

PDF frente a otros formatos de ficheros (EPS, PostScript, XML, ficheros nativos...)

Laurens Leurs

(Diciembre de 1999)

Hay unos cuantos formatos de ficheros informáticos cuyos resultados se acercan a los de PDF. Éstos son algunos de ellos, con sus diferencias y similitudes:

PDF frente a PostScript

Adobe, la compañía que desarrolló PostScript, fue también quien desarrolló PDF. De hecho, PDF está basado en [el lenguaje PostScript](#). PDF Usa el conjunto de instrucciones de PostScript de un modo diferente: Mientras que PostScript es realmente un lenguaje de programación que se puede usar para escribir un juego de ajedrez o un procesador de textos, PDF es de objetivos más limitados. Sólo describe la composición de un documento (usando operadores PostScript). En ese sentido, PDF recuerda más a una base de datos que a un lenguaje de programación.

Las principales ventajas de PDF frente a PostScript son:

- Los ficheros PDF tienden a ser más reducidos debido a los eficaces algoritmos de compresión que se pueden usar.
- Los ficheros PDF se pueden ver fácilmente con el lector Acrobat Reader u otras herramientas.
- Los ficheros PDF se pueden modificar más fácilmente, siempre que se tengan las herramientas apropiadas.
- Los ficheros PDF son más independientes de los dispositivos. Los ficheros PostScript suelen estar creados para dispositivos concretos y pueden generar errores PostScript si se envían a otros aparatos.
- Los ficheros PDF pueden ser más versátiles que los ficheros PostScript: Pueden contener enlaces a otros ficheros y elementos multimedia.

PDF frente a HTML

PDF se suele comparar con HTML (el formato de datos más usual en la creación de páginas *web*). En principio HTML fue creado para describir la estructura de un documento y no tanto para elaborar su apariencia. El aspecto de una página *web* viene determinado por el navegador, no por el creador del documento. Con el incremento en la popularidad de la World Wide Web, el énfasis de las nuevas versiones de HTML pasó de más la estructura a la apariencia. En cierto modo, HTML se desplazó hacia los objetivos de PDF.

Al mismo tiempo, Adobe incluyó cada vez más funcionalidades *web* dentro de PDF (se hizo posible añadir enlaces de Internet dentro de un documento PDF). Al mismo tiempo, se puso a la disposición del público una extensión para los navegadores Netscape Navigator y Microsoft Internet Explorer para poder mostrar ficheros PDF. Acrobat 4 incorporó la capacidad de convertir un sitio *web* (o parte de él) en un documento PDF. Adobe añadió a PDF la capacidad de descarga incremental de bytes (*byte streaming*) para que no fuera necesario descargar todo un fichero PDF antes de poder ver su primera página en Internet.

En cierto modo, HTML y PDF se están convirtiendo en estándares en competencia. Ahora mismo, PDF es mucho más potente en lo que se refiere describir la apariencia de documentos. HTML es más apropiado para accesos de Internet de baja capacidad. Sin embargo, es perfectamente factible usar ficheros PDF en sitios *web* y usar documentos HTML en un catálogo electrónico distribuido en CD-ROM.

PDF frente a XML

XML (eXtensible Markup Language, "Lenguaje Extendido de Marcas"), es un formato que se puede usar para describir el contenido de un documento (similar a SGML, del que es un subconjunto). Ultimamente está en el centro de atención debido a que su gran flexibilidad permite una fácil integración con bases de datos y el intercambio de grandes masas de datos en Internet. En realidad, XML no compite con PDF. Mientras que XML describe el contenido de un documento, PDF describe su apariencia. No es fácil extraer el contenido de un fichero PDF (o al menos, no sin gran cantidad de trabajo manual) ya que las estructura de un documento se pierde cuando se transforma en fichero PDF.

Es muy interesante que PDF 1.3 incluya un mecanismo en forma de estructura en árbol que permite contener datos tipo XML. De este modo, es teóricamente posible que un fichero PDF contenga un resumen estructurado del contenido además de una representación exacta de su apariencia. Desgraciadamente, los programas necesarios para incluir datos XML en un fichero PDF (usando marcas

PDF: *pdfmarks*) como, por ejemplo, una Xtension para Quark XPress, aun no existen. Si necesitas PDF y XML, de momento el único camino es crear dos ficheros separados a partir de un programa de diseño o de un sistema de publicación de bases de datos.

Si quieres saber más sobre PDF y XML, hay un artículo muy interesante que Gordon Kent publicó en [PlanetPDF](#).

PDF y Acrobat

Muchas personas confunden PDF (el formato de los datos) con Acrobat (el programa que la firma Adobe vende para generar, manipular y ver ficheros PDF). Esta confusión parece surgir del hecho de que casi todas las nuevas versiones de las especificaciones PDF van acompañadas de una nueva versión Acrobat. La versión 1.2 de las especificaciones PDF fueron acompañadas de Acrobat 3. Las especificaciones 1.3 de PDF, lo fueron de Acrobat 4.