


Portable Document Format

Материал из Википедии — свободной энциклопедии

Portable Document Format (PDF) — межплатформенный открытый формат электронных документов, изначально разработанный фирмой Adobe Systems с использованием ряда возможностей языка PostScript. В первую очередь предназначен для представления полиграфической продукции в электронном виде. Для просмотра существует множество программ, а также официальная бесплатная программа Adobe Reader. Значительное количество современного профессионального печатного оборудования имеет аппаратную поддержку формата PDF, что позволяет производить печать документов в данном формате без использования какого-либо программного обеспечения. Традиционным способом создания PDF-документов является виртуальный принтер, то есть документ как таковой готовится в своей специализированной программе — графической программе или текстовом редакторе, САПР и т. д., а затем экспортируется в формат *PDF* для распространения в электронном виде, передачи в типографию и т. п.

PDF с 1 июля 2008 года является открытым стандартом ISO 32000^{[1][2]}.

Формат PDF позволяет внедрять необходимые шрифты (построчный текст), векторные и растровые изображения, формы и мультимедиа-вставки. Поддерживает RGB, СМЯК, Grayscale, Lab, Duotone, Bitmap, несколько типов сжатия растровой информации. Имеет собственные технические форматы для полиграфии: PDF/X-1a, PDF/X-3. Включает механизм электронных подписей для защиты и проверки подлинности документов. В этом формате распространяется большое количество сопутствующей документации.

Portable Document Format	
Расширение	.pdf
MIME-тип	application/pdf application/x-pdf application/x-bzpdf application/x-gzpdf
Сигнатура	%PDF
Разработчик	<u>Adobe Systems</u> , <u>ISO</u>
Опубликован	15 июня 1993
Последний выпуск	2.0
Тип формата	<u>Электронный документ</u>
Расширен из	<u>PostScript</u>
Развит в	<u>PDF/A</u> , <u>PDF/X</u> , <u>PDF/E</u> , <u>PDF/UA</u> , <u>PDF/VT</u>
Стандарт(ы)	<u>ISO 32000-2</u>
Открытый формат?	Да
Сайт	<u>iso.org/standard/63534.h...</u> (<u>https://www.iso.org/standard/63534.html</u>)
 Медиафайлы на Викискладе	

Содержание

Общие сведения

Редактирование PDF

История

Версии Adobe PDF

Программы для работы с PDF

Чтение

Редактирование

Библиотеки для работы с PDF

C

C++

Java

PHP

Примечания

Ссылки

Общие сведения

Чаще всего PDF-файл является комбинацией текста с растровой и векторной графикой, реже — текста с формами, сценариями на языке JavaScript, 3D-графикой и другими типами элементов.

Информационные объёмы двух одинаково выглядящих на экране PDF-документов могут значительно отличаться в зависимости от:

- внедрения или связывания шрифтов и мультимедиа;
- разрешения растровых изображений;
- использования встроенного механизма сжатия всего документа;
- используемых алгоритмов сжатия растровых изображений.

Для создания документа минимального объёма необходимо использовать векторную графику и «безопасные» шрифты. Всего имеется 14 таких шрифтов:

- Times (v3) (обычный, курсив, полужирный и полужирный курсив)
- Courier (обычный, наклонный, полужирный и полужирный наклонный)
- Helvetica (v3) (обычный, наклонный, полужирный и полужирный наклонный)
- Symbol
- Zapf Dingbats

Эти шрифты можно использовать без внедрения в документ, так как их должны правильно отображать все программы. Любые другие шрифты, которые не были внедрены в документ и отсутствуют в системе, будут заменены одним из имеющихся, что может стать причиной неправильного отображения страниц, количества символов в строке и других ошибок отображения, связанных с метрикой шрифта.

Существует технология MRC (Mixed Raster Content), которая приближает возможности PDF к возможностям формата DjVu по хранению сканированных изображений с текстом.^[3]

Редактирование PDF

Существуют специальные программы, позволяющие редактировать PDF-файлы, однако их выбор значительно меньше, чем программ для создания или просмотра PDF документов. [Inkscape](#), начиная с версии 0.46, позволяет редактировать PDF^[4] благодаря использованию на промежуточном этапе программы [Poppler](#). В 2017 году функциональность по редактированию PDF была поддержана в российской программе [ABBYY FineReader](#), ранее специализировавшейся только на распознавании и конвертировании PDF в другие форматы. ^[5]

История

В первое время существования данный формат был крайне непопулярен:

- программное обеспечение компании Adobe для чтения и создания PDF было платным;
- в PDF отсутствовала поддержка внешних ссылок, что делало его практически бесполезным во всемирной паутине;
- PDF-документы были большего размера по сравнению с обычным текстом, что означало более длительную загрузку на медленных модемах, широко использовавшихся в те времена;
- на слабых машинах отображение PDF-документов осуществлялось с заметными задержками;
- существовало несколько конкурирующих форматов, таких, как [Envoy](#), [Common Ground Digital Paper](#), [Farallon Replica](#); даже собственный [PostScript](#) являлся конкурентом.

После того как Adobe выпустила бесплатную версию [Acrobat Reader](#) (позднее переименованную в [Adobe Reader](#)) для чтения PDF-документов, популярность этого формата стала возрастать. Формат PDF-файлов несколько раз изменялся и продолжает эволюционировать. Существует несколько спецификаций формата, последовательно расширяющих друг друга. Для каждой новой спецификации создаются новые версии программного обеспечения из пакета [Adobe Acrobat](#). Ниже показана таблица соответствий версий документов и версий программ, в которых впервые была введена поддержка этих документов. Версию любого PDF-документа можно узнать по первым восьми байтам, открыв этот документ в текстовом режиме, например в блокноте.

Версии Adobe PDF

год	версия документа	новые возможности	версия ПО
1993	PDF 1.0		Acrobat 1.0
1994	PDF 1.1	<u>пароли</u> , <u>ссылки</u> , <u>потоки</u> , независимая от устройства цветопередача	Acrobat 2.0
1996	PDF 1.2	интерактивные элементы, обработка событий мыши, мультимедийные типы, <u>Юникод</u> , улучшенное представление цвета и графики	Acrobat 3.0
1999	PDF 1.3	<u>цифровые подписи</u> , <u>цветовые пространства ICC</u> и <u>DeviceN</u> , <u>JavaScript</u>	Acrobat 4.0
2001	PDF 1.4	<u>JBIG2</u> , прозрачность, текстовый слой <u>OCR</u>	Acrobat 5.0
2003	PDF 1.5	<u>JPEG 2000</u> , связанное мультимедиа, объектные потоки, перекрестные потоки, слои	Acrobat 6.0
2005	PDF 1.6	внедренное мультимедиа, 3D, <u>XML-формы</u> , <u>AES-шифрование</u>	Acrobat 7.0
2006	PDF 1.7		Acrobat 8.0
2008	PDF 1.7, <u>AEL3</u>	<u>AES-шифрование</u> 256-битным ключом	Acrobat 9.0
2009	PDF 1.7, <u>AEL5</u>	XFA 3.0	Acrobat 9.1
2011	PDF 1.7, <u>AEL8</u>		Acrobat X (10)

Программы для работы с PDF

Чтение

Свободные:

- MuPDF
- Evince
- Okular
- KPDF (разработка прекращена)
- SumatraPDF
- Mozilla Firefox
- Atril Document Viewer
- Xpdf
- Ghostview из пакета Ghostscript

Проприетарные:

- Adobe Reader
- ABBYY FineReader
- Foxit Reader

- [Google Chrome](#)
- [PDF-XChange Viewer](#)
- PDFMaster
- [STDU Viewer](#)
- Nitro Reader

Редактирование

Свободные:

- PDFmod (только Linux системы)
- PDFsam
- [OpenOffice Writer](#)
- [LibreOffice Draw](#)
- [Pdftk](#)

Проприетарные:

- [Adobe Acrobat](#)
- [ABBYY FineReader](#)
- Master PDF Editor
- Microsoft Word 2013 и выше
- FlexiPDF
- [PDF-XChange Viewer](#) (частично)
- Foxit PhantomPDF
- Nitro Pro

Библиотеки для работы с PDF

C

Свободные:

- Haru

Проприетарные:

- Adobe PDF Library SDK

C++

Свободные:

- PoDoFo
- Poppler
- QPDF

Java

Свободные:

- Apache PDFBox — создание документов, слияние или разделение документа(ов), извлечения текста или другого контента из документа, печать, электронная подпись.
- ICEpdf — просмотр и печать документов, преобразование в изображение, извлечения контента, поиск в документе.

Проприетарные:

- jPedal — просмотр и обработка документов, извлечения контента.

PHP

Свободные:

- pdfparser (<http://pdfparser.org/>) — библиотека PHP для синтаксического анализа PDF файлов и извлечения таких элементов, как текст.

Примечания

1. Формат PDF стал международным стандартом (<http://www.lenta.ru/news/2007/12/05/pdf/>) — [lenta.ru](http://www.lenta.ru)
2. ISO Ballot for PDF 1.7 Passed! — [blogs.adobe.com](http://blogs.adobe.com/insidepdf/2007/12/iso_ballot_for_pdf_17_passed.html) (http://blogs.adobe.com/insidepdf/2007/12/iso_ballot_for_pdf_17_passed.html) (англ.)
3. Как технология MRC уменьшает размер PDF-документов (<http://habrahabr.ru/company/abbyy/blog/119790/>)
4. *Vah, T.* Inkscape: Guide to a Vector Drawing Program. — Prentice Hall, 2011. — ISBN 9780132764148.
5. Четыре в одном: обзор нового ABBYY FineReader 14 (<https://3dnews.ru/947258>) (рус.), *3DNews - Daily Digital Digest*. Дата обращения 24 октября 2018.

Ссылки

- PDF Specifications (http://www.adobe.com/devnet/pdf/pdf_reference.html), including the PDF Reference for PDF 1.7, PDF 1.6 (ISBN 0-321-30474-8), PDF 1.5, PDF 1.4 (ISBN 0-201-75839-3), PDF 1.3 (ISBN 0-201-61588-6)
- Portable Document Format (http://curlie.org/Computers/Data_Formats/Document/Publishing/PDF/) в каталоге ссылок Open Directory Project (dmoz)

Источник — https://ru.wikipedia.org/w/index.php?title=Portable_Document_Format&oldid=107453890

Эта страница в последний раз была отредактирована 3 июня 2020 в 07:26.

Текст доступен по лицензии Creative Commons Attribution-ShareAlike; в отдельных случаях могут действовать дополнительные условия.

Wikipedia® — зарегистрированный товарный знак некоммерческой организации Wikimedia Foundation, Inc.