

Adobe PDF 101

Quick overview of PDF file format

TABLE OF CONTENTS

- 1 Description
- 1 PDF versions
- 1 Working with Adobe Intelligent Document Platform solutions
- 1 Under the hood
- 2 Save and Save As


Description

The Adobe® Portable Document Format (PDF) is a formatting language, first conceived by John Warnock, one of the founders of Adobe Systems. The language is large and complex, but here is a quick overview of the key elements for use with eForms.

PDF versions

The Adobe PDF format is 12 years old. The first version, 1.0, was introduced in 1993. Subsequent releases have added new functionality to the specification, and Adobe’s flagship products — Adobe Acrobat® and Adobe Reader® software — have progressed accordingly. The following chart shows corresponding versions of Acrobat and the PDF specification. The *PDF specification* is freely available. *Acrobat* and *Reader* are applications based on the PDF specification.

| PDF VERSION | ACROBAT VERSION |
|-------------|-----------------|
| 1.3 | 4.x |
| 1.4 | 5.x |
| 1.5 | 6.x |
| 1.6 | 7.x |

 Adding the two numbers in the PDF version column equals the version number of Acrobat where the PDF specification was introduced.

Acrobat and Reader 7.0 can open any previous versions all the way back to PDF 1.0. However, sometimes there are problems opening a newer PDF version (1.6) with an older client (for example, Reader 4.0). Newer functionality or enhanced security settings may have been used in the PDF file such that the older client is unable to interpret the information correctly. The good news is that Adobe Reader is distributed freely by Adobe so that users can upgrade to view newer files.

Working with Adobe Intelligent Document Platform solutions

Adobe LiveCycle™ was introduced in 2004. This release includes tools such as Adobe LiveCycle Designer and Adobe LiveCycle Forms software, which generate PDF 1.5 and PDF 1.6 compliant documents. The ability to embed arbitrary XML in a PDF file was introduced in the PDF 1.5 specification and will work only with PDF 1.5 or higher files.

Under the hood

The general structure of a PDF file is composed of the following code components: header, body, cross-reference (xref) table, and trailer, as shown in figure 1.

The header contains just one line that identifies the version of PDF.

Example: **%PDF-1.6**

The trailer contains pointers to the xref table and to key objects contained in the trailer dictionary. It ends with **%%EOF** to identify end of file.

The xref table contains pointers to all the objects included in the PDF file. It identifies how many objects are in the table, where the object begins (the offset), and its length in bytes.

The body contains all the object information — fonts, images, words, bookmarks, form fields, and so on.

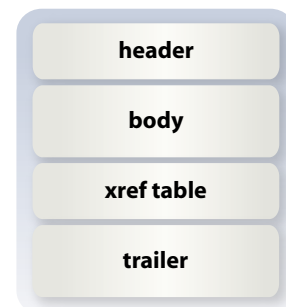
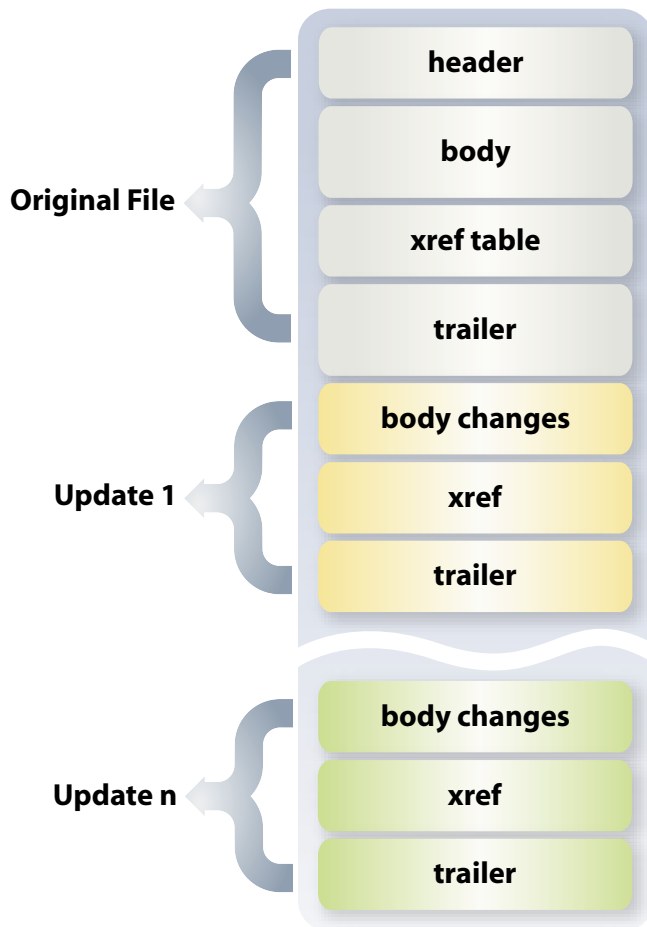


Figure 1. Basic structure of a PDF file

Save and Save As

When you perform a Save operation on a PDF file, the new, incremental information is appended to the original structure (see figure 2); that is, a new body, xref table, and trailer are added to the original PDF file.

Figure 2. PDF structure after Save



If you have applied digital signatures to a PDF file, use the Save function. This preserves the “tracking” of the changes made between original and signed versions. If Save As... is applied to a PDF file with digital signatures, the signatures are invalidated.

You’ll notice that after ten Save operations, you are prompted to Save As... to reduce file size. When you perform a Save As..., Acrobat merges the updated information into the original, reverting to the original structure of one body, one xref table, and one trailer (see figure 1).

For a full description of the PDF file format, visit the Specifications page on Adobe’s partner site.

FOR MORE INFORMATION

If you’d like to provide feedback on this tip or if you have questions, send e-mail to [Lori DeFurio](mailto:Lori.DeFurio@adobe.com), Developer Evangelist.

Adobe Systems Incorporated • 345 Park Avenue, San Jose, CA 95110-2704 USA • www.adobe.com

Adobe, the Adobe logo, Acrobat, Adobe Portable Document Format, Reader, LiveCycle, Adobe Intelligent Document Platform 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.

© 2005 Adobe Systems Incorporated. All rights reserved.5/05