



OOXML and ISO: What's At Stake for Governments?

Background

This paper considers the importance to governments of ISO's vote on DIS 29500 "Office Open XML" (OOXML). Though several thousands of pages of comments were submitted by National Bodies (NB) highlighting problems with OOXML – both technical and non-technical – this paper addresses just a few that were raised by NBs during the ballot period and are specific to the needs of governments.

Introduction

Although the discussions about document formats are technical in nature, there is a lot hanging in the balance for governments. The demand for open standards as a foundation in new government enterprise architectures continues to grow. Typically, these architectures or frameworks identify open standards that support the goal of building a service-oriented architecture which would both enable its IT components to interoperate seamlessly, and, importantly, encourage citizen-centric, user-friendly electronic interaction with the public.

Open standards policies are proliferating as governments seek to create IT architectures that allow multiple vendors to compete directly based on the features and performance of their products, and to enable technology solutions that can be removed and replaced ("plug-and-play") with that of another vendor with minimal effort and without major interruption. Government agencies, in particular, are becoming conscious of the need to provide easy access to electronic documents to all stakeholders, while not requiring them to purchase a particular brand of software in order to view or edit those documents.¹

The growing pressure from governments mandating the use of open standards to facilitate the delivery of eGovernment services and cut costs has raised the ante on openness and standardization. ISO has become the gold standard for global standardization. The organization comprises 140+ national standards bodies representing the broadest possible international consensus of stakeholder groups – business, society, and governments. Importantly for OOXML, ISO approval can influence government IT procurement, as many governments make broad use of recognized standards in support of their policies and legislation requiring "open standards."

Governments Need The Choice That Comes With A Multi-Vendor Format

ISO approved the OpenDocument Format (ISO 26300:2006) as an international standard for document formats in May 2006. Microsoft has chosen not to support the international standard, but rather to develop its own specification, OOXML, and is seeking to have it approved as an alternative ISO standard. The European Union, in the *Pan European eGovernment Services Committee (PEGSCO) Conclusions and Recommendations on Open Document Formats*², noted the increased administrative burdens and incompatibilities

resulting from the potential need to publish documents in multiple formats. “The potential arrival of a second international standard for revisable documents,” the report concluded, “may mean that administrations will be required to support multiple formats leading to more complexity and increased costs.” Apart from these burdens and the dubious notion of having two “standards” covering the same space, OOXML poses many other serious challenges that will force governments to think twice before beginning to use it.

During the five-month ballot period for OOXML, the British Standards Institute commented that “there was no other proven implementation of OOXML apart from Office 2007.” To date, OOXML has been implemented in a single, proprietary product, Microsoft's Office 2007, meaning there would be no true choice for governments and their citizens.

OOXML's complexity, extraordinary length (more than 6,000 pages), technical omissions and single-vendor dependencies combine to make alternative implementations unattractive as well as legally and practically impossible. While some vendors have announced partial support for OOXML through the ODF-OOXML translator, this tool merely saves to OOXML, and, in its current state of development, would not be suitable in a government environment requiring collaboration on a document.

Moreover, OOXML is also either dependent upon, or optimized for, a catalog of Microsoft software applications and platforms and does not function fully with non-Microsoft software. Platform dependencies of OOXML, which are features that can only be implemented or optimized for Windows and will break or not function the same way in non-Microsoft environments, include DevMode and GUID, among others. Application dependencies include VBA macros contained in OOXML documents that will not run when outside Microsoft applications, and an enumerated list of border art means that every application that wishes to fully comply with OOXML must somehow license the use of those graphics.

The practical effect of such dependencies is that the use of OOXML by governments will require the purchase of licenses for Microsoft operating systems on both the desktop and the server as well as Microsoft Office 2007 at considerable cost to taxpaying citizens, and force those very citizens to use a particular brand of software to view or edit public documents.

OOXML Provides Governments No Guarantee Of Compatibility With Existing Microsoft Office Documents

The overview to OOXML provided to National Bodies (NBs) states that “OpenXML was designed from the start to be capable of faithfully representing the pre-existing corpus of word-processing documents, presentations, and spreadsheets that are encoded in binary formats defined by Microsoft Corporation.” Proponents of OOXML argue that this “backwards compatibility” with billions of existing Microsoft Office documents will allow governments to preserve legacy digital material and future heritage material that is now being born digital.

Yet, the American National Standards Institute (ANSI), Brazil's Associação Brasileira de Normas Técnicas, and other NBs noted in their comments submitted during the ballot period that “full compatibility of OOXML with existing Microsoft Office documents cannot be technically analyzed or proven because the present specification does not give a mapping reference between the proposed family of XML schemas and the related legacy binary and proprietary file formats supported by Microsoft Office.” These NBs proposed to “add to the

current specification all necessary mappings between the legacy binary file formats supported by Microsoft Office and the proposed XML schemas, including here all versions of binary formats supported by Microsoft Office (.doc, .xls, and .ppt).”

If the “backwards compatibility” argument were valid, one should be able to open a legacy file in Microsoft’s binary, proprietary format and save it in the new OOXML format without loss of data. Yet it only takes opening a few legacy files to demonstrate that OOXML fails to provide 100% fidelity for legacy formats. Moreover, where exact reproduction is necessary, governments can just as easily use of PDF/A (ISO 19005-1:2005), which guarantees a visually identical document in a preservable format, and one for which any number of free readers exist.

OOXML Does Not Provide Governments With The Cultural & Linguistic Adaptability Necessary In A Document Format

An international standard supported by governments must take a broader view and provide wide cultural and linguistic interoperability which, unfortunately, OOXML does not offer.

An example of a concern is the spreadsheet function NETWORKDAYS()²³. This function is defined by OOXML to return the number of working days between two dates, exclusive of any weekends in that interval. For some cultures, the weekend is Saturday and Sunday. For others, the days of rest are either Thursday/Friday or Friday/Saturday. OOXML does not define “weekend” and does not provide a way for the user to define it either. As implemented in Excel the function assumes the weekend is always Saturday/Sunday. This spreadsheet function is defined in a way which renders an incorrect answer for potentially billions of people across the globe. It is fair to say that OOXML lacks cultural adaptability. Compare this to the same function in OpenDocument Format, where the user may pass in an additional parameter to override the default definition of a weekend.

Second, SpreadsheetML defines a password-protection feature which, as a practical matter, only works for passwords entered in a predefined list of character sets. If you enter a password in another character set, like Armenian, your password will be silently converted into a sequence of '?' characters, making it trivial to break.

Third, WordProcessingML defines a barcode printing feature for printing envelopes. However, instead of being defined flexible, this feature has been defined in a way that allows use only with US Postal Service standards and ignores other national needs, such as the Royal Mail’s “Mailsort” system.

Fourth, WordProcessingML has a feature called “Border Styles” which lists a large number of graphical borders which can be used as page borders. These represent a closed list of specific named border styles with mandated images. An example of two such graphics is shown in the figure below.

earth1 (Earth Art Border)	<p>Specifies an art border consisting of a repeated image of Earth, as follows (showing two repetitions):</p> 
earth2 (Earth Art Border)	<p>Specifies an art border consisting of a repeated image of Earth, as follows (showing two repetitions):</p> 

These are the only two possibilities for displaying a globe in a page border and neither of them show Asia. Similarly, there are graphics for birthday cakes, St. Valentine's Day cupids, painted Easter eggs, Christmas gingerbread men, Halloween Jack O'Lanterns, and other images that are appropriate for a Western cultural milieu, but have limited application elsewhere. The problem here is that this list of page border styles is a closed list. Although it matches exactly what Microsoft Word provides, a would-be implementor of OOXML may not extend this list with additional image types to better suit the cultural milieu of their customers. If they do, their documents will not be valid OOXML and the application that allows non-standard images to be used as page borders will not conform to the OOXML specification. How well does OOXML adapt to other cultures? In the case of page borders, it fails to provide adaptability.

Fifth, as mentioned previously, WordProcessingML defines a number of numeration styles for numbered lists. These numeration styles were essentially only labeled, but not defined. These styles are also defined as a closed list, again matching what Microsoft Word supports, but they are not extensible by other vendors. However, the list of styles provided is incomplete, lacking, for example, support for Armenian, Tamil, Greek alphabetic, Ethiopic and Khmer numerations, as well as the larger number of historic systems used by scholars. Cultural and linguistic adaptability suffers in OOXML because of closed-ended lists which, although they may match perfectly what Microsoft Office offers today, are not extensible by vendors in an interoperable way.

Acceptance Of OOXML Raises Serious Questions For People With Disabilities

When the Commonwealth of Massachusetts moved to adopt an open format, the subsequent important accessibility concerns raised by the disability community with respect to office documents helped raise worldwide consciousness of the impact of information-technology decisions and standards on the lives of people with disabilities.

During the ballot period, the Standards Council of Canada (SCC) and Standards New Zealand (SNZ) both raised serious questions as to the compliance of OOXML with accessibility requirements established under their respective laws. The acceptance of OOXML raises a number of important accessibility questions that should be addressed before ISO recognizes OOXML as an acceptable format. Specifically:

- **OOXML Fails to adequately support Web Content Accessibility Guidelines v.1.0** – According to a white paper authored by Microsoft (http://openxmldeveloper.org/archive/2007/07/02/Accessibility_of_Open_XML.aspx) evaluating OOXML against the Web Content Accessibility Guidelines (WCAG) v.1.0, there are seven WCAG accessibility checkpoints that OOXML failed to support, and another four accessibility checkpoints that OOXML only partially supports. Some of these checkpoints represent significant accessibility issues - including several issues that OASIS ODF Accessibility subcommittee discovered in its public, peer-review of ODF v1.0 and which were fixed in ODF v1.1.
- **OOXML raises additional accessibility concerns that apply to office documents** – Separate from the accessibility failings that Microsoft itself has noted in OOXML, there are a number of important accessibility concerns that apply to office documents not covered in an evaluation based on WCAG 1.0. These include the suitability of the document format for the creation of DAISY format digital talking books for people with print impairments, and the creation of Braille documents for the blind. The OASIS ODF Accessibility subcommittee explicitly addressed these questions in their review of ODF v1.0, and OASIS adopted additions to ODF v1.1 expressly to support DAISY. Governments should be certain of the suitability of OOXML for DAISY and Braille document creation before accepting it as an ISO-approved, acceptable open format.
- **OOXML fails to meet the critical principle of involving the accessibility community in the development of standards** – We have seen no information about whether the Ecma 376 standardization process involved disability experts and people with disabilities in its development - and especially whether this process was undertaken by a peer-review body of such individuals. Standards developed and deployed without a thoughtful accessibility review by accessibility experts and by people with disabilities too often has resulted in the erection of significant barriers to people with disabilities. The principle of involving people with disabilities was clearly articulated by the European Council eAccessibility Resolution of February 2003 (http://ec.europa.eu/employment_social/news/2003/oct/eAccessibility_en.pdf) that people with disabilities should be empowered “to take more control over the development of the mechanisms for delivering eAccessibility”, and “by support for their increased participation in standardisation bodies and technical committees.” Governments should continue to uphold this core principle and insist that all ISO-approved standards that impact people with disabilities must include people with disabilities in the development of those standards.
- **The cost of the expensive assistive technologies that are needed in order to work with the sole program that supports OOXML presents an unfair cost burden** – Governments need to recognize the dimension of affordability in the choice of document formats, not only for the cost burden on government users, but also the burden placed on citizens and individuals who exchange documents with government. That burden is greatly magnified by the cost of the expensive assistive technologies that are needed in order to work with the sole program that supports OOXML - Microsoft Office - on the sole platform where such tools work - Microsoft Windows. ISO approval of OOXML as an acceptable document format could lead to a \$1,500 expense - and the sole use of Microsoft Windows (at additional expense) - for blind

inclusion to access this format.

Conclusion

Expectations for a document format standard are high, and they should be, as governments have interests going well beyond the needs of the marketplace. As a major stakeholder in the ISO standardization process, governments should use their voice and vote to encourage their national standards bodies to withhold approval of OOXML until such time as:

- OOXML's single-vendor dependencies are eliminated to make multiple implementations both legally and practically possible, thereby creating real choice among competing software products for both governments and citizens alike;
- OOXML's "backwards compatibility" with billions of existing Microsoft Office documents can be ensured for both Microsoft Office 2007 and competing vendors by, among other necessary measures, adding to the current specification all necessary mappings between the legacy binary file formats supported by Microsoft Office and OOXML, including all versions of binary formats supported by Microsoft Office (.doc, .xls, and .ppt);
- OOXML's feature set is expanded to take into consideration the needs of a culturally and linguistically diverse global community of users; and
- OOXML achieves nothing less than the high-water mark established by the OpenDocument Format (ISO/IEC 26300:2006) to address the needs of persons with disabilities.

- i See <http://www.jisc.ac.uk/media/documents/techwatch/tsw0702pdf.pdf>.
- ii See <http://europa.eu.int/idabc/en/document/3439>.