

**LJN**LAW JOURNAL  
NEWSLETTERS

LJN's

# LEGAL TECH

Newsletter®

An *incisivemedia* publication

Volume 27, Number 1 • April 2009

## PRACTICE TIP

# Microsoft's New Office 2007 File Format Is More Than an Upgrade

By Randall Farrar

In March 2008, Microsoft received approval on its new OpenXML file format from the International Organization for Standardization ("ISO"), a step Microsoft says is "proof it is willing to make once-proprietary technology work openly with competing programs." Good riddance! The old binary format was a mess and didn't play well with other applications.

The ISO approval is significant because Microsoft is no longer in control of the new file format; responsibility for maintaining standards falls to groups like the ISO, not the companies that developed them.

The new file format makes it easier for many applications to read and write the contents of a file. This helps ensure OpenXML files will

still be readable even if the program used to create them has changed significantly. This means no more backward compatibility issues. In many ways, Microsoft has veered from its common hyper-competitive approach to expose its venerable Microsoft Office to present and future competitive pressures.

The use of OpenXML will soon become the default document standard for law firms. The "ivory tower" shroud that led to both confusion and discouragement over the use of OpenXML is no longer relevant as it is inevitable that OpenXML will become a document standard. Understanding the basics of the format will illustrate that it is a much better format to manage and work with than the old binary format. Understanding is the first step in acceptance.

### OPENXML TO BECOME INDUSTRY STANDARD

In addition to better file format operability, OpenXML is destined to be the required file format for government agencies, courts and

law firm clients. Over the next two years, this will place tremendous pressure on law firms to use OpenXML as the default standard instead of Microsoft Office's old binary format.

A case in point: In 2005, Massachusetts directed state government offices to save all documents as "OpenDocument" format files (backed by Sun Microsystems and IBM), already an ISO standard. Later, Massachusetts reversed its stance after Microsoft promised to turn over responsibility for the OpenXML format to Ecma International, a European standards body. In January 2007, Ecma submitted OpenXML to ISO, to speed-up the approval process. Microsoft saw the writing on the wall — open up the file format or die a slow death of application attrition.

### WHAT IS OPENXML?

Microsoft's new file format is just plain text. Yes, plain text! Theoretically, it can be read by a common text editor. However, that's a broad based definition. The format is Extensible Markup

---

**Randall Farrar** is the president and co-founder of Esquire Innovations, a provider of Microsoft Office integrated practice management software services and applications for the legal market. He can be reached at [randall.farrar@esqinc.com](mailto:randall.farrar@esqinc.com).

Language (“XML”). XML is a general-purpose text specification for creating custom markup languages. By languages, XML allows an application to define the “markup elements” to aid information systems and applications to share a common data structure.

With OpenXML, text can easily be read by many other applications besides Microsoft Word. This opens up the opportunity of new word processors, viewers and converters that can read and write OpenXML documents. There are translators now that allow for interoperability between applications based on ODF (Open Document Format) and OpenXML-based Office applications. Google now supports OpenXML in its Gmail and Google search results.

### ADVANTAGES OF OPENXML

The OpenXML has many advantages over the binary format, such as compact files, improved damaged-file recovery, better privacy and more control over personal information, better integration and interoperability of business data, easier detection of documents that contain macros and more opportunities for different word processors to access the market and become viable competitors to Microsoft.

### THE ANATOMY OF A MICROSOFT WORD OPENXML DOCUMENT

It is relatively easy to inspect a Microsoft Word OpenXML (“.Docx”) document without Microsoft Word. A Docx (OpenXML file) is a combination of 10 or more xml files contained

<b>Part</b>	<b>Path</b>	<b>Description</b>
app.xml	root\docProps	Defines application file properties. These properties include the number of characters, words, lines, paragraphs, and pages in the document.
core.xml	root\docProps	Defines core file properties includes creator name, creation date, title, and description.
styles.xml	root\word	Defines the styles in the document.
footnotes.xml	root\word	Contains the footnotes of the document.
endnotes.xml	root\word	Contains the endnotes of the document.
comments.xml	root\word	Contains the comments in a document.

in one Zip file (a “package”). The only difference between an OpenXML file and a Zip file is the extension. To inspect a Microsoft Word 2007 Docx File a user only needs to save a Word 2007 document to his or her desktop and rename the file by adding a “.zip” extension. This will open it in Windows Explorer (or your default Zip application) and will display the OpenXML root level file package.

Each xml file in the OpenXML file is called a “part.” Within the package part is the main document part. When inspecting the main document part, you’ll notice many elements that you know are in your document, elements like styles and content. By inspecting the main document part, law firm support staff can better inspect trouble documents or documents that are considered “corrupt” at a much more granular level than documents in the old binary format.

### INSPECTING OTHER OPENXML DOCUMENTS PARTS

In addition to inspecting the main document part, there are other document parts of Microsoft Word OpenXML documents that can be

inspected. The chart below shows some of these additional parts.

### SUMMARY

As more government agencies begin requiring OpenXML as the default file format, it will become imperative for law firms to use it as the firm’s document standard instead of the old unwieldy binary format. The format is structured in a text file as XML, which makes it easily understood. Couple this with the advantages of using OpenXML and you now have a compelling reason to upgrade to Office 2007 or at least start saving all Microsoft Office documents in the new format.

