

Digital Preservation



APEX 2017

Dylan Lorenz and Manon Gray

NYU - MIAP

¿Qué son los archivos digitales?

Código Binario (1 y 0)

```
010010000100010111101001001001001110011010010001010000001110101101001001001010000100  
100100100010010100100101100100101000100100101001010010101010010101011011100010011001
```

Guardado en:

1. Magnetismo: disco giratorio, cinta magnética
2. Carga eléctrica: disco duro de unidad estable (SSD) o memoria USB
3. Pits o película: disco óptico (CD o DVD)

¿Por qué es difícil preservar los materiales digitales?

Los archivos digitales son intangibles

Los archivos digitales tienen varios niveles de dependencia

1. Hardware

2. Software

Por eso los archivos digitales exigen más control que los medios físicos.

¿Qué riesgos corren los archivos digitales?

1. Daño físico al equipo
2. Pérdida de datos por compresión
3. Daño de bit o corrupción / error digital
4. Obsolescencia
 - a. Hardware
 - b. Software
 - c. Formato de archivos
5. Metadatos insuficientes
6. Falta de fondos

Los riesgos se pueden mitigar



Riesgo: Pérdida de Datos por Compresión

Una vez que se ha comprimido un archivo, no se pueden recuperar los datos perdidos.

Soluciones:

1. Tener archivos diferentes para preservación y acceso

1. Elegir formatos de archivo sin compresión cuando sea posible para los másters de preservación

- Usar formatos de archivo comprimido para acceso en la Web

Riesgo: Daño físico del Hardware

Los discos duros, ordenadores y otros dispositivos de almacenaje se desgastan o sufren daño de repente.

Soluciones:

- Tener copias suplementarias (en equipo diferente)
- Observar principios de dispersión geográfica

Riesgo: Error digital

Los datos se pueden dañar cuando se hace transferencia, por error humano o por cambios de los bits en los archivos.

Soluciones:

Hacer sumas de verificación (“checksums”) para asegurar la integridad de los archivos

- Un fórmula matemática analiza un archivo y genera un número único que describe un archivo. Se puede usar la secuencia de números para confirmar si un archivo ha cambiado.

Observar principios de dispersión geográfica

- Las copias múltiples disminuyen el riesgo de la pérdida irreparable

Riesgo: La Obsolescencia

Detenimiento de producción del hardware o de la asistencia para un software

Soluciones para el hardware:

- migración
- mantener el equipo viejo

Soluciones para software:

- utilizar software de código abierto
- Emulación

Soluciones para formatos de archivos:

Evaluar la sostenibilidad de los formatos archivos

adopción

- propiedad
- documentación

Riesgo: Metadatos insuficientes

Si no hay información sobre un archivo, no se sabe lo que se tiene.

Soluciones:

- implementar normas para nombrar archivos
- mantener inventarios con los identificadores únicos y ubicación de los archivos
- mantener metadatos técnicos y descriptivos

Riesgo: Falta de fondos

La preservación digital requiere de fondos consistentes para la migración de archivos y almacenamiento.

Soluciones:

- No hay una solución fácil, pero se puede empezar por educar a otros sobre la importancia de la preservación digital
- Establecer prioridades de preservación

Demo

MediaInfo (<https://mediaarea.net/es/MediaInfo/Download>)

- Interfaz Gráfica del Usuario (IGU)
- Ventana Terminal

Suma de verificación

- md5