

UNIVERSIDAD NACIONAL DE SAN JUAN

FACULTAD DE CIENCIAS EXACTAS FÍSICAS Y NATURALES

DEPARTAMENTO DE INFORMÁTICA

LICENCIATURA EN CIENCIAS DE LA COMPUTACIÓN

Asignatura: **“Teorías de la Información”**

**Práctico 1 Máquina**

**ALUMNOS:**

* **Juan Peña 21.151**
* **Nicolás Quiroga Santini 21.444**
* **Baltasar Ortiz Becerra 21.752**

**Equipo de Cátedra**

**Raul Klenzi**

**Manuel Ortega**

**Fabricio Amaya**

**AÑO 2025**

# 

# 

# Ejercicio 6

## Informe: Diferencias Esenciales entre Formatos de Documento de Texto

Este informe detalla las características clave de los formatos

\*.doc, \*.docx, \*.djvu, \*.pdf, \*.epub, destacando sus diferencias en tipo, compresión, tamaño y calidad.

#### **1. Formato .doc (Microsoft Word 97-2003)**

* **Tipo de Documento**: Formato binario y propietario de Microsoft.
* **Compresión**: No es un formato inherentemente comprimido, aunque puede contener objetos comprimidos como imágenes JPEG.
* **Tamaños y Calidades**: Tiende a generar archivos más grandes en comparación con su sucesor, .docx. La calidad es alta para texto e imágenes, pero su estructura binaria lo hace propenso a la corrupción.
* **Diferencia Esencial**: Es un formato "cerrado" y obsoleto, difícil de interpretar por software que no sea de Microsoft.

#### **2. Formato .docx (Office Open XML)**

* **Tipo de Documento**: Estándar abierto basado en XML, desarrollado por Microsoft.
* **Compresión**: Es un formato comprimido. Un archivo .docx es en realidad un archivo **ZIP** que contiene múltiples archivos y carpetas (principalmente XML) que definen la estructura, el texto, las imágenes y el formato.
* **Tamaños y Calidades**: Significativamente más pequeño y robusto que el .doc. La calidad es excelente y es el estándar de facto para documentos de texto editables.
* **Diferencia Esencial**: Su naturaleza de archivo ZIP y su base en XML lo hacen más eficiente, seguro y compatible entre diferentes plataformas.

#### **3. Formato .djvu (DjVu)**

* **Tipo de documento**: Formato de archivo de imagen diseñado principalmente para documentos escaneados.
* **Compresión**: Utiliza una **compresión por ondículas (wavelet)** muy avanzada. Separa el texto, las imágenes de fondo y las máscaras en capas diferentes y las comprime por separado, logrando tasas de compresión extremadamente altas.
* **Tamaños y Calidades**: Produce archivos mucho más pequeños que PDF o TIFF para documentos escaneados en color, manteniendo una alta calidad y legibilidad del texto.
* **Diferencia Esencial**: Es la mejor opción para digitalizar libros y documentos, optimizado para la compresión de imágenes con texto.

#### **4. Formato .pdf (Portable Document Format)**

* **Tipo de Documento**: Estándar abierto (ISO 32000) creado por Adobe.
* **Compresión**: Es un formato mixto. Puede contener texto, fuentes, imágenes vectoriales y de mapa de bits. Utiliza varios algoritmos de compresión (como LZW, FLATE para texto y JPEG para imágenes) internamente.
* **Tamaños y Calidades**: El tamaño varía enormemente según el contenido. Su principal ventaja es que mantiene el **diseño visual exacto** del documento en cualquier dispositivo.
* **Diferencia Esencial**: Su objetivo es la presentación final y la portabilidad, no la edición fácil. Es el estándar universal para compartir documentos.

#### **5. Formato .epub (Electronic Publication)**

* **Tipo de Documento**: Estándar abierto para libros electrónicos (e-books).
* **Compresión**: Al igual que .docx, un archivo .epub es un archivo ZIP que contiene archivos HTML, CSS, imágenes, etc.
* **Tamaños y Calidades**: Generalmente son archivos pequeños. Su característica principal es el **texto redistribuible (reflowable text)**, que se adapta al tamaño de la pantalla del dispositivo (e-reader, tableta, teléfono), mejorando la experiencia de lectura.
* **Diferencia Esencial**: Está diseñado para la lectura de contenido fluido, no para mantener un diseño de página fijo como el PDF.

### Aplicación para Conversión de Formatos

Una de las aplicaciones más potentes y versátiles para convertir entre la mayoría de estos formatos (y muchos otros) es **Calibre**.

* **Calibre**: Es un gestor de libros electrónicos de código abierto, gratuito y multiplataforma. Puede convertir de y hacia formatos como EPUB, PDF, DOCX, etc. Permite ajustar una gran cantidad de opciones durante la conversión para optimizar el resultado para diferentes dispositivos de lectura. Es la herramienta recomendada para manejar bibliotecas de documentos digitales y realizar conversiones masivas.