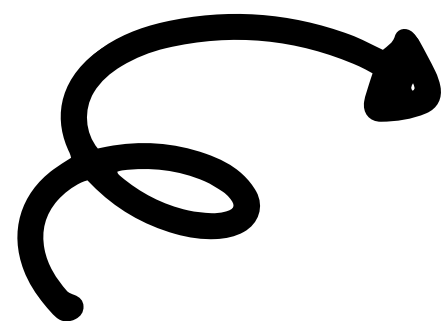




# ELIMINACION DE ARCHIVOS DUPLICADOS

Proyecto Realizado por Juan Pineda y Cisse Santos

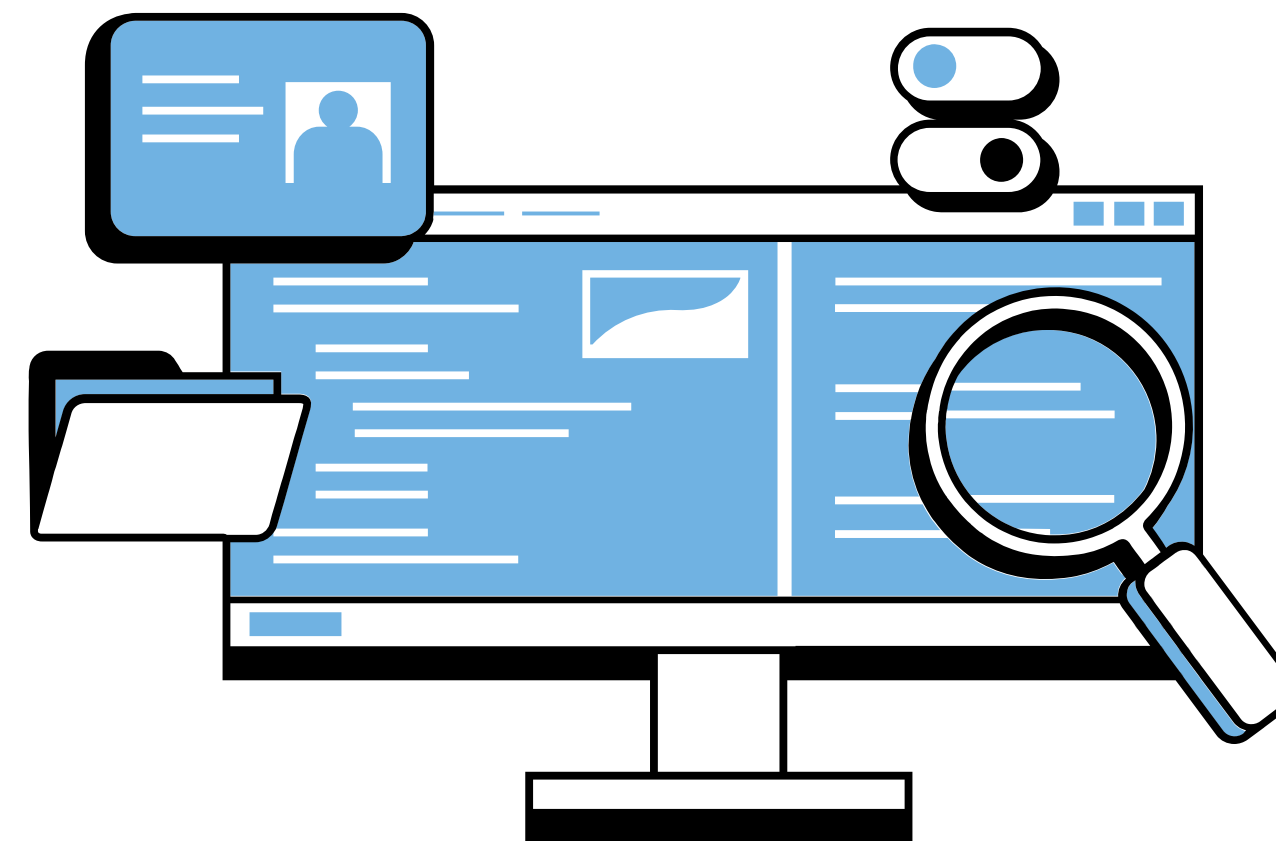


# INTRODUCCIÓN



Los usuarios almacenan grandes cantidades de archivos y, sin darse cuenta, generan duplicados.

- Estos duplicados ocupan espacio innecesario.
- Dificultan la organización de documentos.
- Pueden generar confusión entre versiones.
- Reducen el rendimiento del dispositivo.



# ARCHIVOS DUPLICADOS

En los dispositivos de los usuarios es común acumular archivos duplicados sin darse cuenta. Este problema ocurre por descargas repetidas, copias de seguridad, duplicación de trabajos o nombres similares.

Actualmente, los sistemas operativos no ofrecen una herramienta que detecte duplicados por contenido real, sino solo por nombre o ubicación.

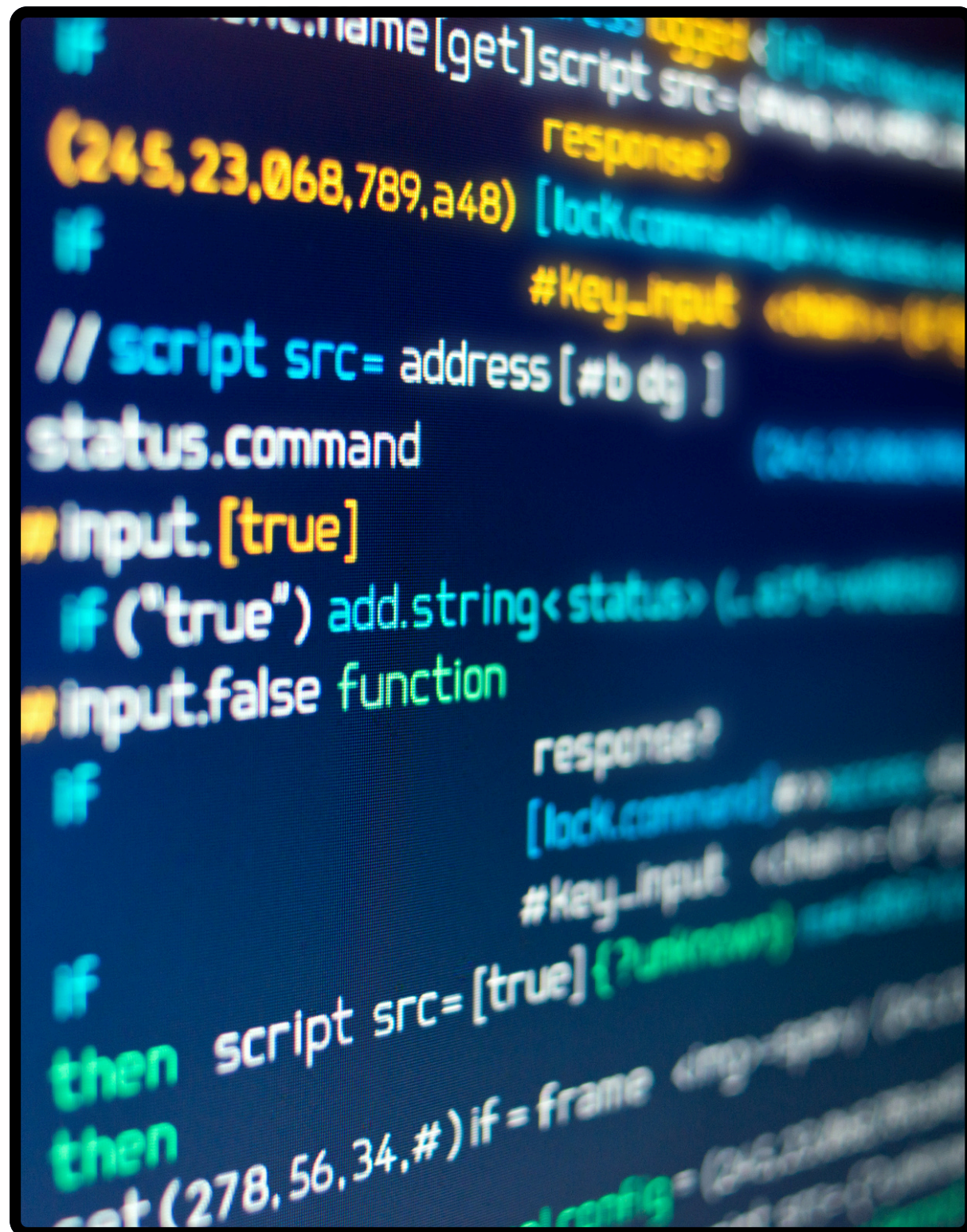


# SOLUCION

La solución es un programa en Python que analiza todas las carpetas seleccionadas, genera una huella digital llamada hash para cada archivo y detecta cuáles tienen el mismo contenido.

El sistema conserva automáticamente el archivo más reciente y elimina las copias antiguas, evitando riesgos y asegurando que el usuario mantenga la versión correcta.





# ¿QUÉ ES UN HASH?

Un hash es un código único que se genera a partir del contenido de un archivo, similar a una huella digital.

El programa utiliza la librería hashlib para comparar archivos:

Si dos archivos producen el mismo hash, significa que su contenido es exactamente igual, sin importar su nombre o ubicación.

# ¿CÓMO FUNCIONA EL SISTEMA?

El sistema sigue un flujo sencillo y preciso:

- Escanea todos los archivos en la carpeta seleccionada.
- Calcula el hash de cada archivo.
- Agrupa los que tienen el mismo hash.
- Identifica cuál es el archivo más reciente.
- Elimina automáticamente las copias antiguas.



# RESULTADOS Y

## BENEFICIOS

Durante las pruebas realizadas, el programa identificó correctamente todos los archivos duplicados presentes en la carpeta de prueba.

Conservó siempre la versión más reciente y eliminó únicamente las copias antiguas.

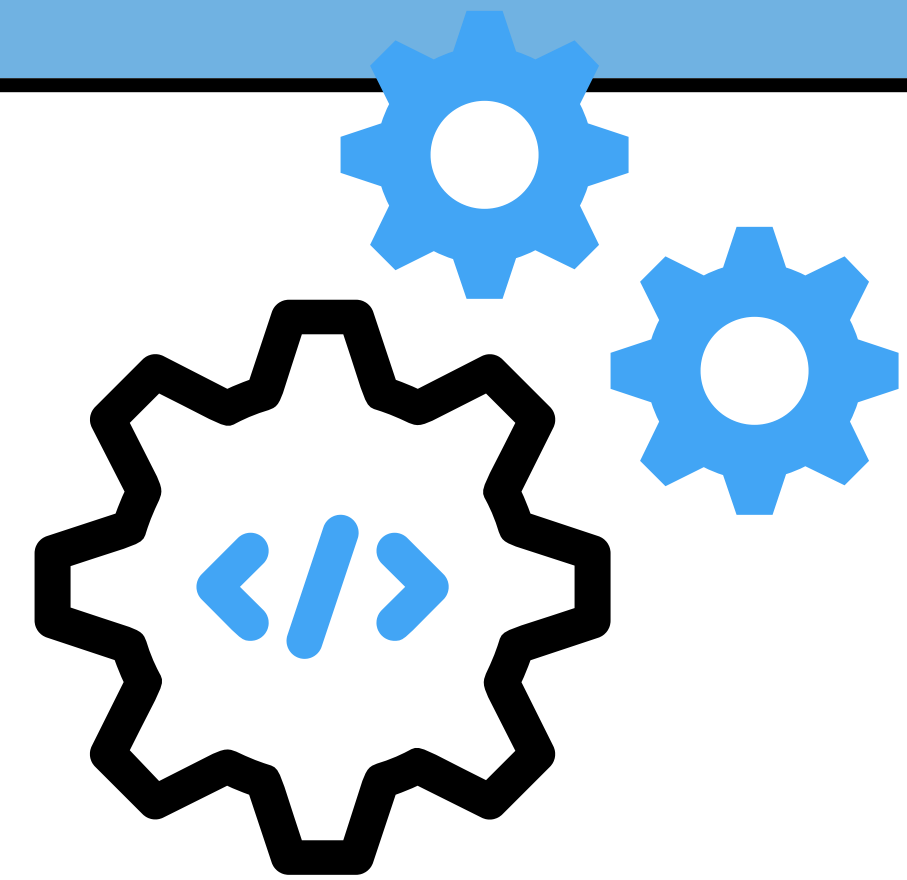
Esto permitió liberar espacio de almacenamiento y mejorar la organización del sistema.

Beneficios clave:

- Optimización del espacio.
- Reducción del desorden digital.
- Automatización de un proceso que normalmente es manual y lento.
- Mayor control sobre versiones de archivos.



# CONCLUSIÓN

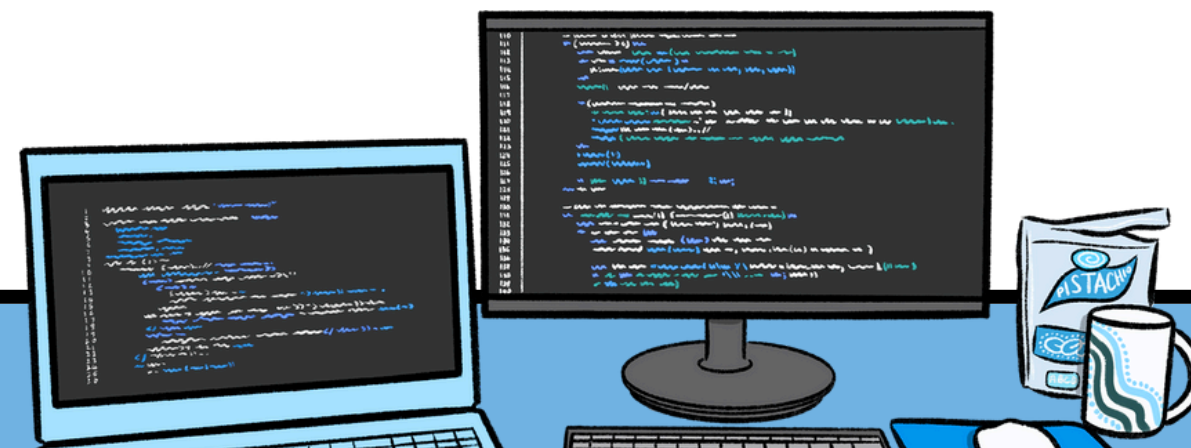


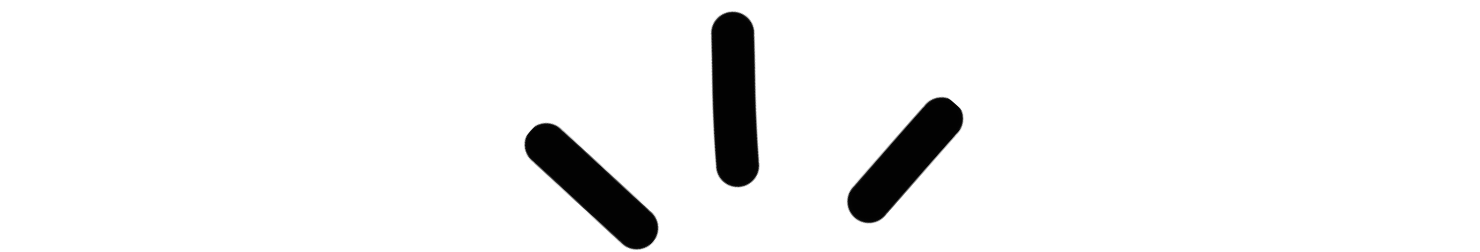
El proyecto demuestra cómo Python puede aplicarse para resolver problemas reales de forma eficiente.

Esta herramienta automatiza la eliminación de archivos duplicados, mejora la organización del usuario y libera espacio de almacenamiento sin riesgos.

## MEJORAS FUTURAS:

- Interfaz gráfica para facilitar el uso.
- Reportes automáticos de archivos eliminados.
- Opciones de filtro por tipo de archivo.
- Historial de acciones para mayor seguridad.
- Opción de elegir los archivos a eliminar





**MUCHAS**

**GRACIAS**

