

Dos chilenos participan en informe mundial que identificó los principales riesgos de la inteligencia artificial

Considerado un hito en la comprensión científica de los riesgos asociados a sistemas avanzados de IA, el primer Informe Internacional de Seguridad en Inteligencia Artificial acaba de ser publicado. Se presentará como referencia clave en el próximo AI Action Summit que se llevará a cabo en París, el 10 y 11 de febrero.

Patricio Lazcano

El AI Safety Report es un informe internacional independiente sobre seguridad de inteligencia artificial que busca convertirse en el primer entendimiento científico integral y compartido de los sistemas avanzados de IA y sus riesgos. Se trata del reporte más esperado a nivel mundial para tomar las medidas de seguridad frente a la inteligencia artificial.

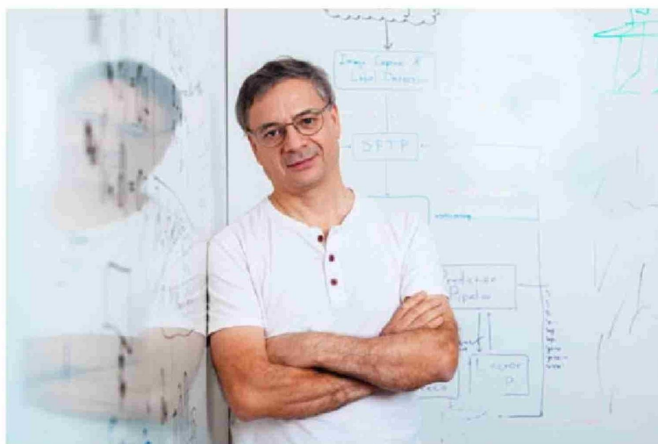
Liderado por Yoshua Bengio, un académico de IA ganador del premio Turing y el científico informático más citado del mundo, el informe fue publicado este martes y en él se reúnen las opiniones de 100 expertos internacionales independientes, entre ellos dos chilenos: Raquel Pezoa y Álvaro Soto, una muestra de los pioneros progresos que ha hecho Chile en esta disciplina.

Redacción del informe

En 2023, Chile fue uno de los pocos países latinoamericanos invitados a participar del AI Safety Summit de Londres. Como consecuencia, nuestro país fue uno de los tres de la región que participó en la redacción de este informe. Gracias a la coordinación por el Ministerio de Ciencia que lidera la ministra Aisén Etcheverry, la profesora Raquel Pezoa, de la Universidad Técnica Federico Santa María (UTFSM), y el profesor Álvaro Soto, de la Universidad Católica de Chile y director del Centro Nacional de IA, CENIA, participaron en el proceso de redacción.

Chile se ha consolidado como referente en esta materia dentro de América Latina, gracias a iniciativas como la publicación de la Política Nacional de IA, y el proyecto de ley de Inteligencia Artificial actualmente en discusión en el Congreso, que busca establecer un marco para el desarrollo seguro y ético de esta tecnología en el país.

“Esto no debiera sorprendernos. Como Gobierno hemos participado activamente de todas las instancias internacionales sobre IA durante dos años. Y el último Índice Latinoamericano de IA, que nos pone en el primer lugar del continente”, afirmó la minis-



► Álvaro Soto, uno de los redactores del informe. Foto: Juan Francisco Lizama.

tra de Ciencia, Aisén Etcheverry.

Sobre el reporte de inteligencia artificial

Escrito por expertos de 30 países, se espera que el AI Safety Report sea una guía para los responsables de las políticas públicas en todo el mundo. Fue encargado en noviembre de 2023 y se inspiró en el trabajo del Panel Intergubernamental sobre Cambio Climático de las Naciones Unidas, para consolidar la experiencia internacional más experta que existe en el mundo sobre la IA y sus riesgos.

El texto identifica tres categorías de riesgos:

1. Riesgos de uso malintencionado: estos incluyen ciberataques, la creación de material de abuso sexual infantil generado por IA e incluso el desarrollo de armas biológicas.

- Fallas del sistema: estos incluyen sesgos, problemas de confiabilidad y la posible pérdida de control sobre sistemas avanzados de IA.

- Riesgos sistémicos: estos se derivan de la adopción generalizada de IA e incluyen interrupciones en la fuerza laboral, preocu-

paciones sobre la privacidad e impactos ambientales.

El informe hace especial hincapié en la urgencia de aumentar la transparencia y la comprensión en la toma de decisiones en materia de IA, a medida que los sistemas se vuelven más sofisticados y la tecnología continúa desarrollándose a un ritmo rápido.

También destaca que, si bien las capacidades de la IA pueden avanzar a distintas velocidades, su desarrollo y sus riesgos potenciales no son una conclusión inevitable. Los resultados dependen de las decisiones que tomen las sociedades y los gobiernos hoy y en el futuro.

De temas generales a temas específicos

Álvaro Soto, director del Centro Nacional de Inteligencia Artificial, CENIA, aseguró que pocos documentos tienen la profundidad científica del AI Safety Report. “

Dijo que al leerlo se ve que tanto en temas generales como temas específicos, por ejemplo, en reconocimiento visual, multimodal o de los modelos de lenguaje, aparecen las referencias clave y eso permite a cualquiera profundizar y saber lo último que se ha

investigado sobre el tema.

“De los muchos informes que he leído creo que por su profundidad, completitud, diversidad y timing -porque incluye mucho de lo que está sucediendo ahora- este es un gran aporte”.

Explicó que el participó en dos capítulos técnicos del estado del arte de la IA en términos de sus principales técnicas y cómo están impactando tanto en el caso de los modelos de lenguaje y los modelos multimodales.

Los autores principales hicieron un primer draft y sobre él tuvimos varias iteraciones para indicar ‘acá falta esto y esto de frentón esto no es así’. Pudimos aportar significativamente porque ese primer draft cambió bastante”.

Colaboración entre países

Raquel Pezoa, académica de la UTFSM, destacó que este reporte fue liderado por Yoshua Bengio, uno de los principales expertos en IA a nivel mundial. “El objetivo es informar sobre los posibles riesgos de esta tecnología, con un objetivo claro: que quienes hacen las leyes puedan estar informados y tomar decisiones de manera colaborativa entre los distintos países.

Explico que sus tareas incluyeron asesorar el alcance, la estructura y los temas por abordar del reporte, además de sugerir revisores de ONG y de la industria, de manera de generar un texto claro y crítico.

La embajadora de Reino Unido en Chile, Louise de Sousa, cuyo gobierno coordinó el reporte independiente luego del Safety Summit de Londres, mencionó: “La publicación es un ejemplo claro de cómo la colaboración internacional puede generar soluciones y guías de trabajo concretas a los desafíos más urgentes de la inteligencia artificial. Chile ha demostrado liderazgo en esta materia, particularmente el Ministerio de Ciencia, con quienes colaboramos continuamente. Estoy orgullosa de ver su contribución en un esfuerzo global tan significativo.” ●