

E.P.E.T. N°4

DOCENTES:

HÉCTOR BRIZUELA

BEATRIZ VIDAÑA

ESPACIO CURRICULAR: CONVERSIÓN Y REPARACIÓN DE DATOS

CURSO: 5 to DIVISIÓN: 2da

CICLO: ORIENTADO

TURNO: MAÑANA

ÁREA CURRICULAR: INFORMÁTICA

TÍTULO DE LA PROPUESTA:

¿QUÉ ES LA CONVERSIÓN DE DATOS?

Conversión de datos

Las empresas a menudo experimentan cambios debido a fusiones, adquisiciones y actualizaciones del sistema. Estas transiciones pueden crear datos inconsistentes y redundantes. La conversión de datos es un paso clave en el proceso que convierte los datos sin procesar en información significativa.

¿Qué es la conversión de datos?

La conversión de datos es el proceso de traducir los datos de un formato a otro, manteniendo su viabilidad y calidad. El proceso implica la extracción de datos desde una fuente, como una base de datos, un archivo o un servicio web, transformándolo y cargándolo al destino requerido.

La conversión de datos puede ser un proceso simple o complejo, dependiendo del número y la complejidad de los formatos de datos involucrados.

Formatos de archivos

Los datos se guardan en archivos informáticos, la forma en que están organizados dichos datos es lo que llamamos formato de archivo.

El formato de archivo está indicado, normalmente, por tres o cuatro letras que aparecen detrás del nombre del archivo y de un punto, por ejemplo, arch_ventas.doc. Hoy en día existen multitud de formatos, cada formato se suele utilizar en un entorno diferente, dependiendo del uso que vayamos a dar a nuestros datos nos interesará guardarlos en uno u otro formato. Generalmente los archivos ocupan mucho espacio, sobre todo los gráficos, sonido y vídeo, por esto la mayoría de los formatos multimedia utilizan métodos de compresión para ahorrar espacio de almacenamiento.

En general, al descomprimir el archivo se pierde algo de calidad, y cuanto mayor es el grado de compresión, mayor es la pérdida de fidelidad de los datos descomprimidos respecto del original.

Cada fabricante de software suele crear su propio formato, con la pretensión de que sea utilizado por todos los demás y pase a ser el formato estándar. Pero esto no ha ocurrido y hoy en día no existe un formato estándar para cada uno de los diferentes formatos de archivos (datos, imágenes, vídeo, sonido, etc.). Los fabricantes suelen dotar a sus aplicaciones de capacidad para trabajar con formatos diferentes al suyo, a esto es a lo que llamamos compatibilidad de un producto software respecto de un formato.

Pero esta compatibilidad no es total, de manera que hay productos que no pueden trabajar con determinados formatos. Por lo tanto conviene conocer con qué formatos pueden trabajar las aplicaciones que vayamos a utilizar y si además otras muchas personas trabajen con nuestros archivos, conviene utilizar formatos con alto grado de compatibilidad.

FORMATOS DE ARCHIVOS

Documentos

- **txt:** Posee texto plano es decir texto sin formato. Esta simplicidad permite que una gran variedad de programas pueda leer y editar ese contenido.

- **doc, docx:** Por la amplísima difusión de Microsoft Word, este formato se ha convertido en estándar con el que pueden transferirse textos con formato o sin formato, o hasta imágenes.
 - **pdf:** (portable document format). Es multiplataforma y es uno de los formatos más extendidos en Internet para el intercambio de documentos.
-

Compresores

- **7z:** es un formato de compresión de datos sin pérdida, superando a las de los populares formatos zip y rar.
- **zip:** formato de almacenamiento sin pérdida, muy utilizado para la compresión de datos como imágenes, programas o documentos.
- **rar:** formato de archivo propietario, con un algoritmo de compresión sin pérdida utilizado para la compresión de datos.

Imágenes

- **bmp:** formato propio del programa Microsoft Paint, que viene con el sistema operativo Windows. El formato BMP o mapa de bits es el que se utiliza para las imágenes tapiz (fondo del escritorio).
- **jpg, jpeg:** formato de imagen muy extendido en internet. Comprime la imagen mediante la aplicación de algoritmos informáticos.
- **gif:** formato gráfico muy extendido usado también para el uso de animaciones simples usado generalmente para la publicidad en tipo banners e imágenes animadas para páginas web. Muy usado por ser el único formato soportado por multitud de navegadores que permita dicho efecto.
- **png:** (*Portable Network Graphics*) es un formato gráfico basado en un algoritmo de compresión sin pérdida.
- **tiff:** (*Tagged Image File Format*) es un formato de archivo para imágenes con etiquetas. Esto se debe a que los archivos TIFF contienen, además de los datos de la imagen propiamente dicha, "etiquetas" en las que se archiva información sobre las características de la imagen, que sirve para su tratamiento posterior.

Web

- **html:** *HyperText Markup Language* (Lenguaje de Marcado de Hipertexto), es el lenguaje de marcado o etiquetas predominante para la elaboración de páginas web.
- **css:** Hojas de estilo en cascada (en inglés *Cascading Style Sheets*). CSS es un lenguaje usado para definir la presentación de un documento estructurado escrito en HTML o XML.
- **js:** es un archivo de texto plano que contiene scripts de *Javascript*, y que puede, ser modificado con cualquier editor de textos. Es ejecutado generalmente por un navegador web.

Audio

Sin pérdida de calidad

- **wav:** apócope de *WAVEform audio format*, es un formato de audio digital normalmente sin compresión de datos desarrollado y propiedad de Microsoft y de IBM que se utiliza para almacenar sonidos en el PC, admite archivos mono y estéreo a diversas resoluciones y velocidades de muestreo.
- **aiff:** *Audio Interchange File Format* (AIFF) es un estándar de formato de audio usado para almacenar datos de sonido en computadoras personales. Es uno de los formatos líderes ya que está comprimido sin ninguna pérdida.
- **ape:** *Monkey's Audio* (Audio del Mono) es un formato de archivo para comprimir información de audio. Siendo un formato de compresión sin pérdida, no elimina información del flujo de audio.
- **flac:** *Free Lossless Audio Codec* (FLAC) (Códec libre de compresión de audio sin pérdida, en español) es un formato para codificar audio sin pérdida de calidad, es decir, el archivo inicial puede ser recompuesto totalmente con la desventaja de que el archivo ocupe mucho más espacio del que se obtendría al aplicar compresión con pérdida o Lossy.

Con pérdida de calidad

- **mp3:** es un formato de audio digital comprimido con pérdida desarrollado por el Moving Picture Experts Group (MPEG).
- **wma:** *Windows Media Audio* o WMA es un formato de compresión de audio con pérdida y es propiedad de Microsoft.

- **aac:** (del inglés Advanced Audio Coding) es un formato informático de señal digital audio basado en el Algoritmo de compresión con pérdida, un proceso por el que se eliminan algunos de los datos de audio para poder obtener el mayor grado de compresión posible, resultando en un archivo de salida que suena lo más parecido posible al original.
 - **ra, rm:** *RealAudio* es un formato de audio creado por RealNetworks, Inc. Permite adaptarse automáticamente a la capacidad de recepción del usuario final.
-

Vídeo

- **avi:** AVI (siglas en inglés de *Audio Video Interleave*) es un formato contenedor de audio y video lanzado por Microsoft en 1992.
- **mov:** *QuickTime* es un framework multimedia estándar desarrollado por Apple que consiste en un conjunto de bibliotecas y un reproductor multimedia (*QuickTime Player*).
- **mpg:** El *Moving Picture Experts Group*, titulado Grupo de Expertos en Imágenes Móviles en español, referido comúnmente como MPEG, es un grupo de trabajo del ISO/IEC encargado de desarrollar estándares de codificación de audio y vídeo.
- **wmv:** *Windows Media Video* (WMV) es un nombre genérico que se da al conjunto de algoritmos de compresión ubicados en el set propietario de tecnologías de vídeo desarrolladas por Microsoft.
- **asf:** *Advanced Streaming Format* es un formato contenedor de audio y video digital propiedad de Microsoft, diseñado especialmente para el streaming.

Conceptos básicos

Antes de que pueda efectuarse cualquier conversión de datos el usuario o programador de aplicaciones debe tener en mente unos cuantos conceptos básicos de teoría de la información e informática:

- Es fácil descartar información usando un ordenador, pero añadirla requiere esfuerzo.

- El ordenador puede usarse para añadir información sólo sobre la base de reglas; la mayoría de adiciones que interesa a los usuarios sólo puede lograrse con la ayuda de humanos.
- Sobre muestrear los datos o convertirlos a un formato con más posibilidades no añade información: sólo hace hueco para dicha adición, que suele tener que hacer un humano.

Por ejemplo, una imagen *truecolor* puede ser convertida fácilmente a escala de grises, siendo la conversión opuesta un proceso complicado. Convertir un archivo de texto Unix a un archivo de texto Microsoft (DOS/Windows) implica añadir información, concretamente un retorno de carro (CR) antes de cada salto de línea (LF), y dicha adición puede hacerse con un ordenador gracias a que se basa en reglas.

Convertir un PNG de 24 bits a uno de 48 bits de profundidad de color no añade información a la imagen, pues sólo rellena los valores RGB de los píxeles con ceros. La conversión permite luego afinar estos valores «expandidos», pero sólo con posteriores manipulaciones manuales. Convertir una imagen o un archivo multimedia desde un formato comprimido con pérdida (como JPEG) a otro sin pérdida (como PNG) o descomprimido (como BMP o WAV) sólo desperdicia espacio, ya que la imagen o sonido resultante será el mismo, con la información original perdida (los artefactos de la compresión con pérdida). Una imagen JPEG nunca puede recuperar la calidad de la imagen original anterior a la compresión.

Debido a estas realidades, la conversión de datos es con mucha frecuencia un proceso complejo y propenso a errores, que requiere de la ayuda de expertos.

GUÍA N° 1 – Resuelva en el cuaderno

1. ¿Por qué una empresa puede tener datos inconsistentes y redundantes?
2. ¿Qué es la conversión de datos y cuál es el proceso que lleva a cabo?
3. ¿Qué es el formato de archivo?
4. ¿Qué tipo de archivos son los que generalmente se comprimen y por qué es conveniente comprimirlos?

5. ¿A qué se refiere la compatibilidad de un producto software?
6. ¿Existe la compatibilidad total entre productos de software?
7. Complete el siguiente cuadro:

Formato de archivo que posee pérdida de calidad o información	Formato de archivo que no posee pérdida de calidad o información
Mp3	7z
...	...
...	...
...	...
...	...

8. Según el formato flac ¿Qué significa codificar audio sin pérdida y con pérdida?
9. Según el formato aac ¿Qué significa compresión con pérdida?
10. Escriba en el cuadro los ejemplos de conversiones, guiándose con los ejemplos:

Se puede convertir una imagen truecolor a escala de grises
Se puede convertir un archivo de texto Unix a un archivo de texto Microsoft
...
...
...
...
...

Directora: Claudia Roldán