

Ejercicio 6

Análisis de Documento de Texto:

DOC (Microsoft Word Document)

- Tipo de documento: Formato propietario de Microsoft Word (versiones anteriores a 2007)
- Compresión: No comprimido
- Estructura: Formato binario complejo
- Tamaño: Relativamente grande debido a la falta de compresión
- Calidad: Preserva formato completo, fuentes, imágenes y elementos multimedia
- Compatibilidad: Amplia, pero puede haber problemas entre versiones
- Características: Soporte completo para formato rico, tablas, gráficos, macros

DOCX (Office Open XML Document)

- Tipo de documento: Formato estándar abierto de Microsoft (desde 2007)
- Compresión: Sí, basado en ZIP con XML
- Estructura: Conjunto de archivos XML comprimidos
- Tamaño: Significativamente menor que DOC (30-75% de reducción)
- Calidad: Preservación completa de formato, mejor que DOC
- Compatibilidad: Excelente, estándar ISO/IEC 29500
- Características: Formato moderno, soporte para elementos multimedia avanzados

DJVU (DéjàVu)

- Tipo de documento: Formato especializado para documentos escaneados
- Compresión: Altamente comprimido, algoritmos específicos para imágenes de documentos
- Estructura: Capas separadas para texto, imágenes y fondo
- Tamaño: Extremadamente pequeño para documentos escaneados (10-20% del PDF equivalente)
- Calidad: Excelente para documentos escaneados, preserva legibilidad con tamaños mínimos
- Compatibilidad: Limitada, requiere software específico
- Características: OCR integrado, búsqueda de texto, ideal para bibliotecas digitales

PDF (Portable Document Format)

- Tipo de documento: Formato estándar para intercambio de documentos
- Compresión: Opcional, múltiples algoritmos disponibles
- Estructura: Formato complejo que puede contener texto, imágenes, vectores
- Tamaño: Variable, depende del contenido y compresión aplicada
- Calidad: Preservación exacta del formato visual original

- Compatibilidad: Universal, estándar ISO 32000
- Características: Multiplataforma, seguridad integrada, formularios interactivos

EPUB (Electronic Publication)

- Tipo de documento: Estándar para libros electrónicos
- Compresión: Sí, basado en ZIP con XHTML/XML
- Estructura: Conjunto de archivos XHTML, CSS e imágenes
- Tamaño: Optimizado para texto, muy eficiente
- Calidad: Texto adaptable, formato responsivo
- Compatibilidad: Amplia en dispositivos de lectura
- Características: Texto refluible, soporte para multimedia, metadatos ricos

Comparación

Aspecto	DOC	DOCX	DJVU	PDF	EPUB
Compresión	No	Sí	Sí	Opcional	Sí
Tamaño típico	Grande	Mediano	Muy pequeño	Variable	Pequeño
Editable	Sí	Sí	No	Limitado	Sí
Formato fijo	No	No	Sí	Sí	No
Compatibilidad	Media	Alta	Baja	Universal	Alta
Uso principal	Documentos editables	Documentos editables	Archivos escaneados	Distribución final	Libros electrónicos

Convertio (convertio.co) es una herramienta online que permite convertir archivos entre los formatos DOC, DOCX, PDF y EPUB sin necesidad de instalar software adicional. Con una interfaz simple de arrastrar y soltar, procesa archivos de hasta 100MB de forma gratuita, manteniendo la calidad del documento original y eliminando automáticamente los archivos después de 24 horas por seguridad, lo que la convierte en la solución más práctica para conversiones rápidas entre estos formatos de texto.