Prevención de Fugas de Datos (Data Leak) en empaquetados

Cuadernillo | Vol. 1

mercedev.es — 2026-04-26 | Fase 10 (Empaquetado y Release)

El Desafío (Síntoma)

Al ejecutar el script de instanciación (`merci-init.py`) para extraer la infraestructura del proyecto hacia un nuevo repositorio público (Merci Boilerplate), se detectó que los archivos PDF generados localmente viajaron junto con la clonación. Esto constituyó una fuga de datos (Data Leak), exponiendo la propiedad intelectual de la autora en un repositorio que debía ser una plantilla en blanco.

La Maniobra (Lógica)

Se parcheó inmediatamente el script de Python encargado de la limpieza destructiva, inyectando una regla explícita para purgar el directorio de artefactos binarios: `purge_directory(REPO_ROOT / "public" / "descargas")`

Adicionalmente, se procedió a eliminar manualmente los binarios residuales en el repositorio de destino y se registró el incidente en la bitácora de seguridad.

El Aprendizaje / Deuda Técnica

Crear una plantilla reutilizable a partir de un proyecto vivo requiere un control de aislamiento exhaustivo. El código fuente es solo una parte del ecosistema; los archivos generados dinámicamente (SSG - Static Site Generation - Generación de Sitios Estáticos) o los binarios (PDFs, imágenes procesadas) arrastran el estado y la identidad del proyecto original.

En DevSecOps, los scripts de empaquetado deben desconfiar por defecto y declarar de forma explícita listas de purga (Allow-list / Deny-list) para garantizar la sanitización total del entorno saliente.