmaceutical, legal, and government sectors.[24]

In fact, the government is now one of the largest, if not the largest, users of PDF. Many of its agencies use it, including a long and distinguished list of well-known agencies such as the FDA, EPA, FBI, SEC, DoD, FAA, US Postal Service, NASA, and the IRS. [25]

Those government agencies primarily use PDF to facilitate submissions processes and make important documents available on the Internet. The previously mentioned other industries also use PDF for those same reasons; however, they also use it for publishing information on CD-ROM, creating and filling in online forms, and streamlining printing processes for greater efficiency.[26]

Due to its prevalence, PDF is no longer just industry-centered. Many companies across all industries, focused on saving money, now place important documents on the web or their intranets, such as Cisco Systems, who recently saved $50 million with PDF.[27] Apple has even decided that all Mac OS X graphics would use PDF as the default standard.[28]

The bottom line is that individuals and organizations across all industries are using PDF to save money throughout several facets of their businesses. To date, more than 110 million users have the free Acrobat Reader on their desktops.[29] PDF has gained widespread acceptance and is now the de facto standard for electronic document distribution, and as such, it also has no direct competitors.

1.3.2 Future Expectations

Currently, there are no new technologies that are likely to render PDF obsolete. It is gaining momentum at an exponential rate; a rate that is only slightly slowed by the time it takes to release each new version of Acrobat. Further, its potential for revolutionizing businesses in the future continues to grow.

One area where organizations still believe that PDF will make an impact is in the unceasing pursuit of the “paperless office.” They are still hopeful that the internal use of PDF can eventually eliminate the need for paper within the workplace; however, this is a vision not likely to be realized anytime in the near future, according to PDF expert, C. Scott Miller. In his article,

"Acrobat, InDesign, Publishing Projects, and the Future of PDF," Miller predicts, "it will take at least a generation to get people away from reading paper."[30]

Taft also agrees that it will take a long time to move away from paper, and he suggests that rather than looking towards PDF to eliminate paper in the workplace, instead, a paradigm shift is a more appropriate perspective. The shift, he says, is from "print and distribute" to "distribute and print."[31] Simply put, historically documents have been printed then distributed among organizations, and now people are moving towards distributing files electronically via PDF, then allowing users to view and print those documents at their own discretion.

One new application that is likely to accelerate this paradigm shift, as well as accelerate the transition from paper to screen, is Acrobat Messenger. This recently released application acts as a standalone scan-and-distribute kiosk. Its basic features include the ability to scan and archive documents, then send searchable PDFs via email, fax, or directly to a network connected machine. In addition, documents can be printed.[32]

Messenger's success, and potentially revolutionary impact, is based upon its very simplistic user interface (just about anybody can use it without any training), and much more importantly, the fact that it distributes documents in PDF. By allowing printing, faxing, and emailing of documents in PDF, Messenger eliminates the need for a fax machine and copier in the office… the two most prevalent paper-based machines in existence, second only to printers.

Messenger helps PDF to become the digital master that people have been touting it to be, and further, it gives organizations a means for quickly and easily converting existing paper documents into PDF. Just as PDF has become a de facto standard, Messenger, too, will become a standard tool within corporations, according to Sally Wiener Grotta of PC Magazine.[33]

Another upcoming application of PDF is in the burgeoning realm of eBooks. In September of 1999, Adobe released its PDF Merchant and Web Buy products, server and client side eBooks software, respectively. Adobe has been involved in the development of the Open eBook Publication Structure, (OEB), a set of standards for eBook software.

According to Karl De Abrew, PDF expert, the major criticism of OEB is that it does not effectively address the issue of digital rights management, that is, it does not prevent eBooks from being easily distributable like mp3s now are; OEB leaves that up to the individual eBook creators. Adobe's eBook software draws upon PDF's already existing digital rights management features, as well as encryption technology, to ensure the protection of authors' and publishers' works. De Abrew also points out that because PDF is already a de facto standard, its creation software is tried and true, not to mention there are already existing and exponentially growing volumes of PDFs already available. For those reasons, he says PDF for eBooks is "a publisher's dream."[34] It is likely that Adobe's eBook software will become the industry leader in this new market.

The aforementioned products exemplify how PDF is revolutionizing businesses and entire markets; however, they do not address the hypothetical issue of how PDF itself may become

revolutionized. Al-Shamma believes that the real future of revolutionizing PDF is with a concept called 'reflow.'[35]

Currently, PDFs are limited to whatever page width they were created with. The basic idea behind 'reflow' is that that because PDF has structure, that structure can also be read and used to guide the physical re-layout of the PDF. This implementation of PDF could then enable PDFs to be "deployed to devices with a wide range of output widths."[36] In other words, a landscape-oriented PDF could be 'reflowed' and printed in portrait orientation, or more importantly, PDFs could be 'reflowed' to fit onto any size device, from large monitors to small hand-held devices. This concept could revolutionize PDF by enabling it to be viewable on all devices of all sizes and in all widths, further supporting PDFs rise towards becoming the digital master.

[1] "Adobe Acrobat Leaps Onto Scene," *Information Week*, June 14, 1993, pg. 78.
[2] Interview with Ed Taft, Adobe Systems, July 27, 2000.
[3] "Adobe PDF," http://www.adobe.com/products/acrobat/adobepdf.html, July 22, 2000.
[4] Sharples, Hadley, "PDF: A 'Magic Carpet' to ease prepress workflow" *Graphic Arts Monthly*, June 1997, p. 76.
[5] Bury, Scott, "Blending the Elements," http://www.dtpjournal.com/current/9906F1.html, July 30, 2000.
[6] Ostertag, Krista, "Internet Acrobats" *VARBusiness*, Dec. 1, 1995, pg. 23.
[7] Concepción, Anne-Marie, "Where are all the PDF Web Pages?" http://www.publish.com/opinion/9901/opinion.html, July 20, 2000.
[8] Ibid.
[9] Bartless, Troy K., and Lisle Reggie, "Getting to the Bottom of the Standards," *LAN Times*, Sept. 11 1995, pg. 86.
[10] Interview with Ed Taft, Adobe Systems, July 27, 2000
[11] Interview with Nabeel Al-Shamma, Adobe Systems, July 28, 2000.
[12] Interview with Ed Taft, Adobe Systems, July 27, 2000.
[13] Interview with Ken Anderson, Adobe Systems, August 3, 2000.
[14] Interview with Nabeel Al-Shamma, Adobe Systems, July 28, 2000.
[15] Ibid.
[16] Interview with Greg Hansen, Adobe Systems, August 2, 2000.
[17] Ibid.
[18] Interview with Ed Taft, Adobe Systems, July 27, 2000.
[19] Ward, Gareth, "Open eBook initiative is gathering pace" *Printing World*, Sept. 6, 1999, p. 16.
[20] Heric, Chris, "The Basics of PDF" http://ep.pennwellnet.com/home/archivearticles.cfm?Section=Archives&ARTICLE_ID=68944&VERSION_NUM=1&KEYWORD=The%20Basics%20of%20PDF, July 30, 2000.
[21] Simon, Luisa, "Agile Acrobat," http://www.zdnet.com/pcmag/stories/reviews/0,6755,2558407,00.html, August 1, 2000.
[22] "Adobe Acrobat 4.0 FAQ," http://www.adobe.com/products/acrobat/pdfs/acr4faq.pdf, July 24, 2000.
[23] Cole, David M., "A PDF for All Seasons," http://www.publish.com/opinion/9708/pov.html, July 30, 2000.
[24] Interview with Sarah Rosenbaum, Adobe Systems, July 25, 2000.
[25] Staff, "Acrobat In Action," http://www.purepdf.com/action/gov.html, July 22, 2000.
[26] "Adobe Acrobat Features," http://www.adobe.com/products/features/acrobat.html August 1, 2000.
[27] "Cisco Systems Saves $50 Million with Adobe Tools," http://www.adobe.com/epaper/spotlights/cisco/main.html, July 25, 2000.
[28] Alvarez, Israel, "Mac OS X: the big deal," http://www.dtpjournal.com/archives/9807d1.html, July 30, 2000.
[29] Davidson, Matt, "Paper in the Pipeline," http://www.adobe.com/epaper/spotlights/ikon/main.html, August 1, 2000.
[30] Miller, C. Scott, "Acrobat, InDesign, Publishing Projects, and the Future of PDF," http://www.performancegraphics.com/Pages/Acrobud.html, July 22, 2000.
[31] Interview with Ed Taft, Adobe Systems, July 27, 2000.
[32] Jefferson, Steve, "From Paper to PDF in a Few Easy Steps," *InfoWorld*, May 1, 2000, pg. 58.
[33] Wiener Grotta, Sally, "Adobe Acrobat Messenger," http://www.zdnet.com/pcmag/stories/firstlooks/0,6763,402097,00.html, July 25, 2000.
[34] De Abrew, Karl, "A Standard Response," http://www.adobe.com/epaper/columns/deabrew/991115kd.html, August 1, 2000.
[35] Interview with Nabeel Al-Shamma, Adobe Systems, July 28, 2000.
[36] Ibid.