



OOXML: Backwards Compatible?

The best way to store old documents is in the binary format itself.

The stated goal of OOXML as amended at the Ballot Resolution Meeting is “to represent faithfully the existing corpus of word-processing documents, spreadsheets and presentations that have been produced by Microsoft Office applications (from Microsoft Office 97 to Microsoft Office 2008 inclusive).”¹ Microsoft claims that “all the features and functions of Office can be represented in XML and all your older Office documents can be moved from their binary formats into XML with 100 percent compatibility.”²

Is “backwards compatibility” actually achieved with OOXML? With 100 percent fidelity, as is claimed? Just how necessary is “backwards compatibility”?

OOXML is a file format, not an application. It cannot faithfully recreate the look of anything. Microsoft Office is the application that interprets OOXML, and it can also render legacy binary file formats. To translate a document in a legacy file format with anything near 100 percent fidelity you need complete and clear documentation for the old format *and* how it translates to the new one (the so-called “mappings”). You also need software that can read and write both with 100% accuracy and completeness. For this, Microsoft would have to grant licenses in perpetuity for use of the legacy applications and associated environments.³ The claim that OOXML – a file format – is the solution for rendering legacy documents is simply false.

No other application supporting OOXML will be able to faithfully or fully recreate the look of Microsoft's legacy binary documents. The binary Office specifications have been posted by Microsoft, but Microsoft has failed to deliver the “mappings” during the BRM that tell you how to translate a binary document into OOXML, or provide any guidance whatsoever on how to “represent faithfully” legacy documents. Given the same binary document, Microsoft Office, Apple iWork, OpenOffice.org, etc., will all produce different OOXML documents, breaking interoperability and preventing the realization of OOXML's stated goal of preserving legacy documents.

The most serious threat to “backwards compatibility” comes from actions taken by Microsoft itself to cut off access to old formats and features. After you install Office 2003 SP3, some Microsoft Office Excel 2003, Microsoft Office PowerPoint 2003, Microsoft Office Word 2003, and Corel Draw (.cdr) file formats, among others, were blocked.⁴ There can be further problems when Microsoft decides not to support legacy features, such as when it removed support for Visual Basic scripting in Office 2008 for

1 http://www.ecma-international.org/news/TC45_current_work/TC45_available_docs.htm

2 Chris Capossela, Vice President, Product Management, Microsoft at http://www.microsoft.com/interop/letters/new_world_of_docs.msp

3 See *Preserving legacy files with Ecma Office Open XML* by Chris Puttick, Oxford Archaeology at <http://www.odf-eag.eu/repository/white-papers/preserving-legacy-files-with-ecma-open-office-xml-msooxml>

4 <http://www.eweek.com/c/a/Windows/Microsoft-Takes-Heat-for-Office-2003-SP3-File-Format-Blocking/>

the Macintosh.⁵

To be sure, Microsoft frequently caveats its claims regarding backwards compatibility – "backwards compatible *with the existing functionality*", "was designed to *faithfully* represent", "*designed to best take the legacy of the past* as a building block toward more powerful solutions in the future" (emphasis added). Whatever OOXML was *designed* to achieve or *faithfully* represent, it clearly does not provide full, 100 percent backwards compatibility.

Not only does OOXML fail to provide sufficient information to provide 100 percent fidelity for legacy formats, but its sole reference implementation, Microsoft's Office 2007, also fails the test of 100 percent fidelity. It takes only opening and saving a few old binary documents and trying to save it in OOXML before receiving the following error message: "This document contains one or more of the following features that are not supported by the selected file format [OOXML]".⁶

The best way to store old documents is in the binary format itself, or in PDF. For preservation of a document in an editable form, you just need a format that can represent the existing content of the document, both visible and hidden. For this Open Document Format (ISO/IEC 26300:2006) is more than sufficient. Where exact reproduction is needed, rather than preservation of content and meaning, PDF/A (ISO 19005:2005), which guarantees a visually-identical document in a preservable format, would be appropriate. 100% fidelity is only achievable with PDF/A because presentation and word processor software packages are affected by issues such as screen resolution, the version or vendor of the software package, the installed printer, fonts, etc., all of which can significantly change a document's visual appearance.

Since the principal justification for having a second ISO-approved standard – backwards compatibility with billions of existing Microsoft Office documents – is not valid, National Bodies should reaffirm their original vote during the ballot period to disapprove OOXML/DIS 29500.

⁵ <http://www.macworld.com/article/54320/2006/12/microsoft.html>

⁶ <http://www.robweir.com/blog/2007/03/compatibility-according-to-humpty.html>