

SEYBOLD REPORT

ANALYZING PUBLISHING TECHNOLOGIES

Workflow

Acrobat-Based Collaboration: How Well Does It Work?

PAPER-BASED APPROVAL PROCESSES ARE generally at odds with shrinking deadlines, multi-departmental reviews and the needs of cross-media production. While numerous vendors have developed network-savvy methods for viewing, annotating and approving electronic files, many of those systems are proprietary in nature and few fully support PDF. One of Acrobat 5's significant new features is the ability to add comments and digital signatures online, forming the basis of a PDF-based collaborative workflow. We tested the new features and examined their strengths and weaknesses.

A Seybold Reprint

This material appeared in *The Seybold Report* (ISSN 1533-9211), Volume 1, Number 11, September 3, 2001. Seybold Publications has prepared this reprint at the request of PlanetPDF.

©2001 by Seybold Publications, P.O. Box 644, 428 E. Baltimore Ave., Media, Pennsylvania 19063, ph. (610) 565-2480. All rights reserved. Reproduction in whole or in part without written permission is strictly prohibited. ISSN: 1533-9211. Subscriptions are available in the U.S./Canada for \$595 per year (24 issues) and Internationally for \$640 per year (24 issues).

For news between issues, visit:

SEYBOLDREPORTS.com

PDF Collaboration In Action

BY BERND ZIPPER AND JOHN PARSONS

Adobe Acrobat 5 will allow document creators, managers and producers to view, annotate and approve PDF files online. We tested these features and examined their strengths—and weaknesses—as collaborative publishing tools.

Traditional, linear approval processes are generally at odds with shrinking deadlines, multi-departmental review processes and the needs of cross-media production. Electronic files, including Portable Document Format (PDF), have often merely exacerbated the problem. Although they can be transferred more rapidly than their paper counterparts, they are still handled in the conventional way: One person looks at the document and the previous comments, makes additional comments, and sends it on to the next person. (In fact, all too many reviewers will print

out a PDF before marking it up, negating much of its value.)

To remedy this problem, numerous vendors have developed network-savvy methods for viewing, annotating and approving server-resident electronic files. Most of these involve the creation of a rasterized version of the page that is viewed on either a standard browser or a customized client application. Authorized users then add comments (usually as electronic “sticky notes”) and other marks. However, these systems’ advantages—centralized data control and relatively simple browser access—are for many users, offset by a combination of negative factors. Most collaboration systems are proprietary in nature, raising doubts about connectivity with other systems, high cost or the life expectancy of the developer or ASP providing the service. In addition, few (if any) of these systems fully utilize PDF or its underlying data structure.

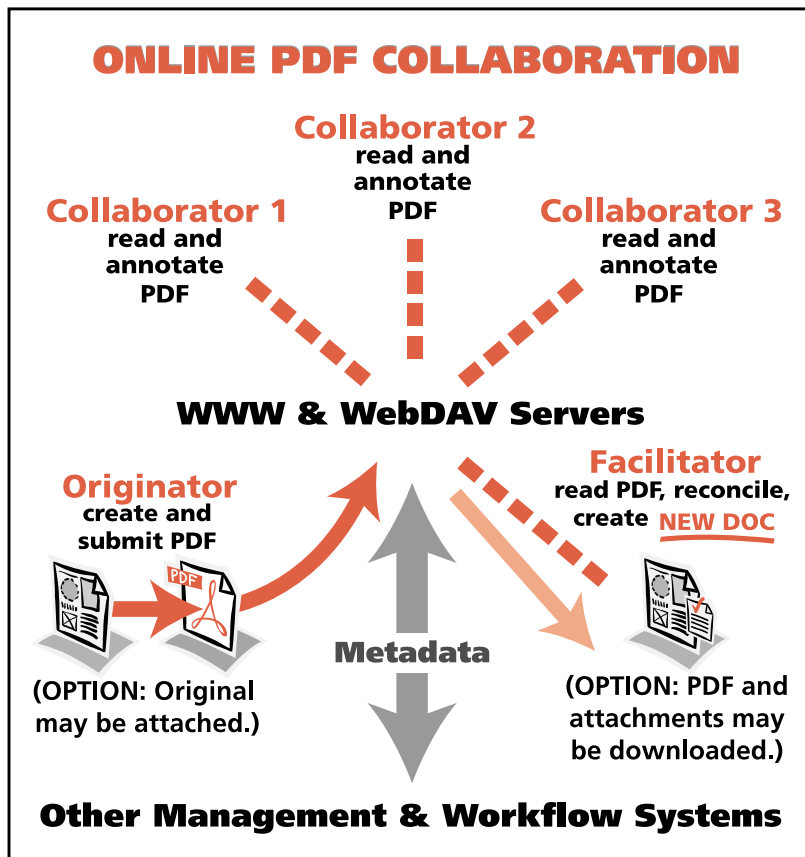
To be sure, earlier versions of Acrobat provided on-screen annotations. Users could even export such comments as small data files, using Adobe’s Form Data Format (FDF), which could then be imported into another copy of the same PDF file. However, until version 5, many users considered this manual process too cumbersome.

Acrobat 5 online

PDF files have long been viewable from a Web browser, using a free plug-in for Communicator or Internet Explorer. Previously, however, the comments or annotations could only be viewed, not created. One of the significant new features of Acrobat 5, released in April, was the ability to add comments and digital signatures online. Although Adobe’s tools are rudimentary, they form the basis of a collaborative workflow that is based, not on a proprietary raster file, but on PDF, which is widely recognized as an open, flexible and data-intensive format.

The basic workflow. Acrobat 5’s online workflow means that a PDF can be uploaded to a Web server, viewed in a browser by any prospective collaborator, and annotated online. Multiple users may view online PDFs, but must upload and download comments as a separate step. After everyone has reviewed the file, a

Collaboration simplified. To begin the process, PDF files are first uploaded to a Web server. With the full version of Acrobat 5, users may add comments to the online PDF, and even attach files—including the original source file, if desired. After everyone has made comments, a facilitator must reconcile the comments, using a variety of tools in Acrobat, and use them as a guide for making the required changes in the source file.



facilitator must act upon the comments—usually by editing a copy of the source file. Optionally, he or she may also download the PDF, with all its comments, for reference.

To participate in this process, users must have the full version of Acrobat 5, not simply Reader or Adobe's new forms review product, Approval. Online digital signatures are possible, but the process is separate from that of online comments. (See the sidebar "Signatures and Forms" on page 12.)

A separate server is required to hold the FDF files created by this commenting process. Adobe provides two methods for doing this: designating a shared folder on a network server, or specifying the URL of a WebDAV server. (Both procedures are described below.) Although it is easy to set up, the shared folder method is problematic because all users must have access to the same server, making it feasible only in an Intranet environment. WebDAV (short for Web Document Authoring and Versioning) is much harder to configure, but it allows anyone with an Internet connection to participate. (For more on WebDAV, see the sidebar on page 9.)

There is a further potential to online collaboration in Acrobat. FDF files—which can contain both comments and forms-based data—are a potential source of valuable information for business or workflow-management systems. Using Adobe's FDF Software Developer Kit (SDK), available at partners.adobe.com, developers can extract meaningful metadata from FDF files and can submit it in a Web environment using an XML flag. However, the SDK is somewhat vague on exactly how this is done.

Although standard forms-data submission is fairly common, it is too soon for developers to have employed similar tasks with online comment data. Undoubtedly, this will start to happen with a variety of business and workflow systems, especially when the Job Definition Format (JDF) becomes more widespread.

Acrobat's toolbox

Before we go into the online aspects of Acrobat 5, it is important to understand the tools themselves. In the full version of Acrobat, a variety of text and graphic markings can be applied to a PDF file, and text passages can be highlighted. The tools for enriching a PDF file with additional information, combined with easy identification of the author of each annotation, are the basis for virtual collaboration on digital documents. Adobe classifies Acrobat's annotation tools into three groups, as shown at right.

The first five tools are for commentary. They include a Notes tool (a digital



Safety First Adobe's Security Mechanisms

While the recipient can quickly judge the authenticity of paper documents by standard identifying marks—signature, letterhead, corporate seal—it has been impossible to do the same for electronic files. Acrobat offers a number of security mechanisms to help with this. First, a file can be protected against possible modification with 40-bit or 128-bit RC4 encryption. The user can choose whether the file is protected against being viewed, changed, printed or copied, as well as whether selected data can be copied and whether annotations can be edited or removed.

Next, the file can be provided with a digital signature—the Acrobat Self-Sign Certificate. The recipient can check for the correct “user certificate.” This requires the sender of the file to send a copy of the certificate separately and also to import the same certificate into the Acrobat application. Then, upon receiving a document with a digital signature, the recipient's copy of Acrobat can compare the certificates and identify the sender. Since the certificate is doubly password-encrypted, the sender need have no concern about forwarding it to the recipient. The sender also can have confidence that the document has not been altered; this would immediately be flagged.

TSR

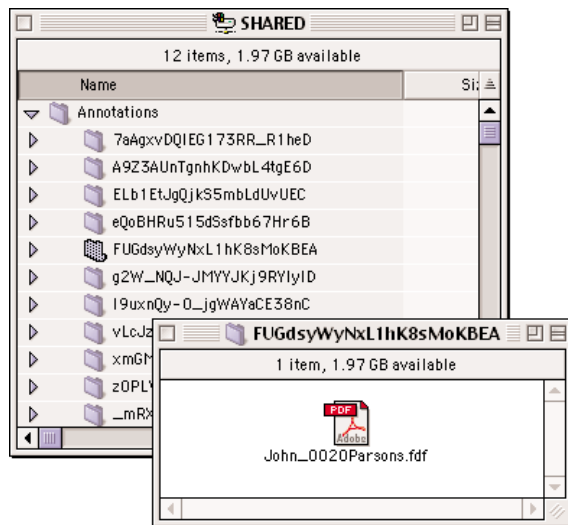
version of the “sticky note”), a Free Text tool for placement of arbitrary lines of text, and a Stamp tool for placing user-defined images, such as “Draft” or “Confidential.” Even audio comments and source files can be appended to the PDF, using the Sound and File Attachment tools. We think audio comments are of dubious benefit, since they add considerably to file size. (Of course, they might have enormous value to blind users.) However, the ability to attach files is certainly valuable, although the attachments can be read only if the recipient has the authoring application. File attachments also increase overall file size and, as actually happened recently, can be used to spread viruses. (See story on page 30 and also *The Bulletin*, Vol. 6, No. 45.)

The second group, graphic marking tools, includes a free-form Pencil tool, and self-explanatory Square, Circle and Line tools. The final group, text annotation, includes a Highlight marker tool, a Strikeout tool and an Underline tool. (The “Commenting” toolbar, shown here, also includes a spell-check tool for comments, and the Digital Signature tool. The latter does not actually create FDF data, as we will discuss below.)

With the exception of the Sound and File Attachment tools, every tool creates data that can be saved as a text file. In addition, the selected text is copied into the annotation window, giving the user an easy opportunity to suggest an improvement. This text can be spell-checked, and the text window can be moved, resized and (in Acrobat 5) made semi-transparent.

Users may specify a unique color for their annotation boxes, and annotations may be sorted (by author, date, page or type) in the Comments portion of the Navigation pane. Pages can be printed with or without

Say what? When a shared folder is used to manage online comments, the designated folder receives FDF files within folders named for cross-platform functionality. A single user's FDF file will always have the same generic name.



comments, and Acrobat includes a variety of tools to summarize, filter (suppress) and find comments.

Role of FDF. Importing and exporting annotations (FDF files) is handled via the Comments pane, or from

the File menu. Prior to Acrobat 5, passing FDF files from users making comments to users following their instructions was the only practical way to conduct a virtual editing cycle. (The alternative, annotating a PDF and sending the entire file to the next user, was impractical because of large file sizes and serious version-control issues. Swapping FDF files was not trouble-free, however. Although the files are small and can be e-mailed easily, the same version-control problems exist.)

Adobe developed FDF as a “transport format” for flexible transfer of information. For example, it is used for transferring the contents of tables or fields. The FDF format is based on the syntax of PDF. Its descriptions of objects and data are similar to those used by PDF itself, and they offer many options for display. Using this transport format, it is possible to forward and exchange data collected from forms, notes and annotations, and even optical markings.

On a small scale, this worked out very well if one had the PDF file as well as the associated FDF file on a local hard disk. However, if the user wanted to leave the PDF

WebDAV: The HTML Communications Standard

Before we can share PDF files on the Web, we must first find a way to get them there. Despite promises in numerous marketing brochures, putting data online is not often an easy matter. Absent any support from CGI scripts, the online data must usually be uploaded to the Web server using an FTP application; the normal browser will not suffice.

With Acrobat, there are two possible solutions for uploading PDF files to a Web server. One is WebDAV, which is an extension of the HTTP 1.1 protocol. It is an interpretation of the IETF working group that developed a proposed standard entitled “RFC 2518: HTTP Extensions for Distributed Authoring.” Adobe, among others, has adopted this approach. WebDAV (Web Distributed Authoring and Versioning) extends HTTP to add the capability of securing data on the server. Further, members of a team can use WebDAV to work on the same document at the same time, without being in the same place. Only an Internet connection is required.

The shared access is implemented by functions such as file locking and version control. The locking feature allows a user to temporarily block access to a file while he or she is working with it. Once the changes are completed, it is unlocked again. Locking and unlocking happen automatically, controlled by WebDAV, to avoid a “collision.” It is not necessary to maintain a network connection during the time the lock (called a “persistent lock”) is applied to a file. Thus, a file can be opened online and edited offline. Subsequently, the changes are “written” to the server.

WebDAV also provides for the association of properties with documents. These properties are metadata encoded as XML. WebDAV distinguishes between “dead” and “live” properties. Live properties are generated by the server itself, including such things as creation date and date of modification. Dead properties are name-value combinations that incorporate a URL and XML coding. In the case of Acrobat, these are online annotations.

But WebDAV does its work just as well in an intranet, and it can be implemented on a network server. Since WebDAV is dealing with “pure” HTTP, up-to-date firewalls and proxy applications do not pose much of a hindrance.

Many software products have recently been developed for WebDAV or have added support for this new standard. For example, Cadaver is a command-line client for DAV, which can be used for uploads, downloads, copies and moves between client and server. PyDAV is a DAV server implemented in Python, and a Python DAV client library is being developed. The Jakarta Slide Project was done with WebDAV. This collection of modules is leading toward a content-management system based on WebDAV. Another open-source tool is Sitecopy, which uses WebDAV and FTP to simplify Web site management. For the Mac O/S, the DAV & Goliath project is in the running. There are four separate projects to develop servers and server-management solutions based on WebDAV. On the server side, WebDAV is supported by market leader Apache (via `mod_dav`) and by Microsoft Internet Information Server 5, among many others.

A long list of commercial products are using the standard as well: Adobe GoLive, Apple Mac OS X, Xythos (a WebDAV server), Macromedia Dreamweaver, ExcOSOFT, Xerox DocuShare 2.0 and Microsoft Office 2000, to name just a few. The implementation in the universally-accepted standard HTTP appears very promising—if the future inclusion of WebDAV on the majority of Web servers could be assured and if several problems get solved. Although WebDAV is still in its development stage, it is clear that the market need is great. Even market leader Microsoft, in spite of its own technologies such as FrontPage, supports WebDAV. If it becomes possible for ordinary users to securely maintain data and documents on network servers and Web servers, that will be good news indeed.

TSR

file on the server without transferring it from one user to the next, matters became complicated and generally impractical.

In version 5.0, Adobe has carried the comments concept a step further. Both PDF files and annotations can be accessed via the Internet or a local Intranet, using current industry standards such as WebDAV and Microsoft's Web Discussions. Data exchange via SQL and folder-based FDFs within a local network can also facilitate the process.

Our test. We undertook a brief "proof of concept" test of several techniques for working with PDF files online. While these were not comprehensive tests, they did indicate some of the obstacles for the do-it-yourself approach, and they revealed a number of issues that Adobe—or its integration partners—must address. Our test scenario—revising an article for a publication—easily could be applied to designing a piece of print collateral or even making remote annotations to a legal document.

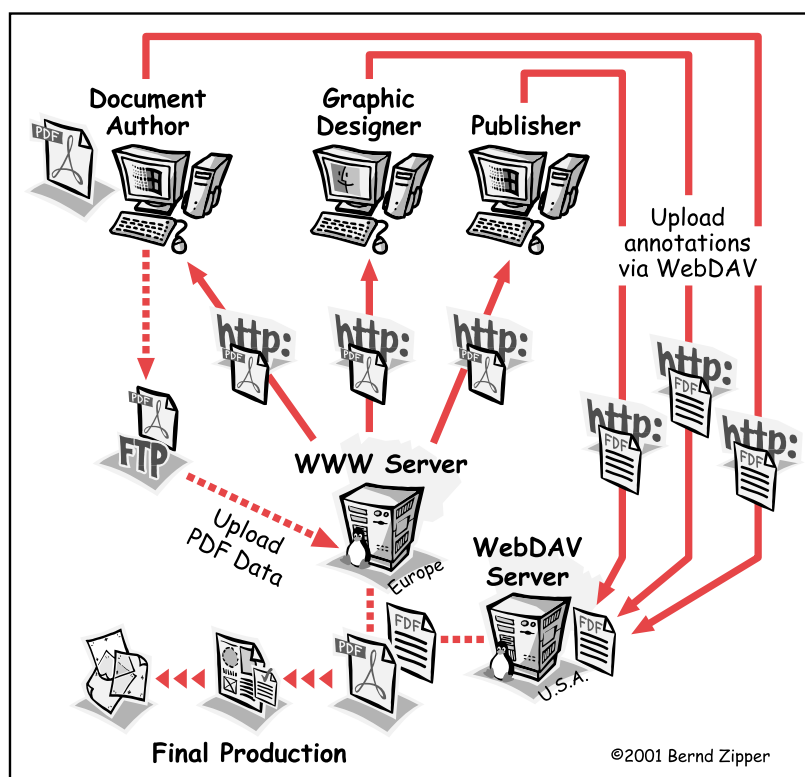
Network folder setup

Adobe's technique for using a shared folder for online comments is extremely straightforward. Each Acrobat 5 user must simply specify a server path in the "Preferences>Online Comments" dialog. In our test, we specified a folder within a Macintosh-accessible volume ("SHARED:Annotations:") on a server running Windows NT 4 ("D:\SHARED\Annotations"). When we opened a PDF file on a separate Web server, Acrobat 5 automatically created FDF files on the server volume.

Acrobat encodes FDF file- and folder-names in a manner that will be usable across different platforms. Although duplicate IDs are theoretically possible, in practice they are exceedingly unlikely. (The math exactly parallels the so-called "birthday problem." The collision probability only reaches the 50 percent level if you have at least 2^{64} files.)

The obvious drawback to the shared folder method is that all users must have the same server access. In most cases, this means access to the same LAN or a robust corporate WAN. Remote users, even those with a reliable Internet connection, are excluded.

Another difficulty is the anonymous nature of the FDF files on the server. Although Acrobat can connect an online PDF with its respective FDF file(s), there is no efficient, external method for doing so. Identifying a specific FDF file can be done only by opening it in a text application, and then searching the File ("F") string. There is also no easy way to delete unused FDF files, which will eventually clutter up the server. A better selection of FDF parsing and management tools is clearly called for. (The SDK mentioned earlier does not deal with FDF files stored on a shared folder or a WebDAV server.)



WebDAV: all that glitters is not gold

Installing and configuring WebDAV was a difficult and frustrating task. The technology is still in its infancy. Numerous suppliers are converting their servers to this new standard, but not without problems. (A technical description of the installation and setup we used for our WebDAV testing, including the problems we experienced, can be found on SeyboldReports.com in the Web version of this article.)

In spite of these problems, we finally managed to configure a WebDAV server, physically located in the U.S., that could manage FDF files and relate them to PDF files on a separate Web server located in Germany.

It's important to note that, in this design, WebDAV is receiving FDF data files containing user annotations, not the actual content of the PDF. Other Adobe WebDAV plans, centering on its content-creation applications (InDesign, Photoshop, Illustrator, GoLive, LiveMotion, etc.), will use WebDAV in an entirely different manner. WebDAV-enabled products such as InScope and Adobe Studio will allow users to "check out" and "check in" various content files, as well as track file versions and a variety of job-related metadata. Acrobat is somewhat unique, in that the job content is static, while the comment metadata is dynamic.

An important advantage of WebDAV is that the implementation of this technology requires few system resources. This improves the overall behavior and performance of the Web server, which explains why many Web administrators are so fond of WebDAV. The nor-

That's easy for you to say! Once configured, the Web server and the WebDAV server act in tandem. One manages PDF files uploaded via FTP, while the other manages corresponding FDF files. Except for PDF uploading, all file transfer and viewing is handled via HTTP.

mal user of a Web server won't get a chance to experience the benefits of a comprehensive WebDAV-HTML environment until sometime around the end of 2001. But when that time comes, it will be Acrobat users who are best equipped to benefit.

Test results

Opening and working with online PDFs was strikingly similar to doing so locally. Pages opened relatively quickly (using an ISDN connection), and file navigation was identical to using Acrobat alone. Naturally, some features, such as Link creation, were not available. Pages could be rotated, but the rotated state could not be saved. (Saving an online PDF copies the file to the local drive.) As with Acrobat itself, users can zoom up to 1,600 percent, revealing vector-quality detail. Rendering of vector objects is

vastly superior to other online systems that rely on bitmap rendering.

Testing the various Acrobat comments was also satisfactory. Multiple users added a variety of annotations, which were successfully preserved using both the shared folder and WebDAV methods.

The annotations are time-stamped and identified with the author's name. However, outside of Acrobat itself, there are no tools (yet) that can extract meaningful data from the stored FDF files and convey it, via XML or otherwise, to an external management or workflow system. During our testing, we learned of several companies in the process of integrating Acrobat and WebDAV, but none were prepared to share their findings. (In a future article, we will detail some case studies, as well as discuss some upcoming products that relate to online PDF collaboration.)

Microsoft's Impact

Although we did not test any Microsoft applications, they will have a significant impact on the use of online PDFs, especially in the corporate enterprise. Since the majority of Acrobat users are not involved in traditional graphic arts but are part of law firms, insurance agencies, government bodies and so on, we can expect Microsoft's and Adobe's Web and collaboration tools to develop a decidedly "corporate" aspect.

Web Discussions

Microsoft's widely used Office Server products offer the possibility of exchanging and requesting documents within an Intranet. The basis for this is the "Microsoft Web Discussions Repository" available in various Microsoft server applications. This technology, available only for the Windows environment, requires the client to be running a recent operating system (Windows 95 OSR 2.0, Windows 98, NT 4.0, Windows ME or Windows 2000) and Microsoft Office 2000. On the server side, an equally up-to-date server package meeting the latest requirements is necessary. In addition, the Microsoft Management Console or the Personal Web Manager must be installed. If these requirements are met, the user can, given the appropriate network and server permissions, select "Web Discussions" in Acrobat's "Online Comments" Preferences, and can then share PDF files with colleagues via the Web browser. The browser must be Microsoft's Internet Explorer, and the discussion server must be appropriately configured.

Adobe uses a simple and direct method of sending commands to the server: the appropriate commands are appended to the URL of the PDF file and are interpreted by the server. But the functionality of this implementation is very limited. For example, binary data has to be converted and transferred to the server in 64K blocks. Because the data is broken up into strings and small blocks of data, occasionally it gets stored on the server in the wrong order.

Database Issues: SQL and ODBC

The incorporation of an SQL database, such as Microsoft SQL Server, Oracle 7 or Oracle 8i, expands the "virtual teamwork" possible with Acrobat's online annotation into a real team effort. The ADBC (Acrobat Database Connectivity) technology that makes its debut in version 5.0 allows Acrobat to communicate with an SQL database via JavaScript. ADBC, based on Microsoft's ODBC, permits easy data exchange. Unfortunately, however, it does not support Macintosh users.

The system requirements for clients using the database repository are identical to those for clients of the Microsoft Web Discussion Server, but the server must be set up with an SQL database and the client must install the appropriate driver. "Repository," in this context, means an archive containing all document annotations and monitored by a server-side version-control system. The required driver can be selected during installation of either Office 2000 or the Microsoft SQL client. If an Oracle database is used, an Oracle license for each client is required.

It is not a trivial task for the system administrator to configure the preferences correctly. After the preferences have been set in Acrobat, the administrator must define the corresponding server and database settings. From then on, all annotations will be stored in the "CollabDB" database, permitting team members to access them. Each document is kept as a separate table, and each annotation is stored hierarchically as a row in the table. The notable aspect of this is that each annotation is stored in the database as JavaScript. The size limit on annotations is around 2 GB, which should be plenty for normal uses. To make implementing the server a bit easier, Adobe offers an FDF that automatically defines the database settings. The necessary files for this are on the CD-ROM of the full version of Acrobat.

The use of PDF in databases is still in its infancy. Several small developers are in the process of developing solutions based on Oracle, with release anticipated in the fall of 2001 or early in 2002.

TSR

Signatures & Forms

When Acrobat 5 was released, online digital signatures were presented as an integral part of a collaborative PDF workflow. We assumed, erroneously, that these would be handled in the same manner as online comments—via shared folders or WebDAV. Instead, we learned that online signatures had to be implemented in a more cumbersome way, which unfortunately we were unable to test.

Due to legal requirements, a digital signature can be applied only after the whole document is saved locally and the encrypted signature embedded in the PDF file. To accomplish this online, the entire file must be downloaded, temporarily stored on a local drive, and then uploaded to the Web server via a “submit” process, generally through a CGI script. The WebDAV approach used for comments could not be used, because it would allow the signature and its PDF file to be separate—potentially allowing substitution and negating the signature’s legally binding value.

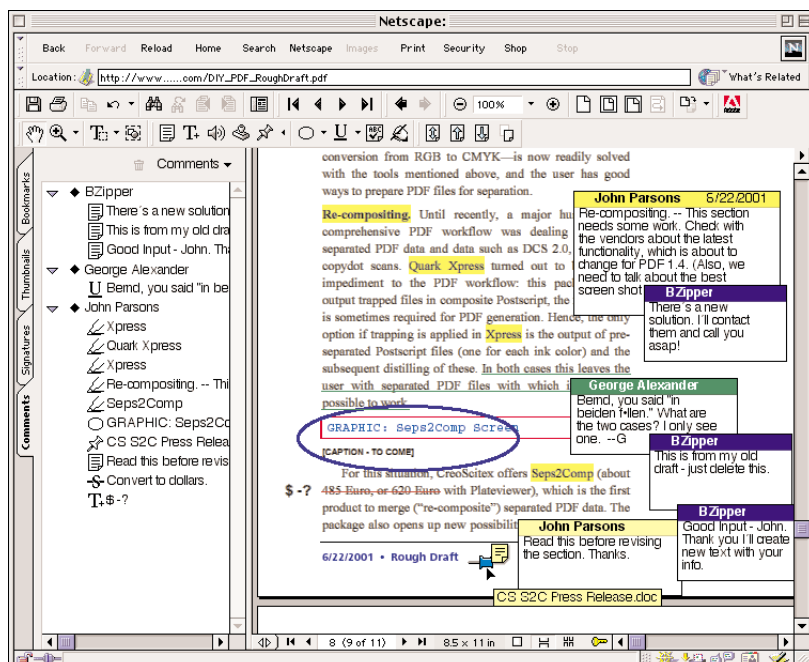
We acknowledge the necessity of this process, especially for legal contracts and other sensitive PDF files. However, we are also disappointed, since the requirement to download and upload the PDF—especially if multiple signatures are required—will be extremely burdensome to graphic arts users. PDFs from these users are often very large, containing full-resolution images. An online PDF “contract proof,” for example, could easily be too large for practical use of digital signatures.

What about forms? The use of PDF forms data is also an intriguing aspect of an online workflow, although it too is separate from the WebDAV approach. Online PDFs could conceivably have a standard “cover page,” including fields and check boxes, that could be integrated with a general workflow system. Although we did not test this feature, it is still quite feasible, according to several sources at Adobe. The “submit” process for PDF forms data is well established for Web-resident documents, and will be expanded by Adobe’s new \$39 application, Approval, slated for release on August 27.

We expect that a number of graphic-arts developers and integrators will add forms capabilities to their online PDF offerings. **TSR**

Conclusion

In general, we were pleased with the results of our tests, and we can certainly see the potential for online comments—not only in graphic arts and publishing, but wherever documents need to be created, reviewed and revised by a distributed workgroup. Dealing with a single PDF file and combining it with multiple comments (in a straightforward Web environment) is attractive, once the initial configuration problems are overcome. The decreased volume of transmitted data and the potential for centralized management makes the process well worth looking at.



We encountered several problems with this workflow. One was the lack of practical administrative tools, although we believe this will be remedied shortly by third-party developers. Another issue was our disappointment—from the graphic arts perspective—over the file-transfer requirements of using digital signatures online. (See the “Signatures & Forms” sidebar.) Acrobat’s commenting tools, both online and in the stand-alone product, were sometimes awkward. (As editors, we found that the electronic tools often required more steps than the manual markup process, or contained annoying limitations such as the inability to highlight a whole paragraph in a multi-column layout.)

One significant barrier to this workflow is price. The full \$249 version of Acrobat 5 is required for users to view online comments, even if the user does not need to make comments. Neither the free Reader nor the recently introduced Approval product (\$39) can view WebDAV-hosted online comments. For users who need only to view and approve comments, we would strongly suggest a read-only WebDAV feature be added to the next version of Approval.

Acrobat’s online capabilities have increased the potential for collaborative workflow, using a common file format and (at least for comments) the obvious strengths and popularity of WebDAV. As tantalizing as this potential is, however, we feel the process is unfinished, and that practical solutions are still to come. **TSR**

About the Authors

Bernd Zipper (zipper@zipcon.de) is a business and technology consultant specializing in PDF, new media and cross-media publishing projects. He is president of ZIPCON NewMedia and Prepress Consultinggesellschaft mbH (www.zipcon.de) in Essen, Germany.

John Parsons is Senior West Coast Editor for *The Seybold Report*.

Notable. Successive users can add a wide variety of comments to online PDF files via Acrobat 5. Comments can be sorted by various criteria. The resulting file may also be downloaded along with its comments, which can be summarized for final review.