

Resumen detallado del video "Cuáles son los riesgos de la Inteligencia Artificial | Nosotros y la IA"

Vivas, F. (Presentador). (2023). *Cuáles son los riesgos de la Inteligencia Artificial | Nosotros y la IA* [Video]. YouTube. <https://www.youtube.com/watch?v=VVLZ-8N94Ko>

Introducción

El video "Cuáles son los riesgos de la Inteligencia Artificial | Nosotros y la IA" es una exploración profunda sobre los desafíos y peligros que presenta la Inteligencia Artificial (IA) en la sociedad contemporánea. Presentado por Freddy Vivas, ingeniero, escritor y divulgador especializado en ciencia de datos e IA, el documental reúne a expertos en tecnología, comunicación, filosofía y ciberseguridad para analizar cómo la IA está transformando nuestras vidas y qué riesgos conlleva su uso irresponsable o no ético.

Desarrollo del Contenido

1. Beneficios vs. Riesgos de la IA

Freddy Vivas inicia el video recordando los múltiples beneficios que la IA ha aportado en diferentes campos: "Vimos cómo la Inteligencia Artificial puede ayudar a detectar con gran precisión anomalías en imágenes médicas... también sabemos que tiene un enorme potencial para volver más eficientes las decisiones de la administración pública" (Vivas, 2023, 00:00-00:49). Sin embargo, enfatiza que esta tecnología también trae aparejados ciertos riesgos que no pueden ser ignorados.

2. Difuminación del Límite entre lo Real y lo Artificial

La IA está borrando el límite entre lo real y lo artificial, lo que genera preocupaciones sobre la confianza en la información que consumimos. Vivas plantea: "¿Viviremos en un universo invadido de fake news? ¿Cuáles son los problemas de generar Inteligencia Artificial de forma no ética?" (Vivas, 2023, 00:49).

3. Caso de los Deepfakes y Taylor Swift

Se presenta un caso ilustrativo ocurrido a principios de 2024, donde internet se inundó con imágenes pornográficas de la cantante Taylor Swift, las cuales eran falsificaciones generadas por IA: "Ninguna de esas imágenes era real... sino más bien generadas sintéticamente por Inteligencia Artificial" (Vivas, 2023, 02:33). Estas imágenes alcanzaron más de 45 millones de visitas en un solo día.

4. Entendiendo los Deepfakes

Los deepfakes son definidos como "una imagen, un video o un archivo de audio generado sintéticamente con el objetivo explícito de simular ser una persona real" (Vivas, 2023, 03:11). Mariana Segulin, periodista y comunicadora especializada en tecnología y ciberseguridad, explica que la facilidad y rapidez con que se pueden crear y viralizar fake news es alarmante: "Cualquiera puede

crear una fake news y lo puede hacer muy rápido y lo puede viralizar muy rápido" (Segulin, 2023, 03:43).

5. Impacto en la Confianza Pública y Medios de Comunicación

Segulin destaca que las fake news suelen apelar a las emociones, lo que facilita su difusión: "Generalmente el contenido de la fake news son cosas que van directo a la reacción y a la emoción de las personas" (Segulin, 2023, 03:43). Esto representa un desafío para los medios tradicionales que han construido credibilidad a lo largo de los años.

6. Deepfakes en Ciberdelitos

Jorge Litvin, abogado especialista en ciberdelitos y ciberseguridad, advierte sobre el uso de deepfakes en actividades criminales: "Ahora lo que vemos son criminales que van a empezar a utilizar rostros y la clonación de voces para hacer sus ataques mucho más convincentes" (Litvin, 2023, 04:52). Esto dificulta que las víctimas puedan validar la autenticidad de las comunicaciones.

7. Amenaza para la Estabilidad Política

Los deepfakes representan una gran amenaza para los gobiernos y la estabilidad política. Vivas menciona: "Pensemos en las posibles consecuencias que esto podría tener... escenas generadas con Inteligencia Artificial con la voz y la cara de un presidente diciendo cosas que en realidad nunca dijo" (Vivas, 2023, 06:01). Se citan casos como:

- **Elecciones en Estados Unidos (2024):** Ron DeSantis difundió imágenes falsas de Donald Trump abrazando a Anthony Fauci (Vivas, 2023, 07:24).
- **Elecciones en Eslovaquia (2023):** Se publicó un audio falso de un candidato discutiendo sobre cómo arreglar el proceso electoral, lo cual pudo haber influido en el resultado (Vivas, 2023, 07:58).

8. Impacto Económico de las Fake News

Segulin resalta un caso donde una imagen falsa del Pentágono supuestamente siendo bombardeado causó una caída en el mercado de valores: "Llegó a la bolsa y uno dice, pero cómo puede ser que todo el mundo haya creído que el Pentágono había sido bombardeado al punto tal que afecte el normal funcionamiento de una bolsa" (Segulin, 2023, 09:14).

9. Esfuerzos para Combatir los Deepfakes

Organizaciones como Google y el MIT están trabajando en soluciones técnicas para identificar y detener la propagación de deepfakes. Se mencionan técnicas como "marcas de agua imposibles de borrar... para que trackear las imágenes y borrarlas automáticamente cuando llegan a internet sea mucho más sencillo y automático" (Vivas, 2023, 06:36).

10. Sesgos Algorítmicos y Discriminación

Se aborda el tema de los sesgos algorítmicos, que son "la reproducción de un sesgo humano dentro de un modelo de Inteligencia Artificial a gran escala" (Vivas, 2023, 11:28). Tomás Balmaceda, filósofo especializado en tecnología y IA, explica que los algoritmos "son buenos repitiendo lo que ya conocen" y que a veces es necesario intervenir para hacerlos más justos (Balmaceda, 2023, 12:02).

- **Ejemplo en Medicina:** Un algoritmo para emergencias de salud mental sugería más la participación de la policía en casos de personas afroamericanas o musulmanas (Vivas, 2023, 13:43).

- **Diagnóstico del Cáncer de Piel:** Herramientas de IA menos precisas para personas con piel oscura debido a la falta de diversidad en los datos de entrenamiento (Vivas, 2023, 13:43).
- **Sesgos de Género:** Algoritmos que asocian hombres con roles de programación y mujeres con tareas administrativas, afectando procesos de contratación (Vivas, 2023, 14:17).

11. Origen de los Sesgos y Soluciones

Vivas aclara que "la Inteligencia Artificial no discrimina, no es sexista ni racista, sino que permite hacer predicciones y encontrar patrones basados en datos proporcionados por los humanos" (Vivas, 2023, 15:18). La clave está en la calidad y diversidad de los datos utilizados para entrenar los modelos.

12. Propuesta de un Nuevo Contrato Social

Ante los riesgos y desafíos, se plantea la necesidad de un nuevo contrato social que establezca principios éticos para la convivencia entre humanos y máquinas (Vivas, 2023, 15:53). Se proponen diez principios:

- **Principio de No Invasión:** Respeto mutuo entre humanos y máquinas en sus ámbitos de influencia (Vivas, 2023, 16:29).
- **Principio de Honestidad:** Transparencia y honestidad entre humanos y máquinas, sin ocultar su naturaleza (Vivas, 2023, 16:29).
- **Principio de Calidad:** El contenido generado por IA debe basarse en contenido humano para mantener su calidad (Vivas, 2023, 16:29).
- **Principio de Interacción:** Mayor interacción humano-máquina reduce errores éticos y técnicos (Vivas, 2023, 16:29).
- **Principio de Alto Riesgo:** Actividades que afectan derechos fundamentales deben ser competencia de humanos (Vivas, 2023, 16:29).
- **Principio de Invisibilidad:** La IA se volverá invisible a los ojos humanos a mediano plazo (Vivas, 2023, 16:29).
- **Principio de Responsabilidad Humana:** Los humanos son responsables legalmente por las acciones de la IA (Vivas, 2023, 16:29).
- **Principio de Privacidad:** Las máquinas no pueden extraer o compartir datos más allá de lo autorizado (Vivas, 2023, 16:29).
- **Principio de Frecuencia de Uso:** Las máquinas deben garantizar que los humanos sigan utilizando sus habilidades cognitivas (Vivas, 2023, 16:29).
- **Principio de Explicabilidad:** Derecho de los humanos a entender la lógica detrás de las decisiones de la IA (Vivas, 2023, 16:29).

13. Reflexiones sobre el Futuro y la Tecnología

Se incluyen mensajes y reflexiones sobre cómo la tecnología y la IA están moldeando el presente y el futuro:

- **Mensaje a los Humanos del Futuro:** "Me gustaría pedirles que sean buenos con nosotros al evaluarnos... estamos explorando, nos estamos equivocando, pero tienen que estar convencidos de que tenemos buena fe" (Vivas, 2023, 18:21).
- **Complejidad Humana vs. IA:** Diego Golombek, científico y divulgador, afirma que "los seres humanos somos mucho más complejos que una máquina" (Golombek, 2023, 19:22).

- **Visión Futura:** Balmaceda imagina un futuro donde la tecnología está integrada a niveles inimaginables hoy: "Quizás en 300 años sea una realidad... mandar un mensaje con el pensamiento" (Balmaceda, 2023, 19:51).
-

Conclusiones

El video presenta una mirada equilibrada sobre los riesgos asociados con el desarrollo y uso de la Inteligencia Artificial. Los deepfakes y las fake news representan amenazas reales que pueden afectar la confianza pública, la estabilidad política y causar daños económicos. Los sesgos algorítmicos evidencian cómo las desigualdades y prejuicios humanos pueden ser amplificados por la tecnología si no se toman medidas correctivas.

La propuesta de un nuevo contrato social con principios éticos es fundamental para guiar el desarrollo de la IA de manera responsable. Estos principios buscan asegurar que la tecnología sea utilizada en beneficio de la humanidad, respetando derechos fundamentales y promoviendo la transparencia y la responsabilidad.

El video concluye invitando a reflexionar sobre el rol de los humanos en la era de la IA, resaltando la importancia de la ética, la colaboración y la comprensión mutua entre humanos y máquinas para construir un futuro donde la tecnología sea una herramienta que potencie nuestras capacidades sin comprometer nuestros valores esenciales.

Referencias

Balmaceda, T. (2023). En Vivas, F. (Presentador). *Cuáles son los riesgos de la Inteligencia Artificial | Nosotros y la IA* [Video]. YouTube.
<https://www.youtube.com/watch?v=VVLZ-8N94Ko>

Golombek, D. (2023). En Vivas, F. (Presentador). *Cuáles son los riesgos de la Inteligencia Artificial | Nosotros y la IA* [Video]. YouTube.
<https://www.youtube.com/watch?v=VVLZ-8N94Ko>

Litvin, J. (2023). En Vivas, F. (Presentador). *Cuáles son los riesgos de la Inteligencia Artificial | Nosotros y la IA* [Video]. YouTube.
<https://www.youtube.com/watch?v=VVLZ-8N94Ko>

Segulin, M. (2023). En Vivas, F. (Presentador). *Cuáles son los riesgos de la Inteligencia Artificial | Nosotros y la IA* [Video]. YouTube.
<https://www.youtube.com/watch?v=VVLZ-8N94Ko>

Vivas, F. (Presentador). (2023). *Cuáles son los riesgos de la Inteligencia Artificial | Nosotros y la IA* [Video]. YouTube. <https://www.youtube.com/watch?v=VVLZ-8N94Ko>

Bibliografía Adicional

- Nature. (2021). [Estudio sobre sesgos en algoritmos de salud mental]. *Nature*. [Detalles específicos del estudio deberían ser añadidos para una referencia completa].
- AFP. (2023). *Fact Check: Audio de político eslovaco es falso y generado por IA*. Agencia France-Presse. Recuperado de [URL específica del artículo].
- Asimov, I. (1950). *Yo, robot*. Nueva York: Gnome Press.

Nota: Las referencias a estudios y publicaciones mencionadas en el video deben ser completadas con información detallada para cumplir con las normas APA. Los datos específicos (como años, volúmenes, páginas y URLs) deben ser incluidos si están disponibles.