

a d v a n c e d

n a v i g a t i o n  
f o r

s u p e r i o r

i n f o r m a t i o n

a c c e s s

chapter

nine

The main point cannot be over-emphasized: the most important advantage of digital documents over paper documents is the *superior access to the information* in the documents. Depending on the nature of the information and the user's needs in handling the information, there are many techniques that add greatly to the value of collection. The publisher can *empower* the user through more *efficient organization* in published digital documents.

The following are currently available features available to expand the navigation capabilities of PDF files.

The point of this section is that documents must at least *equal*, but should intend to *better* the user's access to information in digital collections as opposed to libraries of paper. The plug-in tools listed here are remarkable for the functionality they provide for publishing and researching your information. A dozen are included free with Acrobat 3; others are sold by third-party developers.

Some plug-ins are included on the CD that ships with Acrobat. Adobe also offers a frequently updated collection of free plug-ins that enhance Acrobat Exchange:

## Acrobat 3 Plug-Ins

<http://www.adobe.com/Acrobat/Plug-Ins>

### AutoClose

This option is similar in function to a Web browser in the sense that it closes pages that have been accessed earlier and are still stored in memory. When the 11th document is opened, the last accessed document is closed. It's *always important to conserve RAM*, storage and communication bandwidth.

### AutoIndex

The AutoIndex function immediately opens an associated Catalog-created index when a page with this feature is accessed. It's based on the thinking that if the user has found a hit in the Getting Started guide, he is likely to soon need information in the User Manual and Tutorial and other files. This option immediately fires up the appropriate index and allows easy navigation to related information.

### Capture

Capture performs multi-level recognition of page images, converting not only the text content, but also the appearance of the page. This process involves substituting fonts and font attributes to match the source document, as well as rebuilding the document in the exact dimensions of the original page.

tip

Capture offers two primary output options:

PDF Normal reproduces the page with recognition;

PDF Image & Text offers original image with hidden text, which allows content searching.

## EPS Links

A collection of recipes and examples is available for merging traditional design logos and elements with links.

## ExecMenu

Digital documents don't get much more dynamic than this. This is where sound, movie and all of the other multimedia capability enter the PDF world.

example

With the Execute Menu feature enabled, a PDF link to a virtual reality world could be viewed with Silicon Graphics Cosmo shareware VRML.

<http://www.sgi.com/Products/cosmo/>

A site that will always be dedicated to faster, richer, better online experiences is the virtual reality guys at:

Virtual Reality Modeling Language (VRML) allows representation of objects in interactive 3-D. In a technical manual, for instance, in addition to explanatory words and graphics, virtual working models of mechanical systems can offer a far richer experience.

A VRML lesson on replacing an electronic component might begin with a photo or simple schematic of an aircraft. Choosing a particular part would cause that part of the plane to glow, and the user could easily "walk under" that spot. "Looking up" a couple of screws might appear, and when you touch them, the mouse becomes a screwdriver and opens the access hatch. When you "look into" the hatch, you have a *better than video view* because you can see through physical objects to see how they might be connected on the other side. Through a well-designed VR interface, entire procedures can be practiced and referred to even while doing the job.

The U.S. Air Force is already in prototype development of such systems, with hip pack computers and displays built into lightweight goggles. A crewman could walk out to the aircraft, triple-check everything through the VR-assisted procedures, and confidently and efficiently perform complex tasks.

# Forms

tip

The Forms feature of Acrobat 3 offers significant efficiency gains over HTML Forms because only the variable data is transmitted back to the server. Compared with HTML, which transmits back the entire form, this Acrobat technique is demonstrably more effective than previous means.

The Forms functionality in Acrobat 3 includes a number of classic design elements and also provides for online creation of dynamic files with elements served up from the Web server. This feature is demonstrated at Adobe's Web site. It is fairly amazing that custom forms can be built online based on the initial responses of the user.

## Import Image

This is the doorway for many popular images to enter into the bigger world of PDF files. The files being imported are all widely used in their particular fields, but upon becoming PDF, they achieve universal status.

The importable formats include BMP, CompuServe GIF, PCX and TIF. These formats represent, respectively:

BMP: The Windows bitmap format

GIF: The most widely distributed online image format universally incorporated in HTML

PCX: The original Paintbrush and Paint formats

TIF: The industrial standard of fax and document imaging

tip

The Acrobat viewer technology is in widespread use, exceeded only by Web browsers in number of users. These numbers include many who never use the Web at all.

By importing a traditional image into PDF format, the image becomes universally accessible on all platforms, hardware, software, online and CD media, because it is the first rich viable universal format.

# Movie

This is a Windows version of the Apple QuickTime movie viewer, the one used in Myst and other popular multimedia publications. Using the Movie pull-down tool, it is very easy to create a link to a QuickTime or AVI video file. By activating the link, the user calls up the externally stored movie and QuickTime plays it. QuickTime offers a range of effects to size or even distort the movie presentation.

---

**QuickTime movies are very large video files. The Weezer “Happy Days” video on the Windows 95 FunStuff disk has a running time of about four minutes and a file size of 30 MB. New users should always be informed of file size.**



**For this reason, video is always stored external to a PDF file, so extra care must be used in maintaining these links.**

# OLE Server

This feature brings PDF documents into the mainstream desktop applications by allowing *inline viewing* of Acrobat files. Lucky users of *Lotus Notes* might find rich PDF documents online and can immediately view, search or print them in their full-featured presentation. Microsoft Access applications can query a database and easily retrieve and view PDF contents.

# Optimizer

The key technical breakthrough in Acrobat 3 is the ability to Optimize PDF files. For Web-accessible files, the most dramatic feature is the ability to take advantage of byte-serving Web servers to deliver only certain pages rather than entire documents. This can provide great economies of transmission speed and response time for online users.

Repetitive elements in pages are consolidated and reused, such as background images, text and line art. This leads to significant reduction in file sizes, the key to speedy transmission of information.

To optimize a file, just click the **Optimize** check box in the **Save As** dialog screen.

Or, you can use the **Batch Optimize** command to automatically **Optimize** an entire collection of PDF files in a folder and its subfolders.

## PDF Type Utility

This is a Mac applet that makes Mac-created Acrobat files more accessible to Windows and UNIX Acrobat users.

## Scan

The Scan module includes the ISIS drivers for most popular scanners, as well as the TWAIN interface. ISIS drivers can be selected from the pull-down list and provide instant functionality for scanners as varied as the inexpensive Fujitsu ScanPartner 10 and the 120-page-per-minute Fujitsu 3099 family of duplex scanners.

### **ISIS Scanner Drivers**

ISIS Scanner Drivers are software creations of Pixel Translations, which was acquired by Cornerstone Imaging in 1995. As more and more scanners are connected by standard SCSI connections rather than the traditional proprietary scanner interface cards, most applications that require scanning now include the ISIS drivers. This means that the user can simply choose from a pull-down menu to select a scanner, and the ISIS driver will allow the application to run the scanner.

A driver is a piece of software that translates all of the functions of the scanner to a common set of controls that can be exercised as buttons, bars and so on in the application software. This means that Acrobat Capture and Watermark and other packages that include the drivers offer freedom of choice on scanners.

# Search

The Search features available in PDF collections are extensively covered in Chapters 7 and 12.

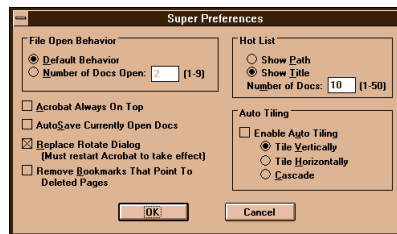
# SuperCrop

SuperCrop provides the ability to eliminate unnecessary white space around PDF documents to improve viewing characteristics. This new crop tool allows rubber-band box selection of the area to be cropped, which is considerably easier to use than the Standard Crop arrow buttons. Users of Acrobat Exchange can also automatically crop to the bounding box of the page.

# SuperPrefs

The SuperPrefs plug-in allows the user to tailor the function of Acrobat Reader to meet specific requirements. The options include:

- File Open Behavior
- Acrobat Always on Top
- AutoSave Currently Open Docs
- Replace Rotate Dialog: Replaces standard rotate dialog to include 180-degree rotation and cancel
- CleanUp Bookmarks: Removes Bookmarks referring to pages that are deleted
- Hot List: A Menu Item listing your favorite files
- Auto Tiling: Choose your layout of multiple open files



SuperPrefs creates a new set of preferences for user convenience.

# TouchUp

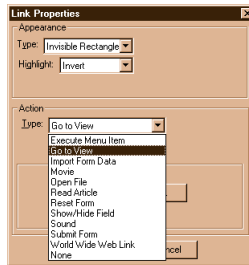
TouchUp is the tool for editing PDF single-line character strings, ideal for editing the results of Capture-created documents. This allows users to clean up and modify the output of the recognition process by modifying text, attributes and other components.

# WebLink

A WebLink is the PDF equivalent of an HTML hyperlink, and WebLinks function across the World Wide Web in the same way hyperlinks do on the Web.

The WebLink plug-in allows the user to add World Wide Web hyperlinks to PDF documents. The Acrobat Exchange Link tool is extended with this plug-in so that a URL can be specified as the target document.

When a WebLink is clicked, the user's Web browser is launched if it's not already running, and the Browser follows the link. WebLink is compatible with Netscape Navigator, Microsoft Internet Information Server, Spyglass Mosaic and Quarterdeck Mosaic.



The user can specify the appearance of links and specify one of a broad array of actions that a link will perform.



# Re:Mark for Digital Annotations

The most obvious difference between digital documents and paper documents is that *it's a lot harder to write on digital documents as opposed to paper documents*. A paper page can be changed and emphasized with a wide range of desk drawer tools.

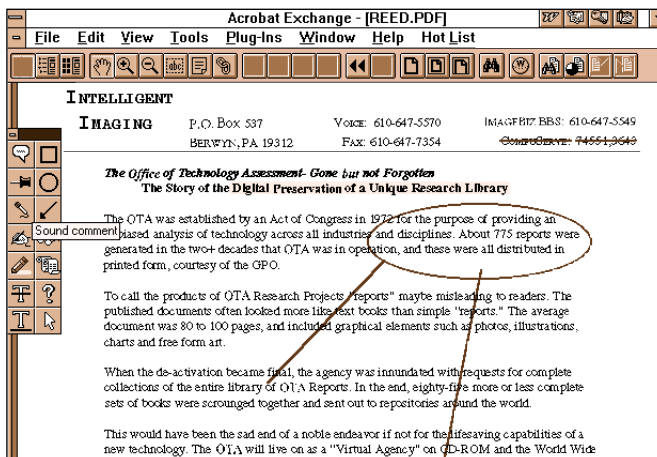
Handwritten notes and drawings can be scribbled with any number of pens and pencils, areas can be circled and pointed at in bright redlining, words or sections can be highlighted with broad strokes of yellow.

Since the introduction of Magic Markers by Carter's Ink in the 1960s, business documents have been augmented, annotated, highlighted, and updated with a variety of colors and writing implements. People have become accustomed to a thrilling array of annotation tools, from highlighters to spray paint.

Digital documents often lack any means for such reader input. Simple e-mail, for example, can be easily copied and quoted, but direct *paper-like comments* are very limited.

Re:Mark, published by Ambia, brings these *markup techniques*, tried-and-true and widely relied upon, to digital PDF documents. The original business process relied on all of the pen-based means of document comments and emphasis. All of these markup features were *required* to make the digital documents match the functionality of the paper documents.

Digital movie-making tools can be used to add visual content to PDF files. Adobe Premiere is designed for digital video editing and production, an excellent enhancement tool for PDF information collections. Searchable, rich multimedia collections on the Web are simple to create.



Acrobat users have a wide range of tools for adding emphasis and comments.

**Now that it's easy to move rich digital pages instead of archaic ASCII e-mail pages, users need conventional markup tools. Digital pages are here, and they need highlighters and sticky notes!**

## Re:Mark's Markup Capabilities

As described above, this set of tools allows the publisher and even the readers of documents to add dynamic input. With paper documents, this number of easily understood comments would be impossible. Markup tools like this allow groups to use digital PDF documents far more dynamically than any paper system. Groupware applications are enabled by this simple ability to add rich "paper-style" markup features to digital documents.

In a way that no paper document could equal, many unique and individual user comments and annotations can be color-coded for immediate identification of the source of the comment. In this digital document, the reader has selective access to *just some* or *all of* the added comments.

The notes should be minimized for clear viewing of the original document, and they are clicked for reading the contents of the notes. The markups can be rendered into the PDF documents so that the commentary can be shared without the need for all users to have Acrobat Exchange.

### Underline

The oldest form of text emphasis is to underline words. Underline is one of the two text emphasis features shared by many handwritten styles and most early machine printers, like typewriters.

### Strike-Out

The other original text emphasis feature is to strike out words. Just as underline promoted the importance of certain words and lines, strike-out deletes certain words and lines from consideration. Even cuneiform in ancient Babylon used strike-out on mistakes and edits.

## Highlight

Though the inspiration for highlighting may hark back to the miraculous dawn of illuminated manuscripts, the function as illustrated here is closer to Magic Marker highlighters.

## Comment Pop-Up

The author of any of these Re:Mark items can insert a simple text comment about the markup itself. For example, the author of a Highlight or Underline may offer an explanation for the emphasis added to the original document.

## Draw On The Digital Document

This ink-drawing tool mimics the simple ability to write on a piece of paper. Of course, if this feature is going to be used extensively, a pen tablet would be an excellent input as it is difficult to write with a mouse. However, even a mouse is sufficient for simple cross-outs or pointers that are used in text and page editing.

## File Attachments

The file attachment feature offers capabilities similar to MIME attachments in e-mail. Any file may be attached to the PDF file, including spreadsheets, word-processing documents, and pure binary documents such as bitmap images, sound and movies.

One of the great benefits of the way this feature works is that the attached file becomes part of the PDF file (except for video and other extremely large files). Since only one file needs to be shipped and retrieved, the danger of lost and missing files is avoided. The end user can then view the Acrobat file, and the attachment can be extracted and viewed in its native application.

For the storage, retrieval and transmission of compound file types, such as a PDF that contains word-processing or spreadsheet files or other binary types, *this method is equal to the best of the MIME-capable e-mail systems*. This *single-file format is simpler* but similar to the methods used to handle compound files in document-management systems.

## Multimedia Attachments

Sound annotation allows verbal comments to be incorporated in the document and played on the increasingly common multimedia workstations.

# Color-Coded User Input For Digital Personality

All comments and annotations are color-coded to represent the author of each note. In this way, the reader instantly recognizes the probable purpose and importance of a particular note based on previous knowledge of the author.

## Review

All of the above modifications and comments can be gathered together from many users. Comments can be filtered by author and annotation type, allowing quick consensus and input aggregation.

All of the input from many sources, identified by User and Note Type, can be viewed in one Master Document. The Master can be reviewed by one editor or shared with the entire group for dynamic access and development.

The contents and comments can be created on either a Mac or Windows platform, and the results are useful for users on either platform. All of the comments and annotations can be rendered into a standard PDF file, viewable with the standard Acrobat Reader.

## Extraction

The Re:Mark Copy Table to Clipboard feature allows the extraction of complex tables in spreadsheet format. This software analyzes the layout of the rows and columns and attempts to rebuild the structure as cells in a spreadsheet.

A PDF-to-RTF converter provides for the transformation of PDF files into the Rosetta stone of word-processing programs: Microsoft's Rich Text Format. The resulting RTF files are widely supported by competing programs and can usually be efficiently loaded into word-processing, spreadsheet and other applications.

## Compose

A recurring theme in this work is that digital documents should be better than paper documents in the way they deliver information. Given the oft-stated 500-year head start that paper books got on digital books, it is usually a good idea to try to reproduce the proven techniques of history in today's digital forms.

Given the fact that digital documents can be text searched and recalled en masse via broad retrieval commands, it would seem that superior access is a given with digital

documents. However, simple practice shows that the ancient book organization of documents is essential and critical to navigation in search of ideas.

Compose, published by Ambia, attempts to restore the organization and structure of paper documents to digital documents. Text searching is not enough; it is important to be able to navigate by page numbers, indexes and the table of contents.

## Automatic Bookmarks And TOC

A great technique called Bookmarks By Example searches for all particular format elements to identify chapter and section markings. By recognizing a combination of font, size and format that is used for a certain heading, multiple-level bookmarks can be automatically generated.

These bookmarks, based on repetitive markup of headings, can also generate an extended table of contents.

It must be noted that the "automatic" quality of this process is limited to the consistency in the source document. If an entire collection of documents shares the same style sheet, it will be possible to perform a highly automated conversion to an orderly digital library form.

If, on the other hand, the documents are inconsistent, this process will require hands-on correction and quality control. For example, files converted by Acrobat Capture may not always assign exactly the same point size or weight to a particular font. Such variations are common in scanned documents. In this case, quality control will be required to generate a predictable and reliable index and table of contents.

## Hyperlink Index And TOC To References

The index and table of contents of any document can, and perhaps should, be hyperlinked to the chapters and sections listed. The Page Linker uses page number recognition to automatically build hyperlinks.

This will give the reader the ability to use these traditional Finder Aids with digital efficiency. After locating a topic of interest in the index, a simple click will take the reader directly to the reference within the body text.

# Automatic Hyperlink Copying

Repetitive links throughout a document are very helpful to the document user, so they should be used generously in the publication of digital documents. The Copy Link tool allows links to be duplicated and populated throughout a long document, providing the utmost in navigation for the user with the most efficient effort by the publisher to add this valuable feature to the digital document.

The most common use of such repetitive links are simple “next page,” “previous page” and “table of contents” buttons, which allow users to quickly move through documents even if they are not familiar with Acrobat’s buttons.

## Link Auditor

The Auditor checks the entire PDF document and fixes broken bookmarks and links.

## Chain Linker

The Chain Linker creates links between every occurrence of a word or phrase. This is like the ultimate index, where every single instance or mention of a certain term is listed. Applied sparingly, this tool can be extremely effective in offering shortcuts to related information. Used profligately, this feature will be so tedious that users will never touch it and its index will just be excess baggage in the PDF file.

## Multi-Document Composition

This feature works just like the Master Document feature in word-processing packages in that it allows the user to build large documents from collections of smaller documents. The smaller documents may be input from several authors, collections of earlier published documents, or conglomerations of several input sources or news wires.

This feature also allows the page numbering of the original documents to be removed, and a common page-numbering scheme is built for the new, compound document. In addition, customizable page headers and footers can be added, which may include document name, date and time fields, page numbering and other components.

## Master Table Of Contents

Hundreds of individual documents can be linked in one title page so that an entire collection can be easily accessed.

# InfoFill

This tool is used to automatically populate the General Information Fields in collections of PDF documents. For example, a simple ASCII file of information can be merged into all of the Title, Subject, Author and Keywords fields. This capability may be exceptionally useful when documents that have already been indexed in another scheme are being converted to PDF and Acrobat Exchange database format.

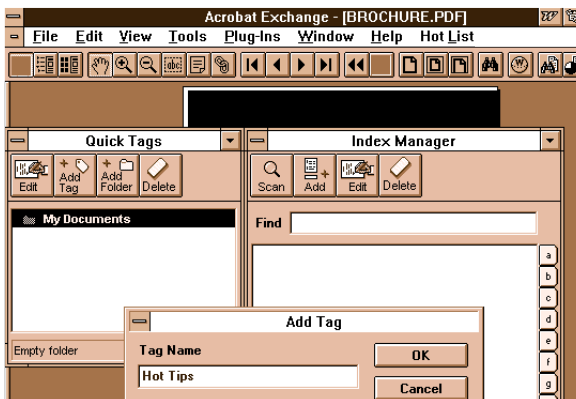
When used with the above Table of Contents Builder, very large numbers of PDF documents can be quickly linked.

# Navigate By Title And Author

This feature allows users of large collections to browse by actual document Title and Author rather than by the cryptic file names found in simple directory listings or FTP-style hierarchical UNIX directories. PDF Launcher allows the user to immediately view the files chosen from the Title and Author listings.

# Aerial

Ambia offers an add-on document navigation tool for Adobe Acrobat called Aerial, which provides many of the common functions of paper documents, such as index searches, page number browsing, and a couple of handy tools for page marking and printing. Aerial reproduces many of the functions that users are accustomed to doing manually on paper documents.



Aerial enhances user cruising through PDF documents by bringing paper document techniques such as page number browsing and paper clip notations to PDF documents.

## Page Number Navigation

Many everyday documents have page-numbering systems that don't match Acrobat's sequential order. This feature restores that ancient Gutenberg facility to Acrobat documents. With its ability to recognize most page-numbering schemes, Aerial effectively upgrades your Go To Page button.

In any project that deals with traditional complex page- and section-numbering schemes, this feature is extremely helpful. Because of the complex nature of the information presented in the source documents, many complex page-numbering schemes are encountered. It is the norm rather than the departure for these documents to include title pages, prefaces, chapter numbering, appendices and so on.

Aerial can recognize page numbers in a variety of formats, including Roman numerals and double folio styles such as 5-15 and so on. In this way the reader can use the Go To Document Page tool to specify an original page number.

Of course, similar functionality could be added on at the time of creation of the digital document through the use of bookmarks or links. However, this tool brings this functionality to virtually all Acrobat documents with page numbers.

## Search Document Indices

The most common value-added navigation or finder/helper in a document, not counting the table of contents, is an index. All but the simplest books or manuals contain an index to guide the user to the information in the document.

An index in a digital document can often be more productive than a full text search. When a reader is looking for specific information, an index allows direct access to the place or places where the information occurs. A text search may produce too many hits, potentially causing the user to review irrelevant pages on the path to the correct page. In addition, an index can help a user find the proper spelling or usage of a term as it appears in a particular document.

The Aerial Index Manager is a Soundex listing, like the Windows Help dialog. As the user types in the letters of a word, the system automatically scrolls to that region of the index. The user can then click on a choice to visit that page.

The index feature can be used to create more refined text searches. By displaying occurrences of terms as they appear in a particular document, an advanced text search user can dispense with Wildcards, Thesauri and Word Stemming to avoid extraneous hits. As documents get larger, ever more precise text searching is highly desirable.



**tip**

**This is the Acrobat Index we are searching here, which is not the index that may appear in a source document. Typically, an index that appears in a book has been created by the author or the editor. That means that an additional level of information has been added in a conventional index.**

By choosing particular terms to include in an index, an author or editor has ranked the important information in the document and excluded the least-valuable terms and words. For this reason, it is often a very effective value to add hyperlinks from the index to the references of a document. Readers will appreciate this improved accessibility. Most readers probably expect it because they intuitively expect digital documents to be better than paper documents in every way.

**tip**

**The conventional index in a book offers an excellent built-in finding aid that should be exploited whenever possible. In new or editable digital documents, applications such as PageMaker and FrameMaker can directly convert such finding aids to bookmarks in PDF output. Alternatively, the author can link the terms in the index to their appearance in the text, for instant access in both directions via hypertext links.**

**Even in documents that are scanned and recognized via Capture, it is simple to use Acrobat Exchange to add hypertext links and bookmarks to table of contents, index, glossary, footnotes, references to graphics, and all other traditional elements.**

**With Aerial, the user can view the original document and then go directly to page numbers indicated by the index.**

## Tag Frequently Referenced Pages

Occasionally we have to admit that paper documents do offer some conveniences not always available in digital documents. When we put a physical book down, we can fold a corner to mark our place. The next time we pick up the book, we can flip directly to our marked spot or spots. Digital documents often aren't as easy to mark.

The Aerial Quick Tags tool provides this feature. Very similar to Netscape Bookmarks, the tags are stored outside the Acrobat documents themselves. Just as Netscape allows the user to click to visit favorite sites on the Web, the Quick Tags allow the user to immediately jump to a particular page.

In many projects, important documents maintain their relevance over a period of time. Previously published information must often be compared to newly arrived documents. Analysts and reviewers find this Quick Tag feature to be an almost absolute necessity. Without such features, a growing digital library can become unwieldy, and research and retrieval can be slowed. Once again, this feature restores a functionality familiar from books on shelves.

## Clip Pages For Quick Reference

Perhaps nothing is more annoying to the user of a digital document than losing his place. It's a common occurrence, however, during the process of tracking down a question, to find answers in separate, unrelated places. Following the train of thought often causes a user to lose track of earlier locations.

In a paper document, the user can stick a finger in related pages or put a pencil in the book. The Aerial Paper Clip tool gives the user this handy ability with digital documents. The user can quickly click back through marked pages for ease of reference.

Users and analysts of large collections of data often have the need to flip between charts and explanations on widely separated pages, and the Paper Clip tool makes this process simple and quick. This feature is essential for every task from a two-part brownie and icing recipe to a brain surgery text. And as usual, users seem to intuitively demand this instant access to information that is expected of digital documents.

## Print Selected Page Portions

Digital documents include all of the special pages of other complex documents. From the simplest oversized spreadsheet to the most complex map or schematic, oversized documents are common and must be accommodated in digital documents.

When it comes to viewing the document on-screen, the Zoom and Pan tools can be very useful and efficient. However, printing oversized documents on common-size printers can result in documents that need to be read with a magnifying glass.

To accommodate the user, who happens to be a human who has eyes that can't zoom in to pixel view of a piece of paper, it is important to be able to print documents in a

useful size. The Aerial Print View tool allows users to print parts of pages or print pages in sections for full-size reconstruction.

Like many differences between paper and digital documents, you never know what you're missing until it's gone. Single-page digital documents fall short on even simple dual-page documents, where a headline or illustration might span the left and right pages in an open book. The ability to print documents at their original dimensions once again rises to meet the functionality of old-fashioned paper.

## Sys-Print

Sys-Print software, published by Sys-Print, Inc., automatically converts high-volume print streams to fully bookmarked PDF documents. This product is aimed at current high-volume printing applications and is therefore designed to accept mainframe output and traditional print languages.

For example, in one demonstration, a 213-MB input file was processed on a 486/66 PC in less than three and a half hours. The resulting PDF file was 61 MB, including bookmarks and reports.

As Jim Ware, VP of Sales & Marketing and co-founder, explains it, "We all come from a publishing background, and we are aiming for completed automated database publishing." The Sys-Print product offers very sophisticated processing of the input streams to produce customized PDF output. We transform SysOut data directly into a bookmarked PDF.

For example, in a payroll application, multiple levels of bookmarks can be generated based on the structure of the mainframe data. Top-level bookmarks might represent regions, second-level bookmarks might represent departments, and third-level bookmarks might represent Social Security numbers.

Compared to traditional greenbar and laser print reports, the user has an infinitely better tool for quickly accessing data.

And for archived data on tape, Sys-Print offers a Data Mining tool because this data can be output as intelligently organized PDF collections, indexed by Catalog for text searching, perhaps printed to CD for distribution, and made available for access in an electronic document system via an Intranet.

## Emerge: Lessons Learned in an Acrobat Capture Service Bureau

### case study

As manager of the EmERGE Acrobat Capture Service Bureau, Glenn Gernert has faced and conquered most of the challenges that any potential user may encounter. In this conversation Glenn graciously shares a wealth of practical knowledge:

"One of the first considerations in any job is whether to use PDF Normal or Image + Text format, and when there are any legal ramifications we always recommend the latter. In terms of cost, Image + Text is about one-quarter the cost of full clean-up in PDF Normal. Of course, I+T files are 80-100K per page, while PDF Normal files are about 10K or less per page. So of course, anything going on the Web, we strongly recommend PDF Normal.

"We use a guillotine for de-binding, and we never scan more than 100 pages in a batch. We normally scan at 300 dpi, but below 8 pt. text we consider scanning at 400 dpi. Our process starts with document preparation, which includes debinding and fixing any mutilated pages. We count out 100 page batches and assign a name to the document during scanning, and we all check the process for 100 page batches. We use four or five character names, and use

and underscore to separate the sequential numbers. If we have grayscale images to scan, we scan the entire document in binary and then insert the grayscale images by replacing pages in the directory, and then Capture processes the entire document.

"To prepare for clean-up, we stitch the batches back together in Exchange. We feel it's very important to have a single ACD file for editing, rather than try to assemble the pieces at the end of the process.

"Ambia's Compose is worth its weight in gold! On clean documents, Compose makes it very simple to add Bookmarks with the Bookmark by Example feature. Another dynamite feature is the Page Linker for an Index or Table of Contents. The Info-fill feature is great for automatically entering all the Document Info fields by merging the files with an ASCII delimited file of information.

"We typically output on CD-ROM, because it is cheap and easy. It's a good idea to create a hybrid CD that can be read on Windows, Mac and UNIX rather than the standard ISO 9660 format. Now that Acrobat 3 includes the Reader with full Search capability, it doesn't make sense to create a CD without using Catalog to create an Index."



## Glenn Gernert's Basic Clean-Up Procedures Overview

Each page is comprised of two basic elements: Text and Graphics. Establishing a base for each is the purpose of this help sheet. The following are methods to increase efficiency and accuracy.

### Step By Step

#### I. Establish your base font.

Before beginning a new document, browse through it and take note of formatting.

Note all column headings, their font/size and bold/italic state.

Note switch of font sizes and styles that are consistent.

Once base font is established, be consistent in its use throughout the document.

#### II. Clean-up of Text

Select all text with your short cut key. (CTRL + A)

Set base font for page. (Change ALL text on page to base font)

Set base font size for page. (Change ALL text on page to base font size)

Bold on/off. (Change ALL text on page to bold, then non-bold)

Italics on/off. (Change ALL text on page to italic, then non-italic)

Re-establish any special font size changes and characteristics to the page.

Page numbers and Table/Figure headings should always be text.

CTRL-TAB through highlighted words as final task. (No highlighted words should remain on document, unless specifically placed there.)

#### III. Dealing with Graphics

Take note of the graphic elements on the page.

Replace Bullets, Squares, Dashes, and other organizational headings with graphic boxes. (But not numbered paragraphs.)

If spending more than a minute or two deciding how to handle a specific page/graphic, ask for help immediately.

**NOTE:** Keep in mind that this is a basic outline of how each page could be handled. There are documents you will not be able to use these methods on but overall, this will decrease time spent on proofing and checking each page.

Thanks to Glenn Gernert, Acrobat Capture Service Bureau Manager,  
EMERGE - <http://www.emrg.com>

# RAELS

The Rapid Access Electronic Library System from Loral Space and Range Systems offers an Acrobat Exchange plug-in called the RAELS PDF Links. This product was developed specifically to save time in creating hyperlinks within very large documents.

Loral estimates that on one task of hyperlinking 250,000 hotspots in a series of related repair manuals that totaled 6,500 pages, almost an entire person-year was saved through the use of PDF Links.

The approach of PDF Links is different from a “link by example” system because it is based on rules rather than examples. The user creates rules in a script that specifies how information is to be connected. For example, complex documents might have several rules to define relations between TOC, illustrations, appendices, footnotes, text and other documents.

The software ships with a Philosophy of Rules to help the user understand how to write simple and effective rules. There is a rule checker that verifies the integrity of user-created scripts. These rules are reusable and can be used on any number of new or updated documents.

Once the rules are created, the PDF Links software automatically applies them and creates links throughout the document and recompiles it to PDF format. PDF Links is also capable of generating automated bookmarks for top-level navigation.

The user manual is intended to educate the user in rule writing and includes plenty of samples to start with.

Complete  
information is  
available from  
Loral at:

<http://www.loral.com>

# PDF Reference Manual: 2nd Edition

Insert link to this epochal document, which is so cool it hardly needs to be updated, just like the pdfmark Reference Manual. Standards don't wiggle like Jell-O, unlike other technology fads.

<http://www.adobe.com/supportservice/devrelations/PDFS/TN/PDFSPEC.TXT>

Many of these products are available on the Web for free evaluation. The best places to start looking for commercial Acrobat plug-ins is:

<http://www.emrg.com>

## Summary

Every potential new user of the new media comes to the keyboard expecting better access to information. Beyond all the economic and process efficiencies of the new media, every individual user judges it in terms of, "What have you done for me lately?"

A world full of creative page and document designers is nimbly evolving into digital document designers. We can expect to see brilliant innovation in the near and ongoing future.

Business projects and procedures are being automated in this new media, and more efficient information-delivery mechanisms are already working and doing a better job than earlier ways.

Automation of digital document-building techniques and procedures can enable rapid site development. These tools allow automatic navigation building and free the Webmaster from tedious, repetitive tasks of meticulously linking interesting info. These tools can build table of contents, index and other links automatically. This is truly enhanced and improved access to information.

part

4

using  
digital  
content

