



**Porqué Chile debería votar  
'*Desapruebo*' a la propuesta de  
estándar ISO 29.500 (ECMA-376)**

**Cristian Bravo Lillo – *cbravo@kind.cl***

# Agenda de esta presentación

1. Descripción general de OO-XML (ECMA-376 / ISO 29.500).
2. Consideraciones normativas sobre ISO.
3. Selección de problemas técnicos presentes en la propuesta de estándar:
  - a) Sobre WordprocessingML
  - b) Sobre SpreadsheetML
  - c) Otros problemas
4. Algunas conclusiones



# ¡Observaciones previas!

1. Los datos y observaciones para los argumentos presentados aquí fueron tomados de múltiples fuentes, indicadas al final.
2. Las observaciones técnicas fueron verificadas sobre la propuesta de estándar ECMA-376, disponible gratuitamente.
3. ISO 29.500 no está disponible en Internet. Se utilizó ECMA-376 como fuente.

# 1. Descripción de OO-XML (1/2)

- OO-XML significa "*Office Open XML*". No tiene relación con "OpenOffice":
  - Es un estándar aprobado por ECMA (ECMA-376).
  - Contiene tres especificaciones principales:
    1. WordprocessingML
    2. SpreadsheetML
    3. PresentationML
  - Más especificaciones adicionales:
    - DrawingML, VML, etc.
    - Esquemas de especificaciones en XML Schema y Relax NG.



# 1. Descripción de OO-XML (2/2)

- La especificación completa está dividida en 5 partes:
  - 1. Fundamentals** (178 págs.):
    - Contiene una descripción general del estándar, sus partes y organización.
  - 2. Open Packaging Conventions** (131 págs.)
    - Describe las convenciones de empaquetado de los archivos OOXML.
  - 3. Primer** (473 págs.)
    - Describe con más detalle las partes del estándar (\*.ML)
  - 4. Markup Language Reference** (5220 págs.)
    - Contiene una referencia detallada de los tags que componen cada especificación.
  - 5. Markup Compatibility and Extensibility** (43 págs.)
    - Describe compatibilidad de elementos y formas de extender el estándar.



## 2. Consideraciones normativas sobre ISO (1/2)

- **Estándar** (def. ISO/IEC 2:2004, def. 3.2):
  - “[Un] documento establecido por consenso y aprobado por un organismo reconocido que establece, a efectos de un uso común y repetitivo, normas, directrices o pautas destinadas a actividades o sus resultados, cuyo objetivo es lograr el nivel óptimo de ordenamiento en un contexto determinado. *NOTA: Las normas deben basarse en los resultados consolidados de la ciencia, la tecnología y la experiencia, y su objetivo es la promoción del beneficio óptimo para la comunidad*”
- **Según Comité Técnico Conjunto N° 1 (JTC1) de ISO/IEC** (directivas JTC1, 5ta. ed., v3.0, sección 1.2):
  - “Uno de los propósitos de la normalización de la tecnología de la información es asegurar que los productos disponibles en el mercado posean características de **interoperabilidad, capacidad de transferencia** y de **adaptación cultural y lingüística**. Por lo tanto, las normas desarrolladas deberán reflejar los requisitos (anteriores)”



## 2. Consideraciones normativas sobre ISO (2/2)

- La normativa de gobierno chilena sobre Doc. Electrónico (Ley 19799, D.S.181/2002, D.S.77/2004, D.S.81/2004, D.S.83/2004, D.S.100/2006, etc.) persigue la interoperabilidad del documento electrónico, el mismo fin anterior.
- Además de las consideraciones anteriores, un estándar ISO debería (características deseables):
  - No entrar en conflicto con otros estándares ISO.
  - Ser modular: referenciar otros estándares existentes en vez de incluir su descripción.
  - Utilizar efectivamente los estándares existentes y vigentes.



### 3. Selección de problemas técnicos (1/10)

- Problemas con WordprocessingML:
  - Contiene elementos cuyo único objetivo es **ser compatible con software específicos, declarados obsoletos por Microsoft**:
    - “autoSpaceLikeWord95”, “useWord97LineBreakRules”, “useWord2002TableStyleRules”.
  - Estos elementos no son definidos en el estándar; el mismo estándar dice:
    - *(autoSpaceLikeWord95) “To faithfully replicate this behaviour, applications must imitate the behaviour of that application, which involves many possible behaviours and cannot be faithfully placed into narrative for this Office Open XML Standard. If applications wish to match this behaviour, they must utilize and duplicate the output of those applications.”* (Parte 4, sección 2.15.3.6, pág. 1378).



### 3. Selección de problemas técnicos (2/10)

- Problemas con WordprocessingML:
  - Enumera gran cantidad de estilos de listas **no definidos**, que representan diversos sistemas de escritura, idioma y convenciones comerciales:
    - Ej.: “chicago”, “chosung”, “ideographDigital”, etc.
    - “*chicago: specifies that the sequence shall consist of characters as defined in the Chicago Manual of Style*” (sin más referencias, hay 15 versiones del Manual de Chicago): Parte 4, sección 2.18.66, pág. 1772.
    - “*ideographDigital: specifies that the sequence shall consist of sequential numerical ideographs enclosed in a circle, using the appropriate character*” (¿qué ideogramas? ¿en qué secuencia van? ¿cuáles son los “caracteres apropiados?”): Parte 4, sección 2.18.66, pág. 1774.



### 3. Selección de problemas técnicos (3/10)

- Problemas con gráficos en WordprocessingML:
  - Existen dos especificaciones: DrawingML y VML.
  - En DrawingML, existen elementos y atributos que contienen información propietaria no definida.
    - Elemento shape, atributo gfxdata: *“Specifies a base-64 package (...) that contains DrawingML content. The contents of this package are application-defined”* (Parte 4, sección 6.1.2.19, pág. 4655).
  - VML:
    - Fue propuesto por Microsoft como estándar W3C, que fue rechazado a favor de SVG, y está cubierto por patente #525857 de Nueva Zelanda (tal vez otras).
    - A pesar de que es opcional, si alguien quiere implementar WordprocessingML debe omitir esta parte o explícitamente infringir una patente.
  - ISO 8632 es un estándar para archivos de “metagráficos”. OOXML utiliza y recomienda “Windows Metafiles” o “Enhanced Metafiles” en vez de ISO.



### 3. Selección de problemas técnicos (4/10)

- Problemas con lenguaje y codificación en WordprocessingML:
  - ISO 639 especifica una lista de lenguajes en el mundo.
  - En la parte 4, sección 2.18.52 (pág. 1784), se especifica el elemento ST\_LangCode, con una lista fija (incompatible con ISO 639) de valores de lenguajes.
  - En la parte 1, sección 11.3.1 (pág. 28), se especifica un “Alternative Format Import Part”, que contiene “any content, support for which is application-defined”. Pero no dice qué codificación de caracteres debe tener esta información. Esto lleva a datos no interoperables.
  - XML establece una codificación por default (UTF-8), que es la misma establecida en la normativa chilena (D.S. 81/2004, D.S. 100/2006, D.S. 181/2002, etc.)



### 3. Selección de problemas técnicos (5/10)

- Problemas con seguridad en WordprocessingML:
  - En términos de estándares criptográficos:
    - ISO escogió y publicó el estándar “Whirlpool” (ISO 10118-3)
    - W3C publicó XML-ENC, que incluye SHA-1, SHA256, SHA512.
    - El proyecto europeo NESSIE recomienda ISO 10118-3, SHA256, SHA384 y SHA512.
    - En EE.UU., NIST recomienda SHA1, SHA224, SHA256, SHA384 y SHA512.
    - En Japón, el CRYPTEC recomienda MD5, SHA1, SHA256, SHA384 y SHA512.
    - **OO-XML no sigue ninguna de las recomendaciones anteriores y propone algoritmos propios de Microsoft**, no revisados ni aprobados por ningún cuerpo colegiado.
    - En países como Chile, la criptografía de llave pública se masificará dentro de los próximos años de la mano de aplicaciones de escritorio (firma electrónica avanzada).



### 3. Selección de problemas técnicos (6/10)

- Problemas de implementación en WordprocessingML:
  - La Parte 4, secciones 2.4.51 y 2.4.52 (entre otras) especifican 'máscaras de bits' a ser aplicadas a los estilos descritos en una instancia de documento.
  - Lo anterior no permite el despliegue de un documento con (por ejemplo) XSLT, que no posee primitivas para aplicación de máscaras de bits.
  - XSLT es un estándar obligatorio establecido en el D.S. 81/2004.



### 3. Selección de problemas técnicos (7/10)

- Problemas con SpreadsheetML:
  - Varias funciones trigonométricas **no especifican si sus argumentos son en grados o en radianes** (haciendo imposible su implementación interoperable):
    - $\sin()$ : Parte 4, sección 3.17.7.287.
    - $\cos()$ : Parte 4, sección 3.17.7.50.
    - $\tan()$ : Parte 4, sección 3.17.7.313.
    - $\text{asin}()$ : Parte 4, sección 3.17.7.12.
    - $\text{acos}()$ : Parte 4, sección 3.17.7.4.
    - $\text{atan}()$ : Parte 4, sección 3.17.7.14.
    - $\text{atan2}()$ : Parte 4, sección 3.17.7.15.
  - La función  $\text{avedev}()$  dice calcular la desviación estándar de un conjunto de valores. ¡Pero en realidad calcula la cantidad de combinaciones de un valor sobre otro ( $n$  sobre  $k$ )! (Parte 4, sección 3.17.7.17, pág. 2545).



### 3. Selección de problemas técnicos (8/10)

- Problemas con SpreadsheetML:
  - Se trata el año 1900 como bisiesto, lo que:
    - Entrega días de semana incorrectos entre el 01/01/1900 y el 28/02/1900.
    - Crea un día inexistente que ¿debe ser ignorado? (29/02/1900).
  - Posee dos bases distintas para valores de fecha (Parte 4, sección 3.17.4.1):
    - La “base 1900”, de rango 01/01/1900 – 31/12/9999.
    - La “base 1904”, de rango 01/01/1904 – 31/12/9999.
    - Los años fuera de una base determinada son considerados “*ill-formed data*”.
  - El estándar dice (misma sección anterior):
    - “For legacy reasons, an implementation using the 1900 date base system shall treat 1900 as though it was a leap year”.



### 3. Selección de problemas técnicos (9/10)

- Otro tipo de problemas:
  - Ya existe un estándar ISO para documentos, hojas de cálculo, presentaciones y mucho más: **ISO 26300**, llamado Open Document Format (ODF). ¿Necesitamos otro estándar? Tal vez sí, pero...
    - ODF posee aprox. 600 páginas de especificación y demoró 1 año y medio en ser aprobado.
    - OOXML posee aprox. 6500 páginas de especificación ... y fue aprobado como estándar ECMA en 6 meses. La especificación misma fue revisada en sólo 30 días (¿~200 páginas por día?)
    - ODF es más compacto porque hace uso intensivo de otros estándares (si existe un estándar, lo ocupa):
      - ISO 639, ISO 8601, XML, XSLT, XML Schema, SVG, PNG, XHTML, Xpath, CSS2, MathML, SOAP, etc.
    - OOXML entra en conflicto con varios estándares ISO (algunos ya nombrados).



### 3. Selección de problemas técnicos (10/10)

- Otro tipo de problemas:
  - El gobierno de Chile posee una serie de normativas basadas en estándares que pretenden favorecer la interoperabilidad.
  - El Instructivo Presidencial 5 (11/05/2001) establece que todos los Jefes de Servicio deberán implementar en sus proyectos de gobierno electrónico “estándares abiertos” (N°6, letra f).
  - OOXML no favorece la interoperabilidad, y definitivamente en varias partes de su especificación interna no es abierto (no define elementos ni comportamientos, no especifica completamente ciertos elementos, etc.)



## 4. Algunas conclusiones (1/2)

- Es notable y positivo que Microsoft haya generado una iniciativa de estandarización de sus formatos de archivos Word, Excel y Powerpoint. Pero OOXML:
  - Es una especificación (todavía) con muchos errores, inmadura y revisada apresuradamente.
  - Posee una serie de partes que están especificadas parcialmente o que no lo están.
  - Está en conflicto excluyente con varias normas ISO aceptadas y en uso.
  - Estimula el uso de sub-especificaciones que no han sido estandarizadas, y para las cuales Microsoft tiene patentes (aunque ha prometido no demandar a nadie).
  - No posee más que una implementación (MS Office), y ésta no puede ser estudiada más que como una caja negra.
- Lo anterior hace que sea imposible (para alguien distinto de Microsoft) implementar completamente el estándar OOXML.



## 4. Algunas conclusiones (2/2)

- Existen muchos escenarios de uso en los que es esencial la interoperabilidad y estandarización de la información:
  - Hoy es común encontrar aplicaciones donde bases de datos extraen información de hojas de cálculo, o Web Services que extraen información de hojas de cálculo y la disponibilizan en línea.
  - Chile estimula y favorece el uso de firma electrónica, a través de certificados X.509. La forma más sencilla de firmar un documento es firmando uno que se acaba de crear en una aplicación ofimática (*“creo una carta en mi procesador de texto y la firmo”*).
- La decisión de qué estándares aceptamos tiene una importancia estratégica, económica y social para los próximos 50 años o más.



# Fuentes (y agradecimientos)

- **Fuentes:**

1. “Microsoft OOXML / ECMA376, Get the Facts”, Anand Vaidya, bajo licencia Creative Commons:
  - <http://www.asianlinux.org/downloads/docs/advocacy/ODF-vs-OOXML-v1.1.pdf>
2. “Microsoft and Open Standards, Can Other vendors Implement Microsoft's Office Open XML?”, 15/08/2007
  - <http://holloway.co.nz/can-other-vendors-implement-ooxml.html>
3. Estándar ECMA-376:
  - <http://www.ecma-international.org/publications/standards/Ecma-376.htm>
4. <http://www.nooolxml.org>
5. <http://www.odfalliance.org>
6. <http://www.iso.org>
7. <http://www.openxml.info>



# Observaciones finales

- Muchas gracias por vuestro tiempo, paciencia y atención.
- Esta presentación fue creada con OpenOffice 2.0 sobre Ubuntu 6.10.
- Esta presentación está disponible bajo petición por email al autor en formatos PDF y ODP bajo licencia Creative Commons del tipo “Atribución-No Comercial-Licenciar Igual 2.0 Chile”:
  - <http://creativecommons.org/licenses/by-nc-sa/2.0/cl/>
  - No puedes usar ni distribuir esta presentación con fines comerciales de ningún tipo.
  - Debes atribuir al autor (Cristian Bravo Lillo, [cbravo@kind.cl](mailto:cbravo@kind.cl)).
  - Si modificas este documento, comparte las modificaciones conmigo y con el resto del mundo.



ubuntu

 OpenOffice.org

