



Deepfakes y Manipulación Digital

Riesgos, detección e implicaciones éticas.

Raúl Miranda Romero

Objetivos de la Conferencia



Comprender Deepfakes

Entender qué son y cómo funcionan estas tecnologías.



Aprender Detección

Conocer los criterios y herramientas para identificarlos.



Analizar Riesgos

Identificar los peligros sociales y democráticos asociados.



Reflexionar sobre Ética

Debatir las implicaciones éticas y legales de su uso.



Introducción a la Era de la Información Sintética

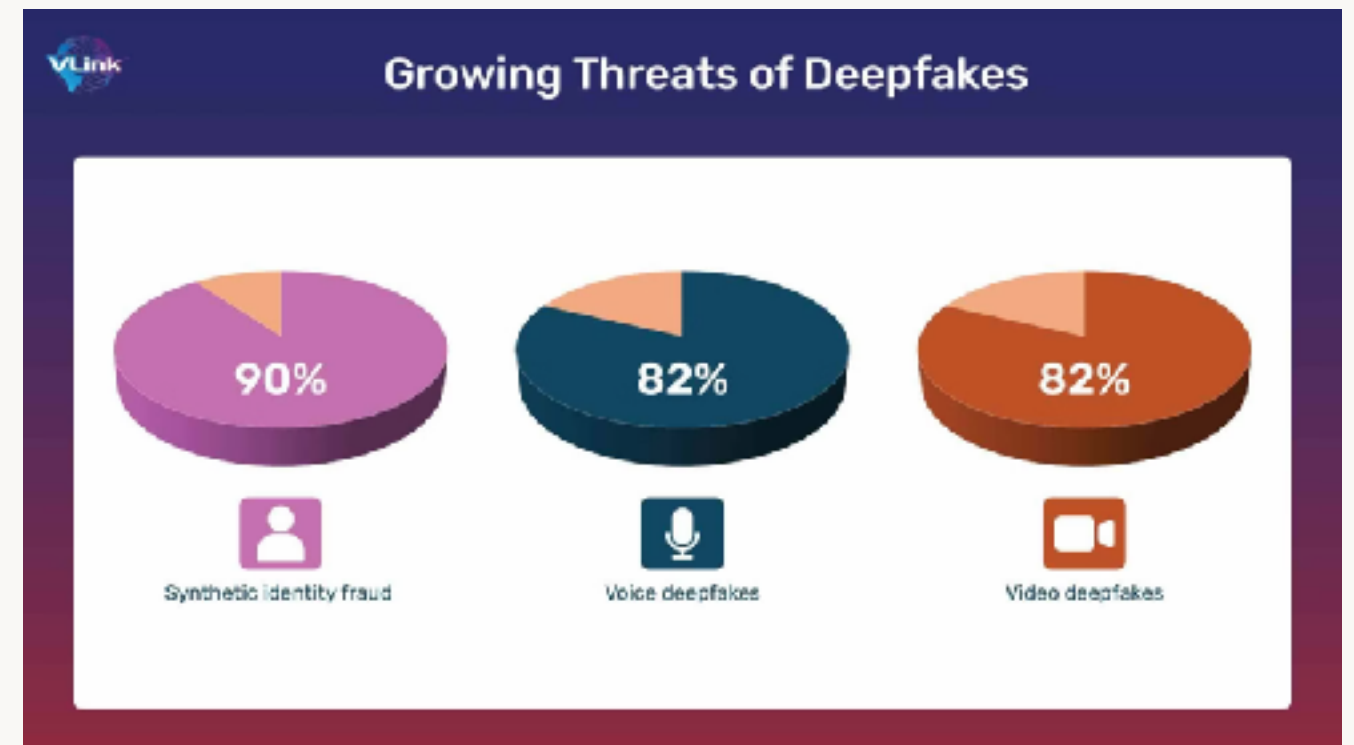
Vivimos en un momento de **crisis de confianza digital**, donde la información sintética se ha vuelto omnipresente. Los deepfakes representan un desafío significativo, con un **impacto social** profundo que exige nuestra atención.

Esta conferencia abordará la naturaleza de los deepfakes, sus mecanismos de creación y las graves consecuencias que pueden tener en nuestra sociedad.

¿Qué son los Deepfakes?

Definición: Los deepfakes son contenidos multimedia (imágenes, videos, audios) generados o modificados con inteligencia artificial para representar eventos o personas que nunca ocurrieron o dijeron algo que no dijeron.

Diferencia con edición tradicional: A diferencia de la edición convencional, que manipula elementos existentes, los deepfakes crean contenido completamente nuevo y convincente, haciendo que la detección sea mucho más compleja.

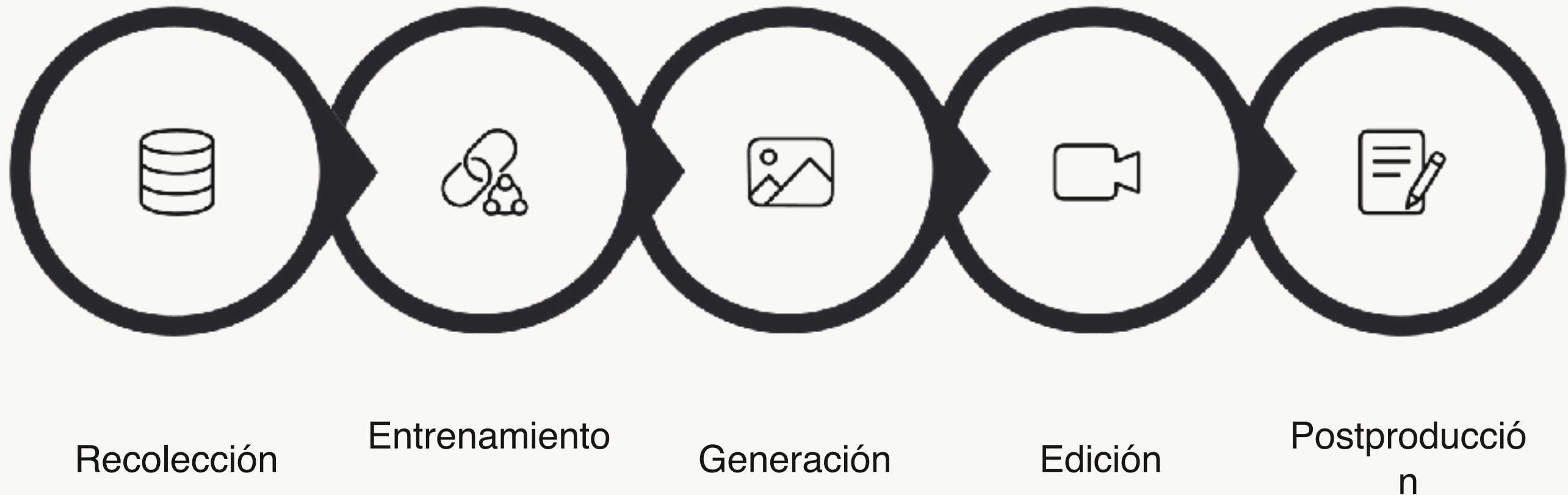


Cómo Funcionan: La Magia de la IA



Estos modelos de IA son el corazón de la creación de deepfakes, permitiendo la generación de contenido sintético cada vez más sofisticado y difícil de distinguir de la realidad.

El Proceso de Creación de un Deepfake



La creación de un deepfake es un proceso meticuloso que comienza con la recolección de vastas cantidades de datos, seguido por un entrenamiento intensivo de modelos de IA, la generación del contenido sintético y, finalmente, una edición para perfeccionar el resultado.

Tipos de Manipulación Digital

Manipulación de Imagen

- **Face swap:** Intercambio de rostros entre personas.
- **Edición semántica:** Modificación de expresiones o características faciales.
- **Restauración:** Mejora o alteración de imágenes existentes.



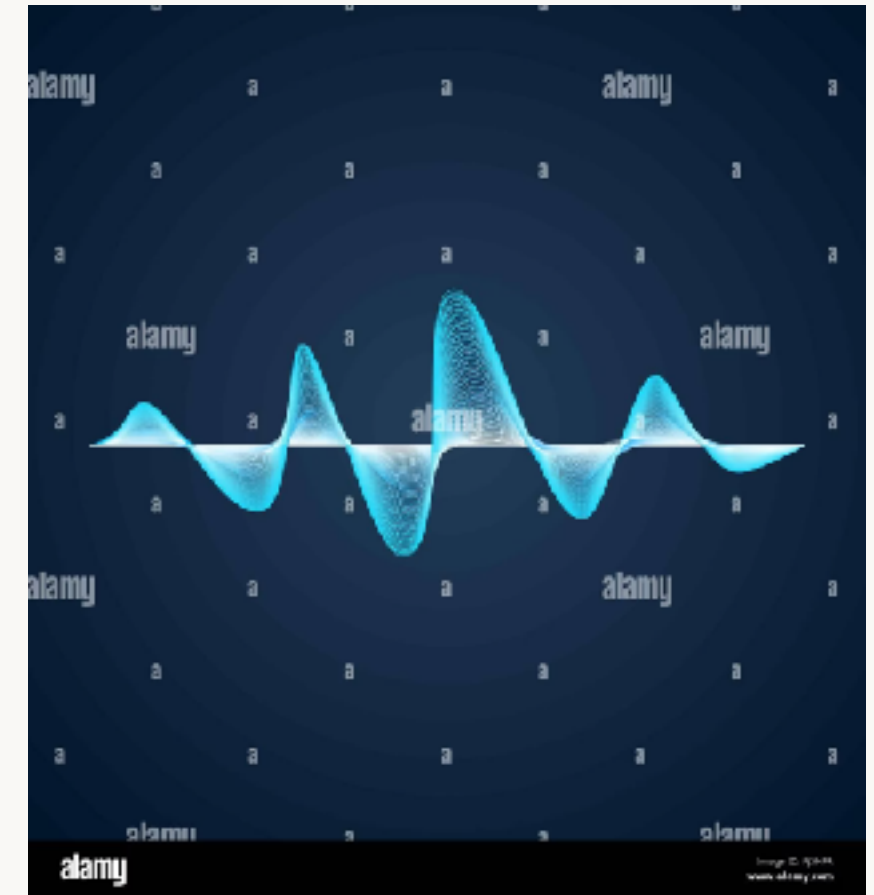
Manipulación de Video

- **Lip sync:** Sincronización de labios con audio falso.
- **Reanimación:** Animación de fotos estáticas o videos antiguos.
- **Avatares:** Creación de personajes digitales realistas.



Manipulación de Audio

- **Clonación de voz:** Replicación de la voz de una persona.
- **Text-to-Speech:** Generación de voz a partir de texto.
- **Voice conversion:** Transformación de una voz a otra.

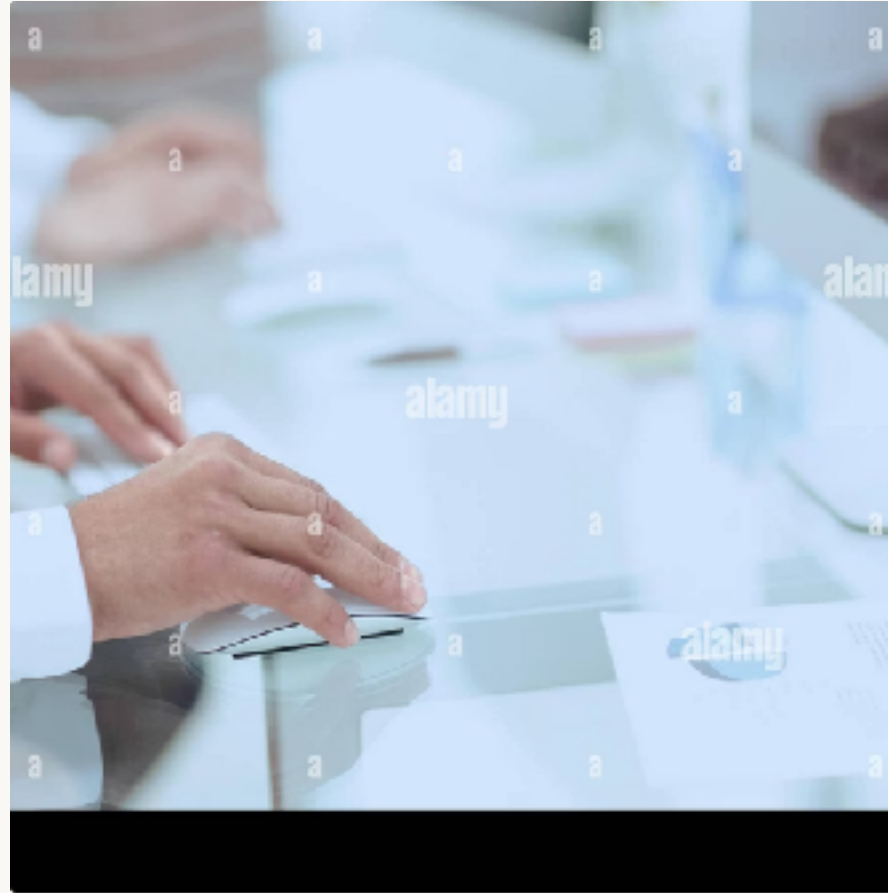


Riesgos y Casos de Deepfakes



Riesgos Políticos

Desinformación, manipulación electoral y suplantación de figuras públicas pueden socavar la democracia.



Riesgos Institucionales

Fraude, daño reputacional y phishing avanzado amenazan la seguridad empresarial y financiera.



Riesgos Sociales

Polarización, desconfianza generalizada y el efecto cínico erosionan la cohesión social.

Los deepfakes ya han sido utilizados en casos políticos (audios falsos), empresariales (fraude por voz) y sociales (difamación, pornografía no consentida).

Detección y Prevención

Indicadores Clave

- **Visuales:** Parpadeo anormal, iluminación inconsistente, bordes pixelados.
- **Audio:** Tono artificial, pausas irregulares, falta de emoción.
- **Contextual:** Evaluar la fuente, el momento y la intención del contenido.

Herramientas de Detección

- Forense digital avanzado.
- Búsqueda inversa de imágenes y videos.
- Marcas de agua digitales y metadatos.



Prevención

- **Institucional:** Protocolos, capacitación y unidades especializadas.
- **Ciudadana:** Pensamiento crítico, verificación de fuentes y alfabetización digital.



Conclusiones y Mensaje Final

Reto Tecnológico y Social

Los deepfakes son un desafío complejo que requiere soluciones multifacéticas.

Necesidad de Ética

El desarrollo y uso de la IA deben guiarse por principios éticos sólidos.

Corresponsabilidad

Gobiernos, empresas y ciudadanos deben colaborar en la lucha contra la desinformación.

Mensaje Final: No todo lo que vemos es real. La verdad se verifica.

Preguntas y Diálogo

Abramos un espacio para la reflexión colectiva y el diálogo sobre este tema crucial.