

OpenDocument en la ofimática

# ¿QUÉ HAY DENTRO?

El formato OpenDocument nos ofrece un nuevo método para almacenar la información e intercambiar documentos para aplicaciones ofimáticas. Pero, ¿qué significa ODF para el usuario? ¿Qué hay dentro de un archivo ODF? ¿Cómo es de portable? Examinamos estas preguntas en el tema de portada de este mes. **POR DIMITRI POPOV**

**E**l proceso de adopción de Open Document Format (ODF) como formato abierto para aplicaciones ofimáticas en la Commonwealth de Massachusetts fue algo parecido a un thriller policíaco. Es complicado imaginar que un humilde formato de documento pudiera causar semejante controversia e intrigas políticas. OpenDocument, sin embargo, es algo más que un simple formato de documento.

ODF representa una nueva manera de afrontar el problema del intercambio de información para aplicaciones de escritorio, y anuncia el fin de la era en la que los fabricantes de software podían ejercer el control del mercado de las herramientas ofimáticas mediante el control exclusivo del conocimiento de los formatos de archivos.

El verdadero objetivo, sin embargo, no es vencer a Microsoft o promocionar OpenOffice.org, y el objeto del debate no es si el formato Word es superior técnicamente al de OpenOffice Writer. El objetivo último es evitar pérdidas de información en nuestra sociedad, que es lo que ocurre a menudo debido a los formatos propietarios. Esta meta sólo podrá alcanzarse mediante la transición a formatos

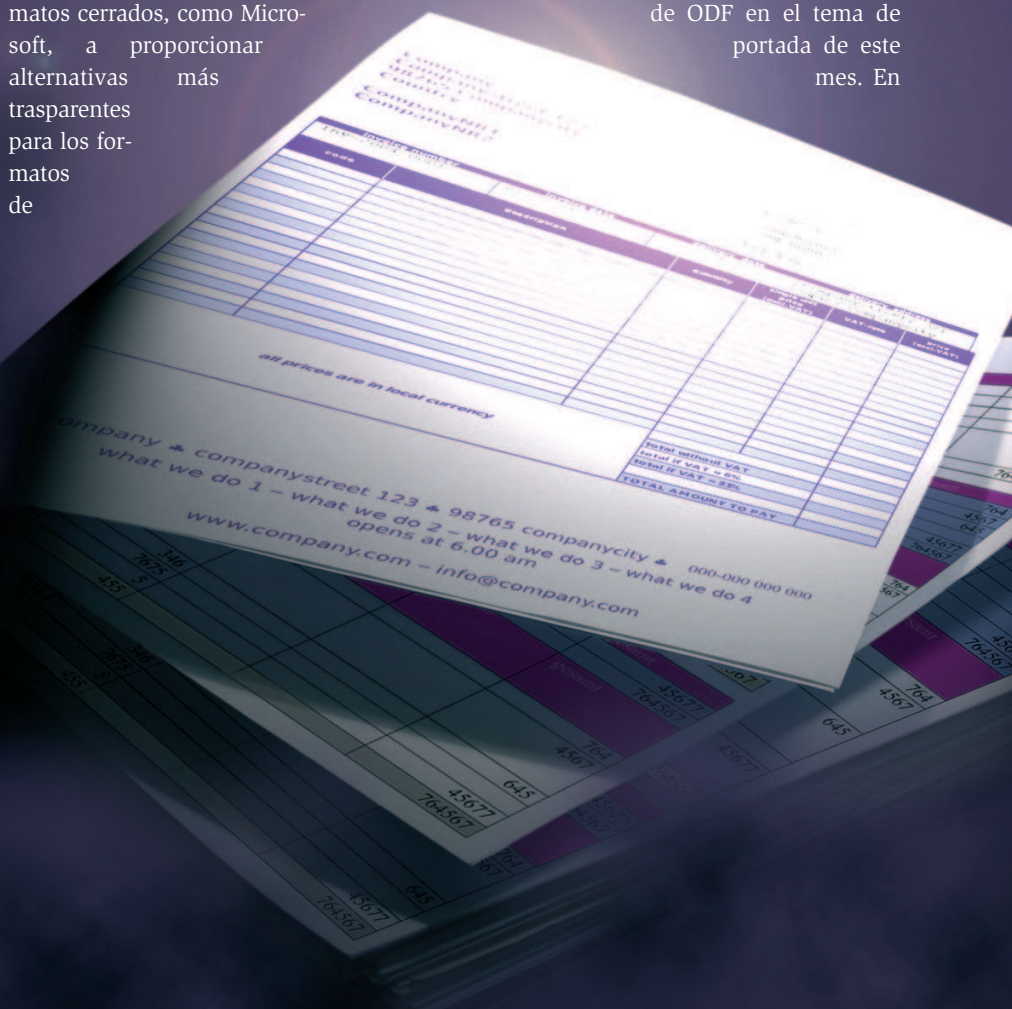
abiertos, como el OpenDocument Format.

El formato OpenDocument es un estándar desarrollado por la "Organization for the Advancement of Structured Information Standards" (OASIS). El propósito original del formato OpenDocument era estandarizar el uso de XML como formato para almacenar e intercambiar la información entre aplicaciones de escritorio. La aparición del formato OpenDocument ya ha forzado a fabricantes de formatos cerrados, como Microsoft, a proporcionar alternativas más transparentes para los formatos de

almacenamiento de información.

Sin embargo, a pesar de toda la atención que ha recibido recientemente ODF en la prensa, y a pesar de que usuarios de todo el mundo ya usan el formato OpenDocument en OpenOffice 2.0 en su trabajo diario, se ha escrito bastante poco de él en sí mismo. ¿Qué es ODF? ¿Cuán compatible es con los paquetes ofimáticos alternativos?

Hacemos un detenido repaso de ODF en el tema de portada de este mes. En



## TEMA DE PORTADA

Compatibilidad ODF .....	20
ODF para Word .....	25
Suites Ofimáticas .....	27
Herramientas Ofimáticas en web .....	33

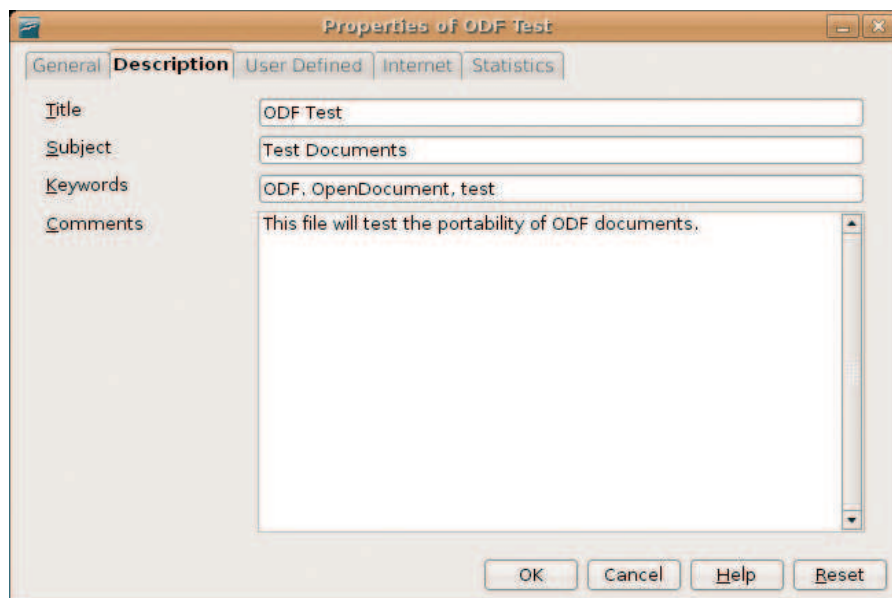


Figura 1: Gran parte de la información guardada en meta.xml se puede ver a través del cuadro de diálogo Properties en OpenOffice 2.0.

este artículo os mostramos cómo bucear dentro de un documento en OpenDocument Format y cómo está organizada la información en su interior.

En posteriores entregas ponemos a prueba la portabilidad de los documentos

ODF, y describimos cómo convertir un archivo ODF de manera que pueda abrirse con Microsoft Word.

Continuando con este tema ofimático, nos detenemos en algunas populares alternativas a OpenOffice y examinamos el

reciente fenómeno de las aplicaciones ofimáticas basadas en Web.

Esperamos que disfrute este repaso al formato OpenDocument y a las alternativas ofimáticas bajo Linux. Para una información más detallada de su anatomía del formato OpenDocument, puede recurrir al libro en PDF *OpenDocument Essentials* escrito por J. David Eisenberg, editado bajo licencia GNU Free documentation, y disponible en [2].

## ¿Qué Hay Dentro?

Los documentos ofimáticos incluyen todo un conjunto de cosas más amplio que simplemente texto y números. Un documento típico contiene también información de estilo y de formato, además de información del usuario que ha creado el archivo y otros elementos útiles para la aplicación.

De hecho, lo que tomamos como un archivo ODF no es simplemente un archivo. Un archivo en formato OpenDocument es en realidad un conjunto de archivos guardados como un Java Archive (JAR), que es un tipo de archivo comprimido zip. Para mirar dentro de un documento en formato OpenDocument, todo lo

### Listado 1: content.xml

```
01 <text:h text:style-name="Heading_20_1" text:outline-level="1">Lorem ipsum</text:h>
02 <text:p text:style-name="Standard"/><text:p text:style-name="Standard">Lorem ipsum dolor sit amet.
03 <text:span text:style-name="T1">Proin velit.</text:span>
```

### Listado 2: manifest.xml

```
01 <?xml version="1.0" encoding="UTF-8"?>
02 <manifest:manifest xmlns:manifest="urn:oasis:names:tc:opendocument:xmlns:manifest:1.0">
03 <manifest:file-entry manifest:media-type="application/vnd.oasis.opendocument.text" manifest:full-path="/"/>
04 <manifest:file-entry manifest:media-type="text/xml" manifest:full-path="content.xml"/>
05 <manifest:file-entry manifest:media-type="text/xml" manifest:full-path="styles.xml"/>
06 <manifest:file-entry manifest:media-type="text/xml" manifest:full-path="meta.xml"/>
07 <manifest:file-entry manifest:media-type="text/xml" manifest:full-path="settings.xml"/>
08 </manifest:manifest>
```

### Listado 3: meta.xml

```
01 <meta:generator>OpenOffice.org/2.0$Unix OpenOffice.org_project/680m5$Build-9073</meta:generator>
02 <meta:initial-creator>Dmitri Popov</meta:initial-creator>
03 <meta:creation-date>2006-11-30T11:47:40</meta:creation-date>
04 <dc:creator>Dmitri Popov</dc:creator>
05 <dc:date>2006-11-30T11:58:51</dc:date>
06 <dc:language>en</dc:language>
```

que debemos hacer es cambiar la extensión *.odt* a *.zip* y simplemente desempaquetarlo. El documento descomprimido contiene, habitualmente al menos, los siguientes archivos:

- content.xml
- META-INF/manifest.xml
- meta.xml
- mimetype
- settings.xml
- styles.xml

El contenido exacto del documento puede variar. Por ejemplo, si éste contiene imágenes y macros, estos elementos se incluyen en carpetas aparte.

## Los Archivos

Cada uno de los archivos dentro de un documento ODF tiene un propósito específico. *content.xml* es el archivo clave que contiene el verdadero texto del documento. Si lo abrimos con un editor de textos, veremos que el etiquetado XML usado en este archivo tiene muchas similitudes con HTML.

Incluso si no somos expertos en XML, podemos identificar fácilmente algunas etiquetas básicas. En el Listado 1, por ejemplo, la cadena “Lorem ipsum” se formatea como Heading 1 y la frase “Proin velit” se marca como negra.

Como en un archivo HTML, *content.xml* incluye tanto el texto como las instrucciones de formateado. Los comandos de formateado asocian el texto con el tipo de letra y la información de estilo. En documentos más elaborados, las etiquetas deberían especificar dónde situar los cuadros o imágenes dentro del archivo.

El resto de componentes dentro del archivo son bien archivos de contenido auxiliar (como imágenes incrustadas en el documento) o archivos que proporcionan alguna información secundaria. El número exacto de archivos puede variar en función de los documentos y la aplicación que los generó. Algunos archivos, sin embargo, son especialmente importantes.

El archivo *manifest.xml* contiene un directorio de todos los archivos del documento (véase el Listado 2). Este archivo se encuentra dentro del directorio *META-INF*.

El archivo *meta.xml* contiene los metadatos del documento, incluyendo la aplicación usada para crear el documento, el nombre del autor, la fecha de creación, el idioma y un contador de palabras (véase el Listado 3). La información de *meta.xml* será familiar para los usuarios que han

examinado el cuadro de propiedades del documento en OpenOffice.org (Figura 1).

Las distintas configuraciones del documento se guardan en *settings.xml*, el cual incluye parámetros como el Field Auto Update, Save Version on Close, Printer Name y demás.

De acuerdo con las especificaciones, *settings.xml* está diseñado para guardar la información que se usa por parte de la aplicación que crea el archivo. Sin embargo, estas configuraciones pueden no

estar soportadas por otras aplicaciones que abran el documento ODF con posterioridad.

La mayoría de los procesadores de texto modernos ofrecen la posibilidad de asociar un bloque de texto o un párrafo con un estilo (véase la Figura 2). El archivo *styles.xml* de ODF contiene la información de estilo asociada con este documento.

Los estilos ofrecen un medio para separar los detalles de formateado

## ¿Qué es un Formato Abierto?

Aunque la industria no tiene aún una lista de criterios oficiales que defina un “formato abierto”, es importante tomar en consideración varios aspectos.

La primera característica de un formato abierto es el desarrollo de la especificación abierto y colaborativo. ODF, como cualquier otro formato abierto, es el resultado de muchas discusiones y un largo proceso de desarrollo que implica a numerosas partes. Estas compañías y organizaciones son miembros de OASIS, una organización que coordina el desarrollo y aprobación de los estándares abiertos en general, y de ODF en particular. Mientras que los miembros de OASIS participan directamente en la creación de las especificaciones de ODF, cualquier persona puede seguir el desarrollo y acceder a la información relevante. OASIS no es un club excluyente: cualquiera puede enrolarse, o al menos, mantenerse al día del desarrollo. Si queremos sumar nuestros esfuerzos, podemos empezar por echarle un vistazo a OASIS OpenDocument Technical Committee [7], a OASIS ODF Adoption Committee [8], o a OpenDocument XML.org [9].

Un estándar verdaderamente abierto debe estar libre de restricciones de patentes o de licencias. Este requerimiento puede parecer muy obvio, pero es uno de los más importantes para un formato abierto. Esto asegura que el desarrollador no tenga que obtener una licencia o pagar royalties a terceros por usar el formato. Como usuario final, podemos usar el formato para guardar nuestros documentos sin preocuparnos de potenciales demandas por infringir alguna oscura patente o por violar los términos de la licencia. Por supuesto, algunos fabricantes y organizaciones usan el término “abierto” para estándares que otros no considerarían como tales. El complejo problema de las licencias y las restricciones por patentes es un tema en constante debate a lo largo de toda la industria del software.

Un formato abierto debe estar libre de dependencias propietarias y de funcionalidades de un fabricante único. ODF no depende de ninguna tecnología propietaria, y esto permite que ninguna compañía añada funcionalidades según le convenga. Esto, entre otras cosas, evita la práctica ejemplificada por el intento de Microsoft de añadir algunas extensiones propietarias a Java, a pesar de convertir a su versión incompatible con el resto del mundo.

Como formato abierto, ODF también asegura una implementación sencilla para los desarrolladores de software en sus propios programas, así como la interoperabilidad con cualquier aplicación que soporte ODF. Como todas las especificaciones de ODF están disponibles y bien documentadas, debería ser extremadamente fácil para cualquier desarrollador añadir soporte ODF a una aplicación existente, o generar nuevo software que use ODF. Puede encontrar una lista exhaustiva de software que soporta ODF en el artículo de la Wikipedia [10] y en la página de software soportado en la página OpenDocument Fellowship.

Con el lanzamiento de Office 2007, Microsoft ha introducido un nuevo formato de documento basado en XML. Para entender la perspectiva del software libre acerca de las diferencias entre el formato OpenDocument y los formatos XML de Microsoft, podemos leer “Format Comparison Between ODF and MS XML” [12]. La Wikipedia proporciona un resumen de este artículo para aquellos que busquen un texto más rápido de digerir [13].

Si queremos comprender mejor los diferentes aspectos de la definición de los estándares abiertos, podemos escuchar la entrevista a Danese Cooper en Open Source Conversations [14].

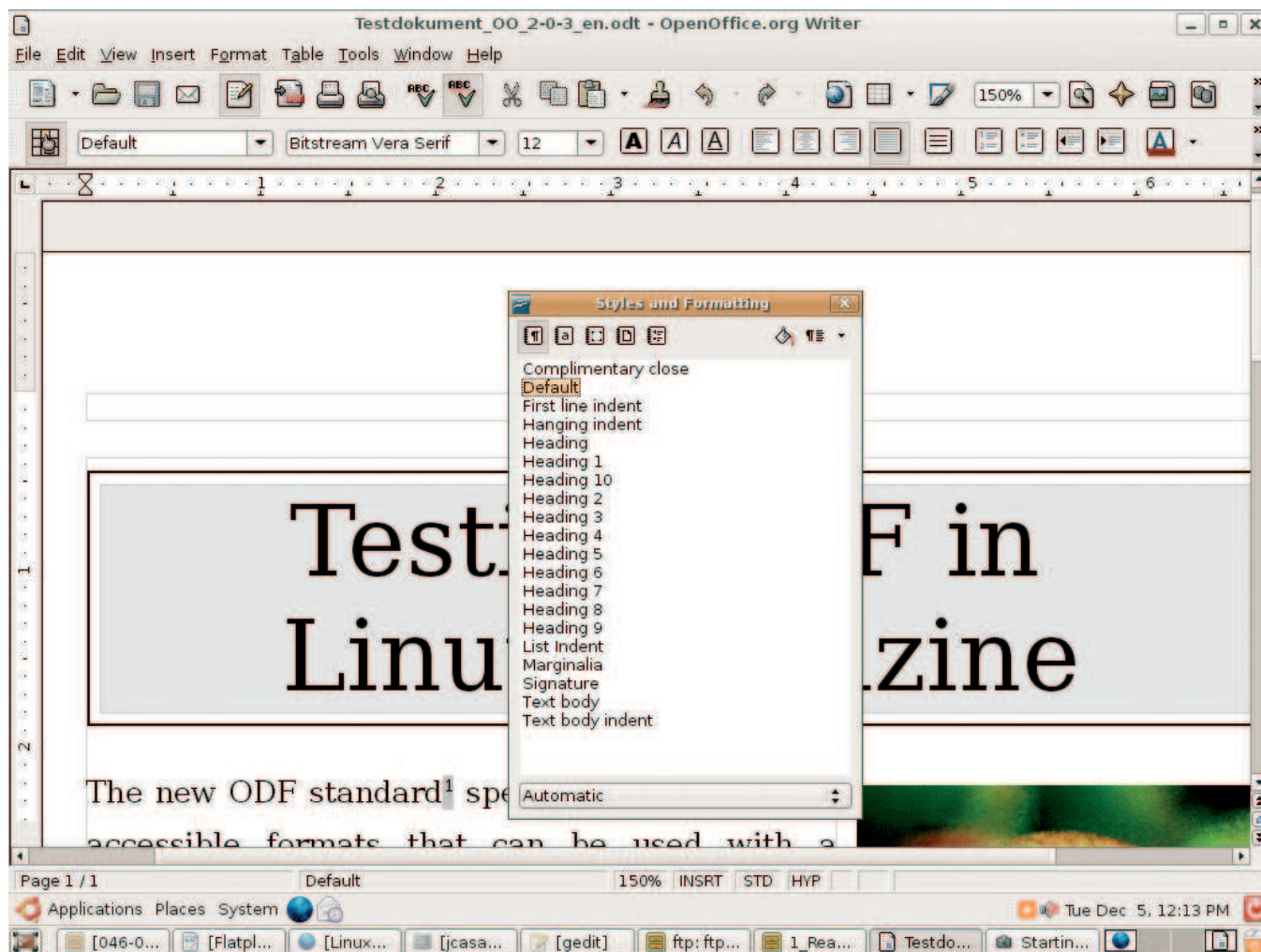


Figura 2: La mayoría de los procesadores de texto permiten al usuario dar formato al documento según estilos predefinidos.

del contenido, lo que mantiene al archivo *content.xml* simple y más manejable.

El archivo *mimetype* contiene el tipo MIME del documento. Algunos de los posibles tipos MIME para el

formato OpenDocument se muestran en la tabla.

En nuestro caso, el tipo MIME de un archivo de texto es *application/vnd.oasis.opendocument.text*.

### OpenDocument Automatizado

XML está diseñado para una fácil automatización, y la total simplicidad y legibilidad del formato OpenDocument permite al desarrollador realizar herramientas

Tabla: Tipos de Archivo OpenDocument

Extensión	Descripción	Tipo MIME
odt	documento de texto	application/vnd.oasis.opendocument.text
ott	documento de texto usado como plantilla	application/vnd.oasis.opendocument.texttemplate
oth	documento de texto usado como plantilla para HTML	application/vnd.oasis.opendocument.textweb
odm	texto global del documento	application/vnd.oasis.opendocument.textmaster
odg	documento de dibujo	application/vnd.oasis.opendocument.graphics
otg	documento de dibujo usado como plantilla	application/vnd.oasis.opendocument.graphics-template
odp	documento de presentación	application/vnd.oasis.opendocument.presentations
otp	documento de presentación usado como plantilla	application/vnd.oasis.opendocument.presentations-template
ods	documento de hoja de cálculo	application/vnd.oasis.opendocument.spreadsheet
ots	documento de hoja de cálculo usado como plantilla	application/vnd.oasis.opendocument.spreadsheet-template
odc	documento de gráficos	application/vnd.oasis.opendocument.chart
otc	documento de gráficos usado como plantilla	application/vnd.oasis.opendocument.chart-template
odi	documento de imagen	application/vnd.oasis.opendocument.image
oti	documento de imagen usado como plantilla	application/vnd.oasis.opendocument.image-template
odf	documento de fórmula	application/vnd.oasis.opendocument.formula
otf	documento de fórmula usado como plantilla	application/vnd.oasis.opendocument.formula-template

tas que soporten ODF en muy poco tiempo. Por supuesto, el número de posibles configuraciones y opciones dentro de un archivo ODF es casi interminable. Pero a menos que estemos escribiendo nuestra propia aplicación desde cero, no vamos a necesitar automatizar todo el espectro de opciones de estilo y configuración.

Una estrategia rápida y sencilla para escribir un programa que genere un archivo ODF es crear un documento que tenga el formato y las funcionalidades del archivo que querríamos que generara el programa, luego desempaquetar este archivo ODF (como se describió anteriormente en este artículo) y examinar el XML.

Podemos experimentar con la configuración de los archivos XML y recargar el archivo en formato OpenDocument para ver el efecto de los cambios.

Por tanto, como escribe J. David Eisenberg en *OASIS OpenDocument Essentials*, “una vez que aprendemos cómo funciona una característica, no dude en copiar y pegar el XML del archivo OpenDocument a su programa. En otras palabras, haga trampas. A mi me funcionó al escribir este libro, y puede servirle también a usted” [2].

XML es mucho más fácil de recuperar y reparar que un formato binario. Incluso si la aplicación falla al recuperar el archivo basado en XML, aún podemos sacar la información directamente, ya que al contrario que con los archivos binarios, los archivos basados en XML se guardan de forma legible.

Un buen ejemplo de la sencillez y efectividad de una herramienta para convertir y reparar archivos de texto ODF es el script `odt2txt` [3], escrito en Python. `odt2txt` convierte documentos de Writer de OpenOffice.org a archivos en texto plano, convirtiendo el formato del documento a una sintaxis Markdown.

## Unas Palabras para Acabar

Aunque ODF ofrece una alternativa viable y abierta a los formatos propietarios, tiene algunos puntos débiles. Por ejemplo, han aparecido algunas dudas referentes a la eficiencia del formato al enfrentarse a grandes cantidades de información (véase [4] y [5]).

A medida que crece la dependencia en el almacenamiento e intercambio de docu-

mentación electrónica en nuestra sociedad, aumenta la preocupación respecto a problemas de interoperabilidad y pérdida de información debido a los formatos propietarios. El formato OpenDocument ofrece una solución a estos problemas, y es sólo cuestión de tiempo el que los gobiernos y empresarios se den cuenta de la importancia de almacenar información en formatos abiertos como ODF [6]. ■

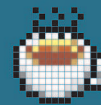
## RECURSOS

- [1] OASIS: <http://www.oasis-open.org/>
- [2] OASIS OpenDocument Essentials por J. David Eisenberg: <http://books.evc-cit.info/>
- [3] `odt2txt`: <http://www.freewisdom.org/projects/python-markdown/odt2txt.php>
- [4] Objeciones a ODF: <http://blogs.zdnet.com/Ou/?p=101>
- [5] Más objeciones a ODF: <http://blogs.zdnet.com/Ou/?p=119>
- [6] “Por qué ganó OpenDocument (y no lo hizo Microsoft Open XML): <http://www.dwheeler.com/essays/why-opendocument-won.html>
- [7] OpenDocument Technical Committee: [http://www.oasis-open.org/committees/tc\\_home.php?wg\\_abbrev=office](http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=office)
- [8] OASIS ODF Adoption Committee: [http://www.oasis-open.org/committees/tc\\_home.php?wg\\_abbrev=odf-adoption](http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=odf-adoption)
- [9] OpenDocumentXML.org: <http://opendocument.xml.org/>
- [10] Artículo en la wikipedia sobre el soporte de OpenDocument: [http://en.wikipedia.org/wiki/OpenDocument\\_software](http://en.wikipedia.org/wiki/OpenDocument_software)
- [11] Software que soporta OpenDocument: <http://opendocumentfellowship.org/applications>
- [12] “Format comparison between ODF and MS XML”: <http://www.groklaw.net/articlebasic.php?story=20051125144611543>
- [13] Comparación de ODF y MS XML en la Wikipedia: [http://en.wikipedia.org/wiki/Comparison\\_of\\_OpenDocument\\_and\\_Microsoft\\_Office\\_Open\\_XML\\_formats](http://en.wikipedia.org/wiki/Comparison_of_OpenDocument_and_Microsoft_Office_Open_XML_formats)
- [14] Danese Cooper acerca de estándares abiertos: <http://osc.gigavox.com/shows/detail1222.html>

# CISL



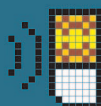
Administración remota de sistemas Linux



Desarrollo de JAVA / JBOSS



Desarrollo de sistemas empujados



Desarrollo SIM / servicios O.T.A.

# CISL

CISL Spain, S.L.

P.T.A  
C/ Ivan Pavlov 1-1  
G Alto 29590  
Campanillas (Málaga)

sales@cisl.es

Tel. + 34 952 020 408  
Fax. + 34 952 020 535