

La advertencia de Jacob Coxon y el desafío de las universidades ante la inteligencia artificial

Por

Ricardo Neira N. Decano de Ingeniería, Arquitectura y Tecnologías para la Sociedad, Universidad Academia de Humanismo Cristiano.

Publicado en Poder y Liderazgo

Una renuncia puede ser apenas un movimiento laboral. La de Jacob Coxon no lo fue. Después de trabajar durante aproximadamente tres años en investigación de preentrenamiento en OpenAI y Anthropic, dos de las organizaciones que lideran la carrera mundial de la inteligencia artificial.

Coxon anunció su salida de Anthropic con una acusación estremecedora: las grandes compañías estarían avanzando hacia sistemas cada vez más poderosos sin garantías suficientes de que podrán mantenerlos bajo control.

La precisión importa. Coxon no renunció simultáneamente a OpenAI y Anthropic. Trabajó primero en OpenAI y posteriormente en Anthropic, empresa de la cual salió en septiembre de 2026. Lo verdaderamente relevante, sin embargo, no es su itinerario profesional, sino el lugar desde donde formula su advertencia: el corazón de los laboratorios que desarrollan los modelos de inteligencia artificial más avanzados del mundo.

En una entrevista con WIRED, el investigador sostuvo que entre algunos especialistas existiría la percepción de que los próximos años constituirán un momento decisivo para la humanidad. No presentó una certeza científica ni habló en

nombre de toda la comunidad tecnológica. Expresó una evaluación personal, basada en su experiencia y en las conversaciones que mantuvo dentro de la industria. Aun así, resulta difícil ignorar una alarma activada por alguien que participó directamente en la construcción de estas tecnologías.

La discusión no debería reducirse a decidir si Coxon es un visionario o un alarmista. Esa es una falsa comodidad. La pregunta verdaderamente importante es otra: ¿qué estamos haciendo, como sociedad, frente a una tecnología que está adquiriendo capacidades a una velocidad muy superior a nuestra capacidad institucional para comprenderla y gobernarla?

El problema no pertenece al futuro

Cuando se habla de los peligros de la inteligencia artificial, el debate suele quedar atrapado entre dos extremos. De un lado, quienes anuncian una futura superinteligencia imposible de controlar. Del otro, quienes consideran esas advertencias una exageración cercana a la ciencia ficción.

Mientras ambos sectores discuten, los efectos concretos ya están aquí. Los algoritmos seleccionan candidatos a un empleo, recomiendan tratamientos, filtran información, producen imágenes, redactan documentos, orientan decisiones financieras y condicionan lo que millones de personas ven en las redes sociales.

También pueden reproducir discriminaciones, utilizar datos sin suficiente transparencia, amplificar noticias falsas, reemplazar tareas laborales y concentrar una enorme capacidad de decisión en unas pocas empresas privadas.

La inteligencia artificial puede contribuir a diagnosticar enfermedades, personalizar la educación, anticipar desastres y aumentar la productividad. Pero esa misma tecnología puede profundizar desigualdades, precarizar el empleo, debilitar la privacidad y alterar el debate democrático.

No estamos, por tanto, ante una elección entre innovación y prudencia. Necesitamos ambas. Una innovación sin control puede convertirse en irresponsabilidad; una prudencia sin conocimiento puede derivar en inmovilismo.

Incluso Anthropic reconoce la existencia de riesgos significativos y cuenta con una Política de Escalamiento Responsable, mediante la cual intenta vincular el aumento de las capacidades de sus modelos con mayores niveles de seguridad. Es un esfuerzo relevante, pero continúa siendo un sistema voluntario diseñado por la propia empresa que desarrolla la tecnología.

Cuando una innovación puede modificar las condiciones de vida de sociedades completas, la autorregulación empresarial no puede ser la única línea de defensa.

La hora de las universidades

Aquí comienza la responsabilidad histórica de la educación superior. Durante demasiado tiempo, numerosas universidades han entendido la inteligencia artificial principalmente como una herramienta de productividad: un asistente para redactar, preparar clases, responder preguntas, procesar datos o automatizar labores administrativas. Eso ya no es suficiente.

Una universidad no puede conformarse con comprar licencias, impartir talleres sobre instrucciones para modelos o celebrar cada nueva aplicación tecnológica. Debe formar profesionales capaces de comprender la inteligencia artificial, cuestionarla, auditarla y decidir dónde resulta legítimo utilizarla.

La IA no es solamente un asunto de ingenieros. Es una cuestión de poder, derechos, trabajo, salud, educación, cultura, territorio y democracia.

Por eso las universidades poseen una ventaja que no tienen las empresas tecnológicas: pueden reunir en un mismo proyecto a

ingenieros, científicos sociales, educadores, profesionales de la salud, artistas, arquitectos, filósofos y organizaciones ciudadanas.

La UNESCO ha establecido que los sistemas de IA deben respetar la dignidad humana, proteger la privacidad, evitar la discriminación, ser auditables y mantener la responsabilidad final en manos de las personas. También recomienda aplicar evaluaciones de impacto y garantizar la participación de los grupos afectados.

Estos principios no deberían permanecer en documentos internacionales. Deben traducirse en instituciones, laboratorios, programas formativos y productos concretos.

La oportunidad de la UAHC

La Universidad Academia de Humanismo Cristiano tiene una posibilidad singular. Su identidad humanista y la presencia de áreas como ciencias sociales, educación, ingeniería, arquitectura, tecnologías, artes y salud le permiten construir una respuesta transdisciplinaria difícil de reproducir por instituciones concentradas exclusivamente en la dimensión técnica.

La UAHC no necesita competir con Silicon Valley en cantidad de procesadores ni en el tamaño de sus modelos. Puede ocupar un espacio más necesario para Chile: convertirse en una universidad capaz de conectar inteligencia artificial, democracia, derechos humanos, innovación y transformación social.

Gobernar antes de implementar

Una universidad que quiera liderar esta transformación debe comenzar por gobernar responsablemente la inteligencia artificial dentro de su propia institución.

Eso significa crear un consejo universitario de gobernanza de IA, aprobar protocolos transparentes, registrar los sistemas utilizados y evaluar cualquier aplicación que pueda intervenir

en procesos sensibles como la admisión, la evaluación académica, la contratación o el bienestar estudiantil.

No se trata de detener el futuro

La advertencia de Jacob Coxon no obliga a aceptar que una catástrofe sea inevitable. Tampoco justifica rechazar la inteligencia artificial. Obliga a abandonar la ingenuidad.

Las grandes compañías poseen los recursos para construir sistemas cada vez más poderosos. Las universidades conservan, o deberían conservar, algo igualmente importante: la independencia intelectual para preguntar a quién beneficia esa tecnología, quién asume sus riesgos, qué derechos puede vulnerar y qué decisiones nunca deberían ser delegadas a una máquina.

Esta puede ser la gran contribución de la UAHC: demostrar que el humanismo no representa una resistencia al progreso, sino la condición necesaria para que el progreso merezca ese nombre.

La renuncia de Coxon desaparecerá tarde o temprano de los titulares. La pregunta que deja abierta, en cambio, permanecerá durante décadas: si algunos de quienes desarrollan la inteligencia artificial están pidiendo controles más fuertes, ¿pueden las universidades seguir observando desde la distancia?

La respuesta debe ser clara. Es tiempo de investigar, formar, auditar e innovar. Es tiempo de construir una inteligencia artificial al servicio de las personas y no una sociedad subordinada a las decisiones de los algoritmos.

Porque el futuro tecnológico no está escrito. Pero si las universidades renuncian a participar en su construcción, otros lo escribirán por nosotros.