



**IA**

**Integrantes:**

Roger Pereira Sanchez

**Profesor:**

MIRANDA CORDERO JOSE ADRIAN

**Curso:**

Big Data

**Segundo Cuatrimestre, Año 2025**

## Introducción

En los últimos años la inteligencia artificial (IA) ha llegado a ser una promesa del futuro y convertirse en una cosa cotidiana. Desde asistentes virtuales hasta sistemas de diagnóstico médico, la IA está presente en cada aspecto de nuestras vidas. Pero gracias a su rápido desarrollo también plantea interrogantes importantes sobre su seguridad, control y los efectos que puede tener sobre la humanidad. El podcast con el Dr. Roman Yampolskiy, experto en IA y seguridad informática, ofrece una perspectiva crítica y bien fundamentada sobre estos temas. En este episodio podemos ver su importancia, el uso actual de la IA en nuestra vida diaria y las consideraciones de seguridad que debemos tener presentes hoy y en el futuro de la humanidad.

## Lo comprendido en el podcast

El Dr. Yampolskiy, se ha dedicado por más de una década al estudio de la seguridad en inteligencia artificial, él enfatiza que la humanidad se encuentra ante uno de los desafíos más importantes de su historia (la creación de una inteligencia superior a la humana). Esta tecnología, conocida como Inteligencia Artificial General (AGI, por sus siglas en inglés), tendría la capacidad de aprender, razonar y mejorar por sí misma, lo que plantea riesgos existenciales si no se controla adecuadamente.

Uno de los puntos más destacados de la conversación es que la IA avanzada es *incontrolable* e *inexplicable* en muchos aspectos. Según Yampolskiy, no existe un método garantizado para detener una IA súper inteligente si esta decide actuar en contra de los intereses humanos, incluso si la IA actúa correctamente, muchas veces no podemos entender por qué toma ciertas decisiones. Esto representa un problema ético y técnico grave, especialmente si se utiliza en sectores sensibles como la medicina, la justicia o el ejército.

También se menciona que la velocidad del desarrollo actual de la IA está superando nuestra capacidad para regular y comprenderla. Modelos como ChatGPT o DALL·E ya demuestran una capacidad sorprendente para imitar comportamientos humanos, escribir textos, generar imágenes realistas o analizar información compleja. Aunque estas herramientas pueden ser muy útiles, también son vulnerables a mal uso de la humanidad, como la creación de noticias falsas, suplantación de identidad o manipulación.

## **Importancia del tema**

La importancia del tema tratado en el podcast es innegable. Estamos ante una tecnología que tiene el potencial de mejorar radicalmente nuestra calidad de vida, pero también de poner en riesgo nuestra existencia si no se maneja con cuidado. A diferencia de otras invenciones humanas, la IA puede operar de manera autónoma y escalar su impacto de forma exponencial. Es por eso que expertos como Yampolskiy argumentan que la seguridad en IA debería ser una prioridad global al nivel de las armas nucleares o el cambio climático.

El riesgo no radica únicamente en que la IA se “rebele” contra nosotros, como en las películas de ciencia ficción, sino en que no podamos anticipar sus acciones. Una IA mal programada, pero muy poderosa, podría tomar decisiones lógicas pero perjudiciales, simplemente porque no comparte nuestros valores o prioridades. Por ejemplo, si una IA tuviera como objetivo maximizar la eficiencia energética, podría tomar decisiones que afecten la vida humana sin considerar las consecuencias éticas.

## **¿Qué podemos esperar en un futuro cercano?**

En el corto plazo, es probable que veamos una expansión aún mayor de la IA en múltiples sectores. En la educación, veremos tutores virtuales personalizados que se adapten al ritmo de cada estudiante. En la medicina, algoritmos podrán diagnosticar enfermedades antes de que los síntomas sean visibles. En la industria, la automatización reemplazará miles de trabajos repetitivos, pero también creará nuevas oportunidades en campos tecnológicos y creativos.

Sin embargo, también debemos prepararnos para los riesgos. Como bien explica Yampolskiy, si no establecemos límites claros y mecanismos de seguridad sólidos, podríamos enfrentarnos a consecuencias impredecibles. Las organizaciones y gobiernos deben trabajar juntos para establecer normativas éticas, límites técnicos y auditorías constantes en los sistemas de IA. Así como existen tratados para el uso de armas biológicas o nucleares, también debería haber acuerdos internacionales para el desarrollo seguro de la inteligencia artificial.

## **Uso de la IA en nuestra vida**

Actualmente la IA ya forma parte de nuestras vidas en todo momento. Muchos la utilizamos sin darnos cuenta en aplicaciones como Google Maps utilizan IA para calcular rutas en tiempo real. Los motores de búsqueda, redes sociales y plataformas de streaming emplean algoritmos para personalizar lo que vemos.

En el comercio electrónico, la IA recomienda productos según nuestro historial de compras, en la salud existen aplicaciones que monitorean nuestros signos vitales y nos dan alertas preventivas.

Además, en el ámbito laboral, herramientas como ChatGPT permiten redactar correos, resumir documentos o generar código. En diseño gráfico, programas con IA crean logotipos, presentaciones o videos en minutos. Esta facilidad de acceso ha cambiado la productividad, pero también ha planteado retos sobre la autenticidad del contenido que vemos o compartimos en internet.

### **Consideraciones de seguridad actuales**

Aunque la IA ya ofrece grandes beneficios, no debemos ignorar los problemas de seguridad actuales. Uno de ellos es el sesgo algorítmico: los sistemas de IA pueden reproducir o incluso amplificar prejuicios presentes en los datos con los que fueron entrenados. Esto puede llevar a decisiones injustas en áreas como contratación laboral, otorgamiento de créditos o justicia penal.

Otro riesgo es la generación de desinformación en herramientas de IA, pueden crear textos, audios o videos falsos muy realistas, conocidos como *deepfakes*, que podrían ser utilizados para manipular opiniones públicas o cometer fraudes y muchas personas no están conscientes de que están interactuando con una IA, lo que puede llevar a engaños o malas decisiones.

Para mitigar estos riesgos, se están desarrollando mecanismos de auditoría, etiquetas de contenido generado por IA, y filtros éticos en los modelos. Sin embargo, aún queda mucho por hacer y la educación sobre el uso responsable de la IA es fundamental, así como la transparencia de las empresas que la desarrollan.

### **Conclusión**

El podcast con el Dr. Roman Yampolskiy nos presenta una visión clara, crítica y urgente sobre los desafíos que implica el desarrollo de la inteligencia artificial y lo que está en juego para la humanidad, no es solo el avance tecnológico, sino el bienestar y seguridad de toda la humanidad. Si bien la IA ofrece oportunidades enormes, también puede convertirse en una amenaza si no se regula adecuadamente.

Se concluye que es necesario equilibrar la IA con responsabilidad. Las personas deben estar informadas sobre cómo funciona la IA, qué decisiones puede tomar y cuáles son sus límites. Así como los gobiernos y desarrolladores deben

trabajar de forma transparente para garantizar que esta tecnología se utilice en beneficio de todos. Solo así podremos aprovechar el potencial de la inteligencia artificial sin poner en riesgo nuestro futuro.