

# Adobe® Acrobat® Capture® 3.0 and Third-Party Solutions

**Adobe Acrobat Capture 3.0**

**3**

**Third-party tools: Expanding the Acrobat Capture solution**

**6**

**Acrobat Capture and Adobe ePaper® Solutions**

**9**

**Appendix**

**10**

In the closing moments of the popular 1982 American film *Raiders of the Lost Ark*, a uniformed guard wheels a pallet bearing a precious religious artifact—the Lost Ark of the Covenant—into a warehouse. The camera pulls back to reveal that the storage facility is a huge, cavernous place, stacked to the ceiling with thousands of creaking wooden boxes.

For many large corporations, this scene is an analogy of their own paper-based archiving system. Paper documents—sometimes accumulated over decades—fill file cabinet after file cabinet, burying important records. These documents contain all types of information produced before the desktop computer revolution: data output from legacy mainframe applications, legal documents, technical drawings and manuals, blueprints, and typewritten documents. Retrieving information contained in these documents can be a tedious, time-consuming process. Then, once the information is unearthed, distributing it to others—particularly if a large number of people require access to it—becomes another logistical problem because the content is in paper, rather than digital, form.

Businesses seeking to bring paper-based information into the digital workflows of today's e-business environment find an ideal solution in Adobe Acrobat Capture 3.0 software. This professional production tool helps users optimize all kinds of paper-based documents and publish them electronically via the Web, intranets, and other media. By converting all the historical and current paper documents in an enterprise into the universally readable, fully searchable Adobe Portable Document Format (PDF), Acrobat Capture software prepares business information for today's e-business environments.

Businesses can also combine the power of Acrobat Capture with products and services delivered by several third-party vendors for a total document management solution. Providing everything from customized front-end software and high-speed scanning technologies to end-to-end document management software suites, strategic partners work closely with Adobe to make the paper-to-Web transition easier for enterprises in every industry worldwide.

### **Adobe Acrobat Capture 3.0**

The newest release of Adobe Acrobat Capture 3.0 software incorporates feedback from Adobe customers, combined with the latest Acrobat technologies, to offer the most robust paper-to-digital application to date. Fundamentally, Acrobat Capture works with scanning devices to convert large volumes of paper documents into universally readable, fully searchable Adobe PDF files as well as HTML files. Enterprises taking advantage of the latest e-business practices can then make their converted documents available via an intranet, Web site, CD-ROM, or other electronic distribution method.

Especially appropriate for conversion are documents such as technical manuals, standard operating procedures (SOPs), contracts, market research reports, and other intellectual property that was previously accessible to an enterprise only on paper. Although any large

organization dealing with paper documents can benefit from the power of Acrobat Capture, specific sectors that effectively use Acrobat Capture to manage paper-based information in their electronic workflows include the financial, legal, insurance, and pharmaceutical industries as well as the government. Acrobat Capture can also be an integral part of the document workflow in any discrete manufacturing process, from aeronautics to automobile production.

In all these industries, immediate access to paper-based information has critical competitive importance. Many technologies, including scanner and OCR technologies, are available for turning paper information into TIFF and other image-compression-based files. But for enterprises that want to keep up with the exploding pace of Internet and intranet use, using Acrobat Capture to convert paper into industry-standard Adobe PDF offers important benefits:

- *Universal viewing in heterogeneous environments.* Documents converted to Adobe PDF can be read anytime, anyplace, on all major computing platforms with the free and broadly distributed Adobe Acrobat Reader™ software. It doesn't matter what application was used to author the document, or even if it was drawn by hand (blueprints, handwritten mathematical notations, and so on). Regardless of the document's origin, colleagues within an enterprise and outside partners can access the information they need via intranet or extranet without proprietary viewing software.
- *Complete visual fidelity.* Adobe PDF files capture the precise look and feel of any printed document, including fonts, color, and layout. Adobe PDF technology makes it possible to capture sophisticated technical drawings in every detail, and ensure that legal documents, such as contracts, retain their fidelity. Visually rich documents—from presentations to marketing collateral to nondigital photographs—can be made part of any knowledge base and shared instantly worldwide.
- *Powerful indexing and searching.* With its robust support for hypertext links and metadata, Adobe PDF allows indexing and searching from a variety of document management systems. Setting Adobe PDF distinctly apart from image-based formats such as TIFF is its patented PDF Searchable Image capability. When this option is selected, Acrobat Capture saves a compound version of the document with a “scanned” image of the text and graphics laid over hidden searchable text. This text can then be rapidly indexed by intranet search engines or document management indexing engines and quickly retrieved by users.
- *Easy navigation through long documents.* Adobe PDF files let you search on any word in a document and see the searched information in context on the page. You can also search for a group of words or a combination of words. And, unlike TIFF files, Adobe PDF files can

contain navigation information, such as hypertext links for tables of contents, indexes, and URLs, which lets Adobe PDF files offer both easy navigation through multipage documents and links to other documents.

- *Fast download times.* The compactness of Adobe PDF files makes them ideal for electronic publication. Multipage documents published on the Web in Adobe PDF means you don't have to wait for the entire file to download before starting to read it. Adobe's support for byte serving, which is incorporated into the Adobe Acrobat Reader software, gives you access to online documents one page at a time. If you want to view a 10-page Adobe PDF file on the Web, for example, click on the document and the first page appears immediately. Meanwhile, in the background, the rest of the document is downloaded.

The general benefits of Adobe PDF are extended by the powerful feature set found in Acrobat Capture 3.0:

- *Robust Software Development Kit (SDK).* The latest release of Acrobat Capture can be more readily integrated with existing enterprise document applications. In-house IT professionals can create customized agents that can be dropped into any Acrobat Capture workflow. And systems administrators can take advantage of new features with watched folders. These features let you customize job files, which enables many of the third-party solutions described in this paper to integrate with other applications quickly and easily.
- *Scalability.* The scalable architecture of Acrobat Capture 3.0 Cluster Edition enables users to process a larger number of documents than ever before. The Cluster Edition can be used in one of two ways. It can be combined with the included Acrobat Capture Assistant software to allow as many as five additional workstations to offload labor-intensive tasks that require human intervention (such as zoning, OCR review, or scanning). Or, additional licenses of Cluster Edition can be installed to run compute-intensive processes (such as OCR and content recognition) on multiple workstations simultaneously. This can be achieved either by purchasing multiple copies of the single-processor edition or by taking advantage of dual- or quad-processor licensing.
- *Expanded content recognition options.* Acrobat Capture 3.0 incorporates a greatly improved recognition engine with enhanced font, table, image, and color recognition. In addition, Acrobat Capture automatically creates intradocument links—tables of contents with nested bookmarks, cross-references, indexes, e-mail links, and URLs—to make documents easy to navigate. With zoning, users can define areas of a page to OCR, to treat as an image, to ignore, or to OCR as an Adobe PDF keyword. As described later, third-party solutions fully exploit the OCR enhancements in multiple, customizable ways.
- *ODMA support.* The in-the-box ODMA support of Acrobat Capture allows the product to integrate directly with more than 20 document management systems, including those from FileNet, OpenText, Intranet Solutions, and Lotus' Domino.doc.

### **Third-party tools: Expanding the Acrobat Capture solution**

In many low-volume environments, using Adobe Acrobat Capture 3.0 with a desktop scanner is sufficient. Individual users can convert paper documents on an as-needed basis and add them to workgroup intranets, compound documents, or small databases. However, in other environments, where legacy paper documents are measured in reams rather than pages, a more sophisticated system might be required. Creating large knowledge bases that can be accessed from within the organization via an enterprisewide intranet or from outside (by external offices, partners, and the like) via an extranet requires a well-considered strategy. Therefore, enterprises planning to convert thousands, even millions, of paper documents to Adobe PDF should consider a comprehensive document management system tailored to fit their IT environment.

There are two main categories of third-party solutions that can be used with Acrobat Capture to maximize an enterprise's IT investment. *Front-end imaging solutions* make sure that paper documents that are converted to Adobe PDF are not only scanned in quickly and accurately but also maintain full data integrity. Then, after files have been readied for electronic management and distribution, a number of *back-end document management applications* can help the document management or IT professional with the complex task of making sure that the converted documents—and the data contained in them—are available to everyone who needs them.

#### **Front-end imaging solutions**

Several hardware products and software applications work with Acrobat Capture to pull large quantities of paper documents out of filing cabinets and place them—as Adobe PDF documents—into a corporate repository. (Information on many of these companies can be found in the appendix.)

- *High-speed scanning.* The batch conversion of volumes of paper documents requires scanning speeds far above that of a typical desktop scanner. A number of vendors address the need for high-speed scanning with hardware and software solutions. High-speed scanners from providers such as Fujitsu and Bell & Howell, as well as plug-and-play scanner accelerator cards from companies such as Kofax, work together with sheet feeders to permit scanning of more than 100 pages per minute. IT professionals can work with Acrobat Capture service providers such as Lason and IKON Office Solutions to install such systems in their enterprises. They can also outsource at least the initial volume scanning to document scanning service providers.
- *Image clean-up.* Third-party software products from Kofax and TMS/Sequoia can clean up poor quality images within Acrobat Capture 3.0. Various methods are used to “clean” images, improving data recognition and ensuring higher scanning accuracy. Along with higher data integrity, cleaner images equal reduced file sizes for better network throughput and lower storage requirements.

- *Advanced document/data recognition.* Processing modules within certain software applications from Kofax, ActionPoint Software, and Captiva Software build upon the OCR features found in Acrobat Capture to also recognize handprinted characters (Intelligent Character Recognition or ICR), mark sense zones (Optical Mark Recognition or OMR), bar codes, and patch codes.
- *Improved data integrity through intelligent forms.* Specialized applications, such as PDF+Forms from Cardiff Software, extend the forms functionality of Acrobat Capture. With such applications, data can be collected from paper-based forms and then verified and exported into data-driven applications. Businesses realize new levels of data integrity, and information that was previously paper-bound becomes instantly accessible for management and customer-service decisions. Documents optimized with PDF+Forms can even be easily linked to XML-driven databases for use in the latest e-commerce applications. For highly complex forms environments, FormWare from Captiva Software provides specialized workflow capabilities to managers who need to rapidly identify processing problems, prioritize jobs, and optimize throughput.
- *Expanded quality control.* Certain applications, from vendors such as ActionPoint Software, provide a variety of integrated quality control mechanisms. For example, automatic QA routines include image clean-up. Still another quality control feature is the program's ability to detect images that fall below a certain quality level. These images are then queued for rescan. Once processing has gone past these behind-the-scenes controls, documents can be viewed page by page, manually indexed, or viewed as thumbnails by QA specialists either in-house or at a service bureau.
- *Extensive scripting capabilities.* Using their own SDK with Acrobat Capture 3.0, many third-party software solutions vendors, including Kofax and ActionPoint Software, deliver customized scripting capabilities to systems administrators. Different software applications allow field validation, database look-up, field formatting, enhanced recognition, and tight integration with outside processes.
- *Data export.* Software solutions from Kofax and ActionPoint Software prepare data extracted from paper-based information for export to workflow, document imaging, or document management systems. Applications extract the data using one or more key entry, OCR, or zone recognition technologies, and then review, validate, and correct the data before merging it into existing data stores.

### **Back-end document management systems**

Once these front-end solutions have worked with Acrobat Capture to convert paper-based information to Adobe PDF, a host of other back-end document management systems help businesses manage, track, and make sense of the newly accessible information. (Information on many of these companies can be found in the appendix.)

- *Business knowledge management.* On the broadest level, document management systems are managing the important business knowledge that lets an enterprise compete at the highest possible levels. Once paper documents have been converted to Adobe PDF, they can be routed, archived, made available on Web sites and intranets, and secured by document management systems in several ways. Some applications—such as LiveLink from OpenText—enable the creation of knowledge management portals where users can share and access Adobe PDF-based information, track document workflows, and manage projects. Panagon software from FileNET, with its extensible, modular architecture, supports managing Adobe PDF files through its Content Services component. Documentum 4i, from Documentum, aggregates content—including Adobe PDF files produced from Acrobat Capture—according to business function to keep knowledge workers continually informed.
- *Automated management of controlled documents.* Some document management systems are ideally suited for the kinds of documents that Adobe PDF captures best: controlled documents. Comprising regulatory documents such as SOPs, material safety data sheets, packaging specifications, and engineering drawings, this class of valuable corporate information must retain precise visual fidelity when converted to electronic form while remaining fully searchable and portable. DocControl Manager from Documentum is one application that was designed to streamline managing these documents. Several pharmaceutical companies and the U.S. Food and Drug Administration (FDA) have made it the standard automated solution for managing new drug applications.
- *Document collaboration workflows.* Several document management systems use the universal portability and support for annotations found in Adobe PDF to create enterprisewide document collaboration workflows. Common mark-up and reviewing tools include concurrent graphical redlining, thumbnails, grayscale, rotation, and zoom. Users can apply watermarks, proofreader's marks, and digital signatures. InfoPortal Team OnLine, a joint Web-hosted effort between CoVia Technologies and Adobe, lets teams collaborate on documents in real time over the Web. The DocViewer application from Documentum treats each annotation as a separate object, with its own attributes and security rules. LiveLink from OpenText adds threaded discussions to the collaborative process.
- *Expanded viewing options.* Some document management solutions, such as Intra.doc! from IntraNet Solutions, let enterprise users access the information found in Adobe PDF documents in multiple ways. The Adobe PDF-to-HTML conversion capability found in Intra.doc! lets enterprises deliver the content optimized by Acrobat Capture and complementary technologies as HTML content that is viewable through Web browsers. This capability addresses environments where Acrobat Reader is not installed on every desktop and allows seamless content distribution to partners outside the company's extranet who may not have Acrobat Reader.

### **Acrobat Capture and Adobe ePaper Solutions**

For large corporations trying to shift operations into Internet speed, the transition to new e-business practices is rarely easy. Updating office workflows and computing environments can consume IT resources rapidly, and keeping up with the latest technologies can be a full-time job for business managers—and that’s just dealing with electronic information. The prospect of having to simultaneously manage volumes of legacy documents—paper-based information that lies like buried treasure in many organizations—can seem impossible. Making this information available to everyone who needs it may require complex solutions. However, Adobe Acrobat Capture 3.0, combined with products and services from third-party vendors, can help.

Adobe Acrobat Capture is just one of several Adobe ePaper Solutions that are designed to aid enterprises in their quest to manage paper and electronic information in today’s complex computing environments. Adobe provides powerful, integrated software tools for enterprises incorporating the latest e-business practices with Adobe Acrobat, which lets you create and interact with Adobe PDF files; and Acrobat Business Tools for workgroup collaboration; and Acrobat Messenger™, which replaces outbound faxing and delivers documents digitally. For more information on Adobe Acrobat Capture, visit [www.adobe.com/acrobat/capture](http://www.adobe.com/acrobat/capture). For more information on other Adobe ePaper Solutions, visit [www.adobe.com/epaper](http://www.adobe.com/epaper). Along with product information, you’ll find case studies, Web-hosted services, and timely articles dealing with topics of interest to enterprises managing paper and electronic information in today’s Internet-driven economy.

## Appendix

The Web sites of our third-party Acrobat Capture partners provide a wealth of information regarding front-end imaging solutions, back-end document management solutions, and other complementary technologies:

ActionPoint Software—[www.actionpoint.com](http://www.actionpoint.com)

Captiva Software—[www.captivasoftware.com](http://www.captivasoftware.com)

Cardiff Software—[www.cardiff.com](http://www.cardiff.com)

CoVia Technologies—[www.covia.com](http://www.covia.com)

Documentum—[www.documentum.com](http://www.documentum.com)

FileNET—[www.filenet.com](http://www.filenet.com)

Fujitsu Computer Products of America—[www.fujitsu.com](http://www.fujitsu.com)

IKON Office Solutions—[www.ikon.com](http://www.ikon.com)

IntraNet Solutions—[www.intranetsolutions.com](http://www.intranetsolutions.com)

Kofax—[www.kofax.com](http://www.kofax.com)

LASON—[www.lason.com](http://www.lason.com)

OpenText—[www.opentext.com](http://www.opentext.com)

TMS/Sequoia—[www.tmsinc.com](http://www.tmsinc.com)



Adobe Systems Incorporated • 345 Park Avenue, San Jose, CA 95110-2704 USA • [www.adobe.com](http://www.adobe.com)

Adobe, the Adobe logo, Acrobat, Acrobat Capture, Acrobat Messenger, Acrobat Reader, and ePaper are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States and/or other countries. All other trademarks are the property of their respective owners.

© 2000 Adobe Systems Incorporated. All rights reserved. Printed in the USA.

BC1625 8/00

