

05 Definición del formato de contenidos: documentos de oficina, gráficos, audio y video.

Antonio Araúzo Azofra
(Área de proyectos de ingeniería)

23 de febrero de 2018

Contenidos

1 Definición del formato de contenidos

- Media types
- Herramientas

2 Ofimática

3 Multimedia

- Imágenes
- Audio
- Video

Tipos de contenidos o archivos

Ejemplo: correo electrónico

- RFC822(1982) → RFC2822(2001) → RFC5322(2008)

¿Cómo determinamos el formato de un archivo?

→ extensiones → problemas

- Multipurpose Internet Mail Extensions (MIME): RFC2045, RFC2046, RFC2047, y otros.
- RFC2046 → **Media types** (aka MIME/Internet media types, or Content-types)
 - Ahora: se usa para identificar tipos de contenidos en general
 - IANA gestiona el [registro oficial de tipos de medios](#)
 - Limitación actual, servidor web tiene que adivinar el contenido para identificarlo (por extensión o magic numbers)

Meta-formatos (o documentos basados en tags)

- SGML
- XML

Compresión

- Reducción del espacio ocupado por los datos
- Usada dentro de muchos formatos
 - Sin perdida (mantiene la información intacta)
 - Con perdida (se descarta la información menos relevante)
→ Gran progreso (imagenes, audio, video, todo en internet...)
- Trabajo del grupo 10: Estándares de compresión
- Ejemplo de estándar de compresión abierto: **zip**
 - Definido para **PKZIP** por PKWARE Inc.
 - Se puede encontrar la definición completa en el [mensaje en el que se registra zip como MIME Media type](#)
 - Libre de patentes en la práctica
 - Implementación libre en **zlib**, en **Python**...
 - Usado en ficheros .zip: WinZip, zip, unzip...
 - Usado en formatos contenedores de otros datos: PNG, OpenDocument...

Open Document

- Ver documento OASIS
- Actividad: abrir un .odt. Descomprimir con unzip
 - ¿Qué indica mimetype?
 - ¿Qué significa vnd en el registro mimetype? (ver RFC6838 del proceso de registro)

Ooxml

Informático especializado en sistemas de documentación:

*“Con respecto al ooxml, efectivamente es un estándar, aunque **no es abierto**, ya que la licencia de patentes que mantiene Microsoft sobre el formato es una licencia tipo RAND (reasonable and non-discriminatory), la cual, irónicamente, no es nada razonable, y menos para el mundo del software libre. Además, **ni siquiera es completo**, ya que, a pesar de las mas de 6000 paginas que lo componen, hay definiciones y referencias que se hacen en el documento que no están definidas en ninguna parte del estándar, con lo que, si además de eso añadimos su (innecesaria, por otra parte) **complejidad**, se hace imposible implementarlo.”*

¿Por qué estandarizar un segundo formato de documentos?

PDF

La gente piensa en PDF como un formato protegido, más cerrado que .doc ...

Objetivo: documentos que se vean igual en cualquier medio en pantalla e impresos con diversos dispositivos

ISO 19005-1:2005 Document management – Electronic document file format for long-term preservation – Part 1: Use of PDF 1.4 (PDF/A-1)

ISO 32000-1:2008 Document management – Portable document format – Part 1: PDF 1.7

Adobe libera el formato dando licencia en la que se compromete a no cobrar por sus patentes

Docbook

(SGML→XML): <http://www.docbook.org/>

Contenidos para enseñanza

- SCORM http://docs.moodle.org/24/en/SCORM_FAQ#What_is_SCORM.3F
- IMS <http://www.imsglobal.org/content/packaging/>

PNG

- PNG vs GIF
- RFC 2083 → W3C
<http://www.w3.org/TR/2003/REC-PNG-20031110/> →
- ISO/IEC 15948:2004 Information technology – Computer graphics and image processing – Portable Network Graphics (PNG): Functional specification

Formatos de audio

- MPEG-1 Audio Layer 3
- OGG
- ...

Formatos de video

- Contenedores (.mp4, .avi ...)
- MPEG-4 Video (ISO/IEC 14496-2)
- H.264/MPEG-4 AVC (ISO/IEC 14496-10)
- ...

Actividad PNG

Elegido porque es un estándar manejable y de complejidad sencilla.

Objetivo: Cómo leer un estándar y encontrar la información deseada.

Encontrar en el texto del estándar la respuesta a las siguientes preguntas:

- 1 ¿Cómo se puede saber si un archivo es una imagen PNG?
→ Nota curiosa: ver la motivación en Wikipedia
- 2 ¿Cómo se guardan las imágenes para su carga progresiva?
- 3 ¿Qué es un chunk?
- 4 ¿Qué es la paleta de colores (o indexado de colores)?
- 5 ¿Qué cosas se han añadido al estándar desde ISO?
- 6 ¿Permiten transparencia? ¿Es Oblig. o auxiliar?

References / Referencias