

Format Background Document: Microsoft Excel 10.0 (XLS)

Date: July 12, 2005

Author: Grace Carpenter

GDFR data

Canonical identifier info:gdfr/fred/f/excel

Description Microsoft Excel 10.0 (XLS)

Alias

Type MIME type

Value <application/vnd.ms-excel>

Alias

Type Informal identifier

Value BIFF8X (Binary File Format)

Alias

Type Informal identifier

Value Excel XP/2002

Version 10.0

Legal or recognized owner

Name Microsoft Corporation

Organization type Commercial (for-profit) entity

Web site <http://www.microsoft.com>

Relationship

Type of relation May be encapsulated by target

Target format identifier

Type GDFR format identifier

Value gdfr/fred/f/ole2cdf

Target registry identifier

Type GDFR registry identifier

Value Fred

Specification

Document title Microsoft Excel File Format

Document type Article

Publisher Name OpenOffice.org

Organization type Non-profit entity

Web site <http://sc.openoffice.org>

Note This is an unofficial, unauthorized spec published by the creators of Open Office software

Access regime Unrestricted access

Identifier <http://sc.openoffice.org/excelfileformat.pdf>

Signature

Signature obligation Mandatory under certain conditions (see notes)

External signature type File extension

Signature value .xls

Note indicates a worksheet

Signature

Signature obligation Mandatory under certain conditions (see notes)

External signature type File extension

Signature value .xlt

Note indicates a template

Signature

Signature obligation Mandatory under certain conditions (see notes)

External signature type File extension

Signature value .xla

Note indicates an add-in (compiled library of functions)

Signature

Signature obligation Mandatory under certain conditions (see notes)

External signature type File extension

Signature value .xlw

Note indicates a workspace

Signature

Signature obligation Mandatory

Signature position Fixed position (requires offset)

Byte offset 0

Signature value 0xD0 0xCF 0x11 0xE0 0xA1 0xB1 0x1A 0xE1

Note This signature identifies an OLE Compound Document, which may or may not contain Excel data

Application

Application name Microsoft Excel

Application version XP.2002

Application release date 2001

Vendor

Name Microsoft Corporation

Organization type Commercial (for-profit) entity

Application's function

Process type Creates objects in the format

Application's function

Process type Renders objects in the format

Software dependency

Application name Microsoft Windows

Application version 95/98/Me/NT/2000/XP

Application release date 1995

Vendor

Name Microsoft Corporation

Organization type Commercial (for-profit) entity

Non-GDFR

1.General

1.1 Description (long) Excel 10.0 or XP, also known as BIFF8X (BIFF stands for Binary Interchange File Format), is the native file format of Microsoft's Excel software. It is a binary file format that is always contained within a Compound Document File ([OpenOffice CompoundDoc])

1.2 Content type: spreadsheet/graphs

2. Category-specific (e.g., color depth, color space, progressive display, etc.)

3.General technical

3.1 Encoding Unicode

3.2 Byte order Little Endian

3.3 Encryption A spreadsheet is encrypted with RC4 encryption if a user chooses to use password protection.

3.4 Human readable no

4. Sustainability

4.1. Proprietary Yes.

4.2 Owner documentation Microsoft does not publish a spec for recent versions of Excel. Although Microsoft publishes very little information about the file format itself, it provides a great deal of information about using Excel on its website (<http://office.microsoft.com/en-us/FX010858001033.aspx>)

4.3 Other documentation OpenOffice has written the most complete spec of the Excel format, and the Jakarta POI project also has some useful documentation. There doesn't appear to be much other up-to-date information on the format, although millions of pages have probably been written on using and coding to Excel.

4.4 Adoption Although I haven't been able to find verifiable numbers, I've seen estimates of Excel's share of the spreadsheet market as being at about 90%.

4.5 Competition Excel's primary commercial competitor is Lotus 1-2-3. In 2003 OpenOffice.org released its OpenOffice suite, which is the main Open Source spreadsheet package.

4.6 Licensing and patent claims

4.7 Other preservation issues Excel spreadsheets can contain various items that complicate preservation strategies, most notably macros in Visual Basic, and other Compound Documents.

5.Lifecycle

5.1 Version duration 1.5 years?

5.2 Version history Because of the lack of published specs, it is challenging to track down release dates of previous Excel versions. A comprehensive list does not appear to be anywhere on the Microsoft site. Pronom lists only versions 4.0 – 9.0. I took this list from the Wikipedia ([Wikipedia]) and website of the consultant John Walkenbach ([Walkenbach]).

| | |
|-------------------|-----------------|
| 1.0 | 1985 (Mac only) |
| 2.0 | 1987 |
| 3.0 | 1990 |
| 4.0 | 1992 |
| 5.0 | 1993 |
| 7.0 (Excel 95) | 1995 |
| 8.0 (Excel 97) | 1997 |
| 9.0 (Excel 2000) | 1999 |
| 10.0 (Excel 2002) | 2001 |
| 11.0(Excel 2003) | 2003 |

5.3 Expected new version Excel 11 was released in late 2003. Microsoft generally releases a new version approximately every 1.5 years.

6.Local use (DSpace at MIT)

6.1 Holdings as of January 2005: 5

6.2 Support level known

7. References

[OpenOffice CompoundDoc] OpenOffice Microsoft Compound Document Format Specification
<http://sc.openoffice.org/compdocfileformat.pdf>

[OpenOffice Excel] OpenOffice Microsoft Excel Format Specification.
<http://sc.openoffice.org/excelfileformat.pdf>

[POI] Jakarta POI project
<http://jakarta.apache.org/poi/>

[Walkenbach]Walkenbach, John, “The Excel Page”, J-Walk & Assoc. web site.
<http://www.j-walk.com/ss/excel/index.htm>

[Wikipedia] Wikipedia: Microsoft Excel
http://en.wikipedia.org/wiki/Microsoft_Excel