Optimizing

Adobe® PDF Files for

Accessibility

Adobe

Adobe Portable Document Format (PDF) is an open de facto standard for electronic document distribution worldwide. Adobe PDF is a universal file format that preserves all the fonts, formatting, colors, and graphics of any source document, regardless of the application and platform used to create it. Adobe PDF files are compact and can be shared, viewed, navigated, and printed exactly as intended by anyone, anywhere with the free Adobe Acrobat® Reader software. You can convert any document to Adobe PDF, even scanned pages, using Adobe Acrobat 4.0 software.

With Microsoft® Windows® software, screen readers keep track of what is happening on the screen by building a screen model in memory and then letting the user navigate through the model. Unfortunately, the screen models typically can represent text only from simple documents that contain a single column of text that flows from left to right and from top to bottom on the page. The models cannot properly represent documents with more complex layouts, such as multicolumn newspaper layouts or documents with text that is not horizontal, such as tax forms that contain vertical labels.

When a screen reader encounters a complex layout in a document, it can make mistakes. With a multicolumn document, a screen reader would probably read the first line in the first column, followed by the first line in the second column, and run the columns together because they are on the same horizontal line. Because Adobe PDF supports rich layouts, screen readers may not reproduce complex documents correctly.

One advantage of Adobe Acrobat software for sighted individuals is that it renders documents with fonts that are either the same as or similar to the author's original fonts. To do this, earlier versions of Acrobat software would sometimes encode fonts in a way that is incompatible with screen readers. This problem has been fixed in Acrobat 4.0.

access.adobe.com

The first solution for accessibility is a free, Web-based service at *http://access.adobe.com*. Through *access.adobe.com*, any Adobe PDF document on the Internet can be converted to HTML or plain text. Because the service converts Adobe PDF files into formats that can be represented by ASCII text, *access.adobe.com* works best on documents in English. Documents in European languages, such as French or German, should convert fairly well except for symbols that are not represented by ASCII text, such as umlauts. Languages with double-byte characters, such as Japanese, Chinese, Arabic, Hebrew, and Russian, are not supported.

The *access.adobe.com* site approximates the logical reading order of the text in an Adobe PDF document and reformats it into a single column of text that can be read easily by screen readers. This solution provides three benefits:

- Visually disabled users can use familiar applications, screen readers, and Web browsers.

- PDF files can be accessible on any platform that supports Web browsers and screen readers.

- Some of the problems screen readers have with complex layouts are eliminated.

The *access.adobe.com* site provides two conversion options. The first is a Web-based form that can be used to convert PDF documents that are on the Internet. Users type the URL to an Adobe PDF document and click the Get This PDF Document as HTML button. The document is converted to HTML on-the-fly and is returned immediately to the Web browser. With the

second option, users can submit the URL of an Adobe PDF file via e-mail to convert it to HTML or ASCII text. Adobe PDF files on a local hard disk, local CD-ROM, or local area network (LAN) also can be converted by attaching the Adobe PDF file to an e-mail message. The converted results are returned in the body of a new mail message in a few minutes.

When *access.adobe.com* converts Adobe PDF to HTML, all existing hypertext links are converted to HTML links. This includes intradocument links as well as links to other documents on the Internet. HTML links are also created to enable easy navigation among pages. The first line of the document contains at least two of these special links:

- **Document Body** is a special link to the beginning of the document.

- **Page Navigation Panel** is a part of the document that contains links to each page of the document. For example, if the document has five pages, Page Navigation Panel contains links called 1, 2, 3, 4, and 5, each corresponding to its respective page.

- **Document Outline** contains HTML links that correspond to the Acrobat bookmarks found in the original document. Acrobat bookmarks are an electronic table-of-contents feature and are different from the bookmarks associated with the Web browser.

- At the end of each page of the document, links are inserted to make it easy to move to the next and previous pages.

Acrobat Access™ Plug-In

Adobe Acrobat Access 4.0 for Microsoft Windows software is an alternative solution for reading Adobe PDF files. This plug-in for Adobe Acrobat software provides accessibility to Windows screen-reading programs. The Acrobat Access plug-in converts Adobe PDF files on a user's local system into plain text, which then can be read by screen-reading programs.

As with *access.adobe.com*, the Acrobat Access plug-in extracts the text of a document in approximate reading order and in a single column, enabling screen readers to follow the flow of the document. Acrobat Access also can preserve navigational elements, such as hypertext links and bookmarks. Acrobat Access 4.0 is available now and can be downloaded for free from *access.adobe.com*.

Demonstration of Accessible PDF Forms

Adobe Acrobat 4.0 software lets users create, fill in, and submit electronic Adobe PDF forms. Adobe PDF forms can include fields with text boxes, buttons, checkboxes, combo boxes, list boxes, radio buttons, and signature fields. When all the proper software and hardware components are in place, Adobe PDF forms data can be submitted over the Web and collected in a database, in the same way that HTML forms are processed.

The *access.adobe.com* Web site and the Acrobat Access plug-in do not provide full accessibility to Adobe PDF forms. Adobe now has a technology demonstration of accessible Adobe PDF forms for Acrobat 4.0.5 software for Windows. The forms accessibility demonstration works with the free Acrobat Reader 4.0.5 for Windows as well as with the full Acrobat 4.0.5 for Windows product.

The forms accessibility demonstration is not intended to be a full solution, but it provides an illustration of future plans to enhance the accessibility of Adobe Acrobat software. Users must download and install additional components to see the Adobe PDF forms accessibility demonstration. These components can be downloaded from *http://access.adobe.com.*

Unlike *access.adobe.com* or the Acrobat Access plug-in, the Adobe PDF forms accessibility demonstration does not rely on screen readers. Instead, it uses the JavaScript TTS (text-to-speech) object, which renders machine text as digitally synthesized speech. To use the TTS object, users must have a TTS engine installed on their computers.

You can download the Microsoft Speech Application Programming Interface (MSAPI) 4.0 Software Development Kit (SDK) from the Microsoft Web site, which provides a TTS engine. Any speech engine that is MSAPI 4.0-compatible should also work.

With the Adobe PDF forms accessibility demonstration, users can create Adobe PDF forms that offer a rich set of accessibility options. These include:

- **Document Title.** When an Adobe PDF document is opened, saved, printed, or closed, the document title is spoken to the user. If the title has not been specified in the File > Document Info > General dialog box, then the filename is used. Because filenames are usually abbreviated or changed frequently, specifying a title usually results in the document being more accessible. For example, if a document's filename is *IRS1040.pdf*, a good document title would be *Form 1040: U.S. Individual Income Tax Return for 1998.*

- **Short Description of a Form Field.** Authors creating Adobe PDF forms can specify a field name and optionally provide a short description of the field. When a user interacts with the field, the name is spoken. If a description has also been provided, this is spoken instead. Descriptions usually provide a more accessible experience for disabled users because field names tend to be short or cryptic. For example, if a field is labeled *name.1*, a good short description would be *First Name.*

- **Sound and Speech Cues.** Authors creating Adobe PDF forms can provide sound and speech cues to indicate context. A sound cue is a *.wav* file that plays when certain events occur. A speech cue is a text string that is spoken by the TTS engine when certain events occur. Some

events can have both sound cues and speech cues attached to them. For example, when a user goes to the second page of a three-page document, a sound cue indicating that a page has been turned can be played, followed by a speech cue that says, "Page two of three." The Acrobat forms accessibility demonstration has a default set of sound and speech cues that can be modified by users. Events that can have sound cues or speech cues attached to them include:

- Application initialization

- Document open, close, activate, save, and print

- Page turn

- Keystroke handling when interacting with a field

- Alert dialog boxes

Optimizing Adobe PDF Files for Accessibility

Adobe Acrobat 4.0 software lets users create Adobe PDF files that contain logical document structure ("logical structure" refers to the organization of the document, such as title page, chapters, sections, and subsections). When Adobe PDF files contain logical structure, users can easily create bookmarks, print specific pages in a document, or extract information.

In the near term, Adobe PDF files that contain logical structure are not more accessible than files without logical structure. Adobe intends to develop versions of Acrobat software that will use logical structure to increase the accessibility of Adobe PDF. Logical structure can indicate precise reading order and improve navigation, particularly for longer, complex documents. For example, an Adobe PDF document may be a newspaper, where the first page contains the beginning of three articles, each of which is continued on a separate page. With a knowledge of the logical structure, Acrobat software's accessibility solutions will be able to follow a single article from beginning to end, much as a sighted person would flip through the newspaper pages to continue reading an article.

Adobe Acrobat 4.0 software enables users to create Adobe PDF files with logical structure

Users can set Web Capture preferences to "Add PDF Structure." This setting stores logical structure in the Adobe PDF file that corresponds to the HTML structure of the original Web pages. If this option is selected, users can create structured bookmarks for paragraphs, list elements, table cells, and other items that use HTML tags.

PDFMaker for Microsoft Word for Windows

Acrobat PDFMaker 4.0 allows users to create Adobe PDF files directly from Microsoft Word for Windows 97 or 2000. PDFMaker installs automatically when Acrobat 4.0 software is installed and is available as an icon from the Microsoft Word toolbar.

PDFMaker allows users to specify the Word headings and paragraph styles for logical structure. Once the structure is specified in the Adobe PDF file, users can automatically create bookmarks and easily reuse information.

Optimization Guidelines

The following guidelines are designed to help optimize Adobe PDF documents to be as accessible as possible with current Adobe solutions and with future versions of Adobe Acrobat software:

- Use Acrobat 4.0. If users have legacy Adobe PDF files that include TrueType fonts and were created with earlier versions of Acrobat for Windows, they can convert the original source files to Adobe PDF using Acrobat Distiller® 4.0 and the Adobe PostScript® drivers AdobePS™ 4.3.1 for Windows 95 and Windows 98 or AdobePS 5.1.1 for Windows NT® 4.0. Converting documents to Adobe PDF using Acrobat 4.0 and the AdobePS drivers solves the problem that screen readers had in the past of recognizing certain kinds of text represented by Windows TrueType fonts. Acrobat 4.0 includes features such as the ability to include logical structure in an Adobe PDF file to optimize a document for future accessibility. Legacy PDF documents that contain only Type 1 fonts, or that contain TrueType fonts and were created with Acrobat for Macintosh, should be fine.

- Use Acrobat bookmarks. For visually disabled users, an electronic table of contents or set of links can provide an overview of a document that is easy to navigate. Bookmarks can be automatically generated when creating Adobe PDF documents from popular desktop-publishing programs, such as Adobe PageMaker® and Adobe FrameMaker® software, or from Microsoft Word. The service at *access.adobe.com* and the Access plug-in convert Acrobat bookmarks to hypertext links that can be accessed by visually disabled users.

- Use PDFMaker 4.0 with Microsoft Word. PDFMaker lets Word users create Adobe PDF documents with logical structure. Logical structure can be used to automatically create bookmarks and enhance accessibility with future versions of Acrobat software.

- Use "Image with Text" or "Normal" settings when converting scanned paper documents to Adobe PDF files. Acrobat 4.0 and Adobe Acrobat Capture® 2.01 software let users convert scanned paper documents to Adobe PDF. Sometimes scanned documents are converted as image-only Adobe PDF files, which are bitmapped images contained in an Adobe PDF file. These images are inaccessible to visually disabled users.

- Test your Adobe PDF files for accessibility. Use *access.adobe.com* or the Acrobat Access plug-in to check how accurately your Adobe PDF file converts to text. Pay particular attention to documents that contain graphics or images, multiple columns of text, or text on a path.

- Provide links to *access.adobe.com*. This Web site is a resource to any potential visually disabled users visiting your Web site. Details about linking to *access.adobe.com* can be found at *http://access.adobe.com/access_faq.html#09.*

- Check Adobe PDF documents for the following types of content, which may not be fully accessible through *access.adobe.com* or through the Acrobat Access plug-in. Future versions of Acrobat software will improve accessibility to the following types of content:

  - **Text on a path.** Adobe PDF can represent graphically rich documents with complex layouts, such as text that follows the edge of a circle. Current Adobe solutions can usually extract this text, but the extracted text might not be organized in the proper reading order.

  - **Bitmapped images.** Bitmapped images are not accessible through *access.adobe.com* or the Access plug-in. Usually a descriptive caption that accompanies an image can be accessed through Adobe's current solutions and can provide sufficient context for visually disabled users.

  - **Certain types of annotations.** Adobe Acrobat 4.0 software lets users annotate documents with electronic sticky notes, highlighting, and other forms of mark-up. Some of these annotation types are not currently accessible. Acrobat annotations are not usually required when publishing documents for mass distribution.

  - **Information in tables.** Adobe's current accessibility solutions can extract text from tables, but the extracted text might not always be in the logical reading order.

  - **Image-only PDF files.** Adobe Acrobat 4.0 and Acrobat Capture 2.01 software let users create Adobe PDF files from scanned paper documents. One option lets users convert scanned documents to Adobe PDF files of bitmapped images without any text. These image-only Adobe PDF files are inaccessible to visually impaired users.

  - **Forms.** Adobe Acrobat 4.0 software lets users create, fill in, and submit electronic Adobe PDF forms. The *access.adobe.com* Web site and the Access plug-in do not provide full accessibility to Adobe PDF forms. Adobe intends to enhance accessibility to forms in future releases of Acrobat software.

Future Enhancements to Adobe Acrobat Software

Adobe Systems intends to enhance accessibility features in the next versions of Acrobat software. These enhancements will be designed to achieve the following objectives:

- Enable organizations to publish Adobe PDF documents that are accessible through the free Adobe Acrobat Reader software for users with low visual acuity or blindness, motion impairment, and other disabilities.

- Provide tools for organizations publishing Adobe PDF documents to review and optimize those documents for accessibility.

In addition, Adobe will provide guidelines for creating accessible Adobe PDF documents with future Acrobat software releases. Adobe intends to achieve these goals through several methods. One method is to enhance the Acrobat software's ability to exploit logical document structure. Logical structure lets Acrobat software automatically create accessible navigation, such as bookmarks, and provides a clear indication of logical reading order. Enhancing the accessibility of Adobe PDF forms will also be a primary focus for future versions. By achieving these goals, Adobe believes it can simplify the process of creating Adobe PDF documents that will be accessible to a broad range of people with disabilities.