

Document file format

A **document file format** is a text or binary file format for storing documents on a storage media, especially for use by computers. There currently exist a multitude of incompatible document file formats.

A rough consensus has been established that XML is to be the technical basis for future document file formats, although PDF is likely to remain the format of choice for fixed-layout documents. Examples of XML-based open standards are DocBook, XHTML, and, more recently, the ISO/IEC standards OpenDocument (ISO 26300:2006) and Office Open XML (ISO 29500:2008).

In 1993, the ITU-T tried to establish a standard for document file formats, known as the Open Document Architecture (ODA) which was supposed to replace all competing document file formats. It is described in ITU-T documents T.411 through T.421, which are equivalent to ISO 8613. It did not succeed.

Page description languages such as PostScript and PDF have become the *de facto* standard for documents that a typical user should only be able to create and read, not edit. In 2001, a series of ISO/IEC standards for PDF began to be published, including the specification for PDF itself, ISO-32000.

HTML is the most used and open international standard and it is also used as document file format. It has also become ISO/IEC standard (ISO 15445:2000).

The default binary file format used by Microsoft Word (.doc) has become widespread *de facto* standard for office documents, but it is a proprietary format and is not always fully supported by other word processors.

Contents

Common document file formats

See also

References

External links

Common document file formats

- ASCII, UTF-8 — plain text formats
- Amigaguide
- .doc for Microsoft Word — Structural binary format developed by Microsoft (specifications available since 2008 under the Open Specification Promise)^{[1][2]}
- DjVu — file format designed primarily to store scanned documents^[3]
- DocBook — an XML format for technical documentation
- HTML (.html, .htm), (open standard, ISO from 2000), in combination with possible image files referred to.
- FictionBook (.fb2) — open XML-based e-book format
- Office Open XML — .docx (XML-based standard for office documents)
- OpenDocument — .odt (XML-based standard for office documents)
- OpenOffice.org XML — .sxw (open, XML-based format for office documents)

- OXPS — Open XML Paper Specification (Windows 8.1 and above, older version is XPS used in Windows 7)
- PalmDoc — Common Handheld document format
- .pages for Pages
- PDF — Open standard for document exchange. ISO standards include PDF/X (eXchange), PDF/A (Archive), PDF/E (Engineering), ISO 32000 (PDF), PDF/UA (Accessibility) and PDF/VT (Variable data and transactional printing). PDF is readable on almost every platform with free or open source readers. Open source PDF creators are also available.
- PostScript - .ps
- Rich Text Format (RTF) — meta data format being developed by Microsoft since 1987 for Microsoft products and cross-platform document interchange^{[4][5][6][7][8][9]}
- SYmbolic LinK (SYLK)
- Scalable Vector Graphics (SVG) - Graphics format primarily for vector-based images.
- TeX — Popular open-source typesetting program and format. First successful mathematical notation language.
- TEI — XML format for digital publication
- Troff
- Uniform Office Format — Chinese standard
- WordPerfect (.wpd, .wp, .wp7, .doc) (Note: possible confusion with Word format extension)

See also

- List of document file formats
- List of document markup languages
- Comparison of document markup languages
- Open format

References

1. "Microsoft Office Binary (doc, xls, ppt) File Formats" (<https://web.archive.org/web/20090308012827/http://www.microsoft.com/interop/docs/OfficeBinaryFormats.mspx>). 2008-02-15. Archived from the original (<http://www.microsoft.com/interop/docs/OfficeBinaryFormats.mspx>) on 2009-03-08. Retrieved 2010-03-18.
2. Microsoft Corporation (2010-07-23). "MS-DOC - Word Binary File Format (.doc) Structure Specification" (<http://msdn.microsoft.com/en-us/library/cc313153.aspx>). Retrieved 2010-08-08.
3. "What is DjVu - DjVu.org" (<http://djvu.org/resources/whatisdjvu.php>). *DjVu.org*. Retrieved 2009-03-05.
4. Microsoft Corporation (May 1999). "Rich Text Format (RTF) Specification, version 1.6" ([http://msdn.microsoft.com/en-us/library/aa140280\(office.10\).aspx](http://msdn.microsoft.com/en-us/library/aa140280(office.10).aspx)). Retrieved 2010-03-13.
5. "4.3 Non-HTML file formats" (<https://web.archive.org/web/20100218033318/http://archive.cabinetoffice.gov.uk/e-government/resources/handbook/html/4-3.asp>). e-Government Unit. May 2002. Archived from the original (<http://archive.cabinetoffice.gov.uk/e-government/resources/handbook/html/4-3.asp>) on February 18, 2010. Retrieved 2010-03-13.
6. "RTF (.rtf)—Wolfram Language Documentation" (<http://reference.wolfram.com/mathematica/ref/format/RTF.html>). *reference.wolfram.com*.
7. "WD: Rich Text Format (RTF) Specification 1.7" (<http://support.microsoft.com/kb/86999>). *support.microsoft.com*.
8. "Techtree.com India > Technology News, Reviews of Mobile Phones, PC Hardware and Electronics, Free Downloads, Forums, Helpdesk, Ask Tech Tree" (<https://web.archive.org/web/>

[20131029225243/http://archive.techtree.com/techtree/jsp/article.jsp](http://archive.techtree.com/techtree/jsp/article.jsp)). 2013-10-29. Archived from the original on October 29, 2013.

9. Ranjan Parekh, Ranjan (2006). *Principles of Multimedia*. Tata McGraw-Hill. p. 87. ISBN 0-07-058833-3.

External links

- [Lost in Translation: Interoperability Issues for Open Standards - ODF and OOXML as Examples \(http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1201708\)](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1201708)
 - [Secure document storage \(https://web.archive.org/web/20100222102019/http://www.the-stockroom.co.uk/\)](https://web.archive.org/web/20100222102019/http://www.the-stockroom.co.uk/)
-

Retrieved from "https://en.wikipedia.org/w/index.php?title=Document_file_format&oldid=946180951"

This page was last edited on 18 March 2020, at 16:19 (UTC).

Text is available under the [Creative Commons Attribution-ShareAlike License](#); additional terms may apply. By using this site, you agree to the [Terms of Use](#) and [Privacy Policy](#). Wikipedia® is a registered trademark of the [Wikimedia Foundation, Inc.](#), a non-profit organization.