Presentación

En las últimas semanas han circulado diversos documentos y opiniones acerca de los potenciales riesgos de que los agentes de la Inteligencia Artificial (IA), actúen sin control. Estos agentes son sistemas o programas capaces de realizar tareas de forma autónoma mediante la percepción de su entorno, la toma de decisiones y el uso de herramientas.

Aunque para algunos expertos, estas alertas son exageradas, es interesante conocer lo que al respecto han reflexionado distintos grupos de especialistas convocados por organismos internacionales. Este es el caso del *Reporte Preliminar del Panel Científico Internacional sobre la IA: Evaluación basada en la evidencia de las oportunidades y los riesgos de la inteligencia artificial**, el cual fue publicado a principios de julio de este año por la ONU. El panel está integrado por científicos y expertos independientes de todas las regiones de las Naciones Unidas.

Al principio del informe se señalan las enormes ventajas potenciales que ofrece la IA. Si ésta se despliega y aplica de manera reflexiva, la IA puede contribuir hacia la consecución de los Objetivos de Desarrollo Sostenible (ODS), así como impulsar el progreso en las ciencias de la salud y en un mayor acceso a la educación.

A su vez, la rápida evolución del desarrollo tecnológico y la amplitud de sus posibles aplicaciones representa importantes desafíos a quienes diseñan políticas. Además, el despliegue rápido y descontrolado de esas tecnología a gran escala, plantea también riesgos importantes, tales como los efectos negativos para la salud mental de los individuos, al igual que su posible uso como herramienta destructiva, y sus efectos en los sistemas sociales, económicos y ambientales, así como las dificultades ligadas al control de la tecnología.

En cuanto a las capacidades y adopción de la IA, el informe establece que recientemente se han producido rápidos avances, como el análisis de datos a gran escala y la ejecución de proyectos complejos, que se han aplicado de manera exitosa en la investigación científica, la salud, la agricultura, la accesibilidad, el trabajo intelectual y la tecnología de la información, en la cual se incluye la IA.

Así, se ha acelerado el diseño de fármacos, el desarrollo de vacunas y la investigación sobre la resistencia a los antibióticos. Adicionalmente, la adopción de la IA se ha acelerado de manera generalizada, aunque de manera desigual entre países y sectores. En este sentido, se subraya que a nivel mundial, más de mil millones de personas utilizan actualmente la IA conversacional cada semana. No obstante, el acceso a la IA y su uso varían notablemente en todo el planeta, y su adopción en el sur global va muy por detrás del norte global.

Por otro lado, el informe hace notar que si bien la transición hacia los agentes de IA está ya en desarrollo, es posible que en el futuro su adopción y sus repercusiones económicas dependan de las mejoras continuas en su capacidad de realizar tareas intelectuales con poca o ninguna supervisión humana.

En cuanto a la comprensión y gestión de los riesgos, el informe plantea que el desarrollo de la IA conlleva diversos riesgos, tales como las posibles repercusiones negativas en los derechos humanos, los sistemas sociales y el medio ambiente. Un ejemplo de estas implicaciones es el material de abuso sexual de niños generado por IA y la violencia sexual facilitada por las ultrafalsificaciones que con mucha frecuencia circulan por internet, afectando principalmente a mujeres y niños.

Asimismo, el desfase entre unas capacidades en rápida evolución y unos métodos de gestión de riesgos que puedan ser eficaces podría tener consecuencias muy graves. Sería el caso en que las habilidades técnicas avanzadas permitieran que actores privados utilizaran la IA con fines maliciosos en ámbitos como el fraude, la ingeniería social, la ciberseguridad, la desinformación, la biotecnología y la manipulación financiera. Además, los riesgos de la IA se distribuyen de manera desigual entre las poblaciones y los países, en tanto que el desarrollo de la IA y la riqueza que genera están concentrados en unos cuantos países y empresas.

El informe destaca que para aprovechar al máximo las ventajas que ofrece la IA y reducir sus riesgos es indispensable contar con una buena gobernanza. Así, la IA puede ampliar de manera muy amplia las capacidades humanas al ofrecer educación personalizada, herramientas de salud mental accesibles y tecnologías de apoyo mejoradas, pero para ello se requiere la realización de inversiones específicas y disponer de políticas que fomenten el acceso equitativo y estimulen la innovación.

Aunque las capacidades para actuar en función de las pruebas empíricas disponibles sobre los riesgos y las repercusiones de la IA ya existen, su distribución es desigual. El acceso a las herramientas de la IA no garantiza que todos se beneficien de igual manera, pues se requieren inversiones complementarias en datos, habilidades, flujos de trabajo e instituciones que permitan pasar del acceso al despliegue útil, eficaz en términos de costo y de seguridad, aunque su distribución sigue siendo desigual.

El informe concluye señalando que actualmente existen medidas concretas para subsanar las carencias que el documento menciona, pero cada una de ellas requiere de una inversión sostenida en la capacidad de los Estados miembros para configurar, evaluar y desplegar la IA. Al cierre del documento subrayan que, "las salvaguardias actuales no pueden seguir el ritmo del crecimiento de las capacidades de la IA".

Las consideraciones y reflexiones elaboradas por el grupo de expertos internacionales que contiene el *Reporte Preliminar...* acerca del presente y el futuro, así como las potencialidades y riesgos de la IA, sirven de marco referencial para el conjunto de trabajos que componen este número de *Universidades*. Los coordinadores del dossier, Francisco Tamarit y Marcelo Knobel también abonan en su muy puntual "Introducción" a esta reflexión necesaria en torno de los alcances y limitaciones del IA para el desarrollo de la educación superior en América. Latina y el Caribe.

*El documento completo puede consultarse en: https://www.un.org/files/es_Preliminary_Report

Armando Alcántara Santuario
Director