

# Suplantación de identidad mediante IA y deepfakes

El papel de la IA en el ataque y en la defensa

# **Ataques de presentación**

La inteligencia artificial puede tener un papel importante tanto en la preparación y ejecución de ataques de presentación como en su prevención.

A continuación, veremos algunos ejemplos.

# **Spoofing mediante imagen o vídeo**

## **Mecánica:**

Se muestra al sensor una foto, un vídeo o una pantalla reproduciendo el rostro de la víctima.

## **Rol de la IA en el ataque:**

1. Mejora de la nitidez y el realismo.
2. Corrección de iluminación.
3. Estabilización del parpadeo.
4. Generación de micro-movimientos para simular “pruebas de vida”.

## **Rol de la IA en la defensa:**

1. Detección de artefactos digitales, patrones de pantalla, reflejos no humanos.
2. Detección de modelos de “liveness” basados en micro-expresiones y análisis de profundidad.

# Deepfakes de rostro y voz

## **Mecánica:**

Generación sintética de cara o voz para suplantar a alguien.

## **Rol de la IA en el ataque:**

1. Generación de deepfakes hiperrealistas (face-swap, voice cloning).
2. Sincronización labial (“lip-sync”) en tiempo real.

## **Rol de la IA en la defensa:**

1. Detectores de deepfake (análisis de inconsistencias biológicas, parpadeo, compresión).
2. Modelos que verifican concordancia entre audio, vídeo y contexto.



# Caso Arup

## Ataque mediante deepfake de IA

### Estafa de 25 millones de dólares



|                       |                                                                                                                                                                                                                                                     |
|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Año                   | 2024                                                                                                                                                                                                                                                |
| Víctima               | Filial en Hong Kong de Arup, empresa global de ingeniería con sede en Londres                                                                                                                                                                       |
| Tipo de ataque        | Suplantación de directivos mediante ataque de presentación con deepfake                                                                                                                                                                             |
| Mecánica del ataque   | <div>1. Videollamada con deepfake en tiempo real que suplantaba al CFO.</div> <div>2. En la videollamada participaban otros deepfake de directivos.</div> <div>3. El empleado creyó estar viendo al director financiero y autorizó los pagos.</div> |
| Datos biométricos     | Sí, obtenidos a partir de las fotos y vídeos del director financiero y otros.                                                                                                                                                                       |
| Uso de sistema de IA  | Sí, para crear un deepfake en tiempo real a partir de los datos biométricos.                                                                                                                                                                        |
| Resultado             | 15 transferencias que sumaron 25,6 millones de dólares                                                                                                                                                                                              |
| Origen de los datos   | Fotos y vídeos de los directivos publicados en LinkedIn y en YouTube.                                                                                                                                                                               |
| Dificultad del ataque | Muy baja: Herramientas de bajo coste + investigación en LinkedIn.                                                                                                                                                                                   |

# Máscaras 3D y prótesis realistas

## **Mecánica:**

Uso de máscaras de silicona o de impresión 3D para engañar a sistemas de reconocimiento facial.

## **Rol de la IA en el ataque:**

1. Diseño asistido por IA.
2. Aproximación de volúmenes, texturas y tonos de piel.

## **Rol de la IA en la defensa:**

1. Detección de profundidad, termografía, micro-texturas de piel.
2. Modelos que aprenden indicadores de materiales no humanos.

# **Simulación de huellas dactilares mediante moldes, pegatinas o impresión**

## **Mecánica:**

Presentación de moldes de silicona, impresiones 2D/3D sobre el lector..

## **Rol de la IA en el ataque:**

1. Reconstrucción de patrones a partir de imágenes y muestras.
2. Generación de huellas sintéticas creíbles.

## **Rol de la IA en la defensa:**

1. Detección de indicadores de vida (sudor, elasticidad, presión).
2. Análisis de poros y micro-arrugas mediante modelos entrenados.

# Iris y retina

## **Mecánica:**

Presentación de lentillas impresas, imágenes del iris, proyecciones..

## **Rol de la IA en el ataque:**

1. Reconstrucción de patrones a partir de imágenes.
2. Generación de patrones de iris sintéticos con alta similitud estadística.

## **Rol de la IA en la defensa:**

1. Medición de indicadores de vida en el iris a través de la contracción pupilar y los reflejos especulares, entre otros.
2. Modelos que distinguen entre impresiones y estructuras orgánicas..



# ISO 30107

La norma ISO/IEC 30107 es una serie internacional que establece el marco para la **detección de ataques de presentación (PAD)** en sistemas de autenticación biométrica, definiendo cómo probar la "detección de vida" para prevenir el fraude con máscaras, fotos o vídeos, dividiéndose en partes que cubren el marco (Parte 1), los formatos de datos (Parte 2) y las pruebas y reportes de rendimiento (Parte 3). Su objetivo es asegurar que los sistemas biométricos sean fiables frente a intentos de engaño, como el uso de réplicas o imágenes falsas presentadas al sensor.

Esta norma es muy importante para la seguridad en los sistemas biométricos, ya que permite a las organizaciones validar que sus lectores de huellas o reconocimiento facial pueden distinguir a un usuario real de un impostor que utiliza una falsificación a través de un ataque de presentación.

# **Formación obligatoria en materia de IA**

Si eres usuario de un sistema de IA puedes realizar la formación obligatoria del artículo 4 del Reglamento de IA en nuestro campus:

<https://www.campus-ribas.com/p/ia-formacion-obligatoria>

# Datos de contacto

|                     |                                                                                                 |
|---------------------|-------------------------------------------------------------------------------------------------|
| Nombre del despacho | Ribas                                                                                           |
| Domicilio           | Diagonal 640 1C - 08017 Barcelona                                                               |
| Persona de contacto | Xavier Ribas                                                                                    |
| Correo electrónico  | <a href="mailto:xavier.ribas@ribastic.com">xavier.ribas@ribastic.com</a>                        |
| Teléfono fijo       | 934940748                                                                                       |
| Teléfono móvil      | 639108413                                                                                       |
| LinkedIn            | <a href="https://www.linkedin.com/in/javierribas/">https://www.linkedin.com/in/javierribas/</a> |
| Web                 | <a href="http://ribas.legal">http://ribas.legal</a>                                             |
| Blog                | <a href="http://xribas.com">http://xribas.com</a>                                               |