

Lo que ahorra empaquetar un documento

Un PDF es sobre todo diccionarios —el catálogo, el árbol de páginas, cada página, cada fuente, cada anotación— y un diccionario no es un flujo, así que cada uno se escribe entero y ninguno se comprime.

Los flujos de objetos los reúnen en flujos que se comprimen juntos, y la tabla que dice dónde está cada objeto se vuelve también un flujo. El documento que ve el lector es el mismo en los dos casos.

El documento pesado aquí tiene 24 páginas de texto, cada una con su título.

El mismo documento, pesado dos veces

Sin empaquetar	22 724 bytes
Empaquetado	18 370 bytes
Ahorrado	4 354 bytes

Empaquetar este documento ahorró el 19 por ciento. Lo que ahorra crece con el número de páginas, porque una página cuesta unos cuantos diccionarios y los diccionarios son lo que comprime el empaquetado.

Este mismo archivo está empaquetado, así que quien lo haya abierto ha leído uno de los dos documentos pesados.

El mismo documento, pesado de otras formas

Sin compresión pesa 31 089 bytes; al nivel 1, 26 750; al nivel 6, el de por defecto, 22 724; al nivel 9, 22 713.

Con un título, pesa 22 826 bytes con su diccionario de información y 25 273 sin él, cuando el título figura solo en sus metadatos; el diccionario se escribe: no.

Una simulación, que no escribe nada, cuenta 24 páginas y 55 objetos, escritos como PDF 1.7 aunque bastaría PDF 1.4, y 0 avisos.

Escrito con sus avisos notificados pesa 22 724 bytes, y 22 724 escrito igual en un flujo; avisos: 0.

Empaquetado en flujos de objetos: no; con la primera página primero: no; la versión que necesita: PDF 1.4.

Con una portada delante y un marcador hacia ella puesto primero, tiene 25 páginas, y su primer marcador dice «Portada».

Una línea de cuatrocientos caracteres medida dentro de una pasada de medida no le cuesta nada al archivo: 7 449 bytes, y 7 449 con la pasada mantenida abierta a mano, frente a 24 826 si se mide fuera de una.