Persistencia.

Se llama persistencia a la capacidad de guardar la información de un programa para poder volver a utilizarla en otro momento. Es lo que los usuarios conocen como Guardar el archivo y después Abrir el archivo. Pero para un programador puede significar más cosas y suele involucrar un proceso de serialización de los datos a un archivo o a una base de datos o a algún otro medio similar, y el proceso inverso de recuperar los datos a partir de la información serializada.

Por ejemplo, supongamos que en el desarrollo de un juego se quiere guardar en un archivo la información referente a los ganadores, el puntaje máximo obtenido y el tiempo de juego en el que obtuvieron ese puntaje.

En el juego, esa información podría estar almacenada en una lista de tuplas:

[(nombre1, puntaje1, tiempo1), (nombre2, puntaje2, tiempo2), ...]

Esta información se puede guardar en un archivo de muchas formas distintas. En este caso, para facilitar la lectura del archivo de puntajes para los humanos, se decide guardarlos en un archivo de texto, donde cada tupla ocupará una línea y los valores de las tuplas estarán separados por comas.

La persistencia de datos es la existencia residual de datos digitales que permanecen incluso después de que se ha intentado eliminarlos o borrarlos. La persistencia de datos provoca que los datos queden intactos incluso después de una operación de eliminación de archivos, del reformateo de medios de almacenamiento que no eliminan los datos previamente escritos, o a través de propiedades físicas de los medios de almacenamiento que permiten recuperar los datos, con el riesgo de la divulgación inadvertida de información sensible.

Qué es la persistencia de datos

La persistencia de datos se refiere a los datos residuales que quedan de un archivo luego de que se han hecho intentos para eliminarlo o borrarlo. Estos residuos pueden ser causados por el proceso de eliminación llevado por el sistema operativo que deja los datos intactos, por el formateo de medios de almacenamiento que no elimina los datos previamente existentes en el medios, o debido a las características físicas del medio de almacenamiento que permiten recuperar datos previamente escritos.

La persistencia de datos puede permitir la recuperación de la información y la puede dejar expuesta si no se hace un correcto manejo del dispositivo y se libera en un entorno no controlado, por ejemplo si se pierde, se arroja a la basura o se deja en manos de terceros.

Muchos sistemas operativos, gestores de archivos y otro software proporcionan métodos de borrado en los que un archivo no se elimina inmediatamente cuando el usuario solicita esa acción. En su lugar, el archivo se mueve a un área de almacenamiento temporal, denominado usualmente “Papelera de reciclaje”, lo que facilita al usuario deshacer un error.

Del mismo modo, muchos productos de software crean automáticamente copias de seguridad de los archivos que se están editando, para permitir al usuario restaurar la versión original o recuperarse de un posible fallo, característica denominada usualmente “Autosave” (“Autoguardado”).

Incluso cuando no se proporciona un recurso de retención temporal de archivos borrados o cuando el usuario no lo utiliza, los sistemas operativos no eliminan realmente el contenido de un archivo cuando se elimina, a menos que se utilicen conscientemente herramientas de borrado seguro. En su lugar, simplemente eliminan la entrada del archivo del directorio del sistema de archivos porque es más rápido y requiere menos trabajo, y el contenido del archivo permanece en el medio de almacenamiento hasta que el sistema operativo reutilice el espacio para guardar nuevos datos.

En algunos sistemas, por diversos motivos también se dejan atrás suficientes metadatos del sistema de archivos como para permitir una fácil recuperación de archivos mediante utilidades de software. Incluso cuando la recuperación es imposible, los datos, hasta que se han sobrescrito, pueden ser leídos por un software que lea los sectores del disco directamente. La informática forense emplea a menudo tal software.

Del mismo modo, reformatear o reparticionar un sistema es poco probable que sobreescriba todas las áreas del disco, aunque parezca que el disco está vacío. Lo mismo ocurre en caso de efectuar una copia imagen de otro soporte de datos, aunque el medio aparezca vacío excepto los archivos presentes en la imagen, muchos datos no habrán sido sobrescritos.

Además, incluso cuando el medio de almacenamiento se sobrescribe, sus propiedades físicas pueden permitir la recuperación del contenido anterior. Sin embargo, en la mayoría de los casos esta recuperación no es posible simplemente leyendo el dispositivo de almacenamiento con herramientas convencionales, sino que se requiere el uso de técnicas de laboratorio, tales como desmontar el dispositivo y acceder/leer directamente sus componentes.

Se han desarrollado diversas técnicas para contrarrestar la persistencia de datos y minimizar los riesgos de exposición de la información, tales como sobrescribir, desmagnetizar, cifrar y destruir los medios.