

DEEPFAKE

Posted on 11. febrero 2026 by Alexander

Un **deepfake (ultrafalso)** es un archivo de imagen, audio o vídeo generado o manipulado mediante IA que representa a personas reales de forma engañosamente realista o simula declaraciones.

¿Qué significa deepfake?

El término se compone de "deep learning" (una forma de aprendizaje automático) y "fake" (falso). Los deepfakes se crean con ayuda de redes neuronales que analizan grandes volúmenes de datos y generan contenidos multimedia de apariencia realista.

Formas típicas:

- Videos manipulados con el rostro sustituido
- Voces generadas artificialmente
- Imágenes realistas de personas que no existen

La calidad de los deepfakes modernos suele ser tan alta que las manipulaciones son difíciles de detectar.

¿Cómo funciona un deepfake?

Los deepfakes suelen basarse en **Generative Adversarial Networks (GANs, redes generativas antagónicas)** u otros modelos de IA similares. Se combinan dos sistemas:

1. Un generador crea contenidos artificiales.
2. Un discriminador evalúa lo realistas que parecen esos contenidos.

Mediante entrenamiento repetido, la calidad mejora hasta que los contenidos generados apenas se distinguen de grabaciones reales.

Ejemplo práctico

En un vídeo manipulado, parece que un individuo conocido hace una declaración que nunca ha realizado. En realidad, el rostro se ha sustituido digitalmente o la voz se ha generado de forma sintética.