

Asignatura:

Teoría de la información

Tema:

Representación de la información de los datos

Trabajo para realizar:

investigar como la computadora representa cada formato
(texto,ascii,imagen,video,audios)

Alumno:

Pérez Barahona Pedro Luis

Programa educativo:

Ingeniería en datos e inteligencia organizacional

MATRICULA: 190300395

Presentado a:

Jiménez Sánchez Ismael

Fecha:

28/01/2026

1. Representación de Formatos (Codificación)

A. Texto: Del Símbolo al Bit

Las computadoras no entienden letras, solo números. Para esto usamos mapas de caracteres:

ASCII: Utiliza 7 bits (128 combinaciones). Representa el alfabeto inglés, números y caracteres de control. Es insuficiente para la diversidad global.

Unicode (UTF-8): Es el estándar actual. Asigna un "Code Point" único a cada símbolo del mundo (incluyendo emojis). UTF-8 es eficiente porque usa un número variable de bytes: 1 byte para inglés, pero hasta 4 para símbolos complejos.

B. Imágenes: Píxeles y Vectores

Mapas de Bits (Raster): La imagen se divide en una rejilla de **píxeles**. Cada píxel tiene un valor numérico que representa su color.

Modelo RGB: Cada píxel usa 3 canales (Rojo, Verde, Azul). Si usamos 8 bits por canal (24 bits en total), podemos representar más de 16 millones de colores.

Compresión: **JPG** elimina frecuencias de color que el ojo humano no nota (con pérdida), mientras que **PNG** agrupa áreas de color idéntico para reducir tamaño sin perder calidad.

Vectores (SVG): No guardan píxeles, sino ecuaciones matemáticas (ej. "dibuja un círculo de radio 5 en la coordenada X,Y"). Esto permite escalar la imagen infinitamente sin que se "pixelee".

C. Audio: Muestreo y Cuantización

El sonido es una onda analógica continua. La computadora la representa mediante:

Muestreo (Sampling): Se toma la "altura" de la onda miles de veces por segundo (ej. 44,100 Hz para calidad CD).

Cuantización: Se asigna un valor binario a cada muestra.

WAV: Guarda todas las muestras (pesado, alta fidelidad).

MP3: Usa modelos psicoacústicos para eliminar sonidos que el cerebro humano no percibe (como un tono bajo oculto tras uno fuerte).

D. Video: La Ilusión del Movimiento

Un video es una secuencia de imágenes (cuadros) más una pista de audio.

Códecs (H.264, HEVC): Son algoritmos que comprimen el video. En lugar de guardar cada cuadro completo, solo guardan los **cambios** entre un cuadro y el siguiente (redundancia temporal).

Contenedores (MP4, MKV): Son "cajas" que organizan el flujo de video, las múltiples pistas de audio y los subtítulos.

2. Los Tres Estados de la Información

Desde la perspectiva de la ciberseguridad y la ingeniería, la información se comporta distinto según donde esté:

I. Información en Reposo (Data at Rest)

Es la información almacenada de forma persistente. No se está moviendo ni procesando.

Ubicación: Discos duros (HDD/SSD), bases de datos, cintas de respaldo, almacenamiento en la nube (S3, Azure Blob).

Reto de Seguridad: El cifrado de disco (AES-256) y el control de acceso físico. Si alguien roba el disco, no debe poder leer los datos.

II. Información en Tránsito (Data in Transit)

Es la información que viaja por una red.

Ubicación: Cables de fibra óptica, ondas de Wi-Fi, routers y switches.

Reto de Seguridad: El uso de protocolos seguros como **TLS/SSL (HTTPS)**. Aquí la información es vulnerable a ataques de interceptación (*Man-in-the-Middle*). La teoría de la información aquí se enfoca en la integridad (que el mensaje no cambie por ruido o ataque).

III. Información en Uso (Data in Use)

Es la información cargada en la memoria volátil para ser procesada.

Ubicación: Memoria **RAM**, registros del **CPU** y memoria **Caché**.

Reto de Seguridad: Es el estado más difícil de proteger. Se requiere proteger el espacio de memoria para que un programa malicioso no lea los datos de otro (ej. ataques tipo Spectre o Meltdown).

3. Conexión: Teoría de la Información y Seguridad

En todos estos estados, aplicamos los conceptos vistos:

SQL Injection: Un ataque que ocurre cuando la información pasa de **Tránsito** a ser **Procesada** (en uso) sin ser validada, alterando la base de datos (en reposo).

Protocolos: Son las reglas (sintaxis) que aseguran que el Emisor y Receptor entiendan el código (binario) de la misma forma, manteniendo la semántica (significado).

