

EDITORIAL

LATAMGPT Y LA SOBERANÍA DIGITAL

El Centro Nacional de Inteligencia Artificial (Cenia) y el Gobierno presentaron esta semana la primera versión de LatamGPT, un modelo de lenguaje grande de inteligencia artificial (IA) de código abierto, entrenado con datos aportados por 70 instituciones, de 15 países de América Latina y el Caribe. El proyecto, liderado por Chile, es un logro tecnológico y de cooperación multilateral, en un ámbito inédito en la región, que busca posicionar al país como desarrollador de IA, más que como mero consumidor. En el actual entorno dominado por grandes tecnológicas globales, el desarrollo de capacidades propias para entrenar y desplegar modelos de esta naturaleza es crucial para garantizar la soberanía tecnológica y reducir la dependencia de proveedores externos.

En ese marco, LatamGPT se concibió como un modelo que, a diferencia de los sistemas comerciales cerrados que actualmente concentran el mercado, no compite en razonamiento complejo con plataformas como ChatGPT o Gemini, sino que ofrece una base sobre la cual terceros puedan construir sus aplicaciones. Para ello, el entrenamiento se realizó inicialmente en la nube de Amazon Web Services, con créditos canalizados a través del Data Observatory y las próximas versiones

se levantarán sobre infraestructura local.

Hasta ahora, solo entre el 2% y el 3% de la información con que son entrenados los modelos globales proviene de América Latina, pese a que la región representa cerca del 8% de la población mundial. La consecuencia son sesgos en idioma, historia y diversos asuntos locales, que de no abordarse tenderán a una homogenización cultural en el largo plazo.

Con el fin de reducir estas brechas, LatamGPT fue entrenado con datos de bibliotecas, archivos históricos y colecciones académicas de América Latina y el Caribe, lo que abre un espacio para el desarrollo de aplicaciones -con contexto regional- en educación, atención ciudadana y políticas públicas, entre muchas otras.

Sin embargo, el éxito de la iniciativa dependerá también de que el modelo acredite estándares exigentes de calidad y representatividad en sus datos de entrenamiento, así como mecanismos robustos de auditoría y mitigación de sesgos. Por otro lado, hasta ahora, el proyecto se ha financiado con recursos estatales y del Cenia, y su acceso permanece aconditado. Su consolidación exigirá ampliar de manera sustantiva la capacidad de cómputo local y asegurar financiamiento estable, en un contexto de restricciones presupuestarias. Será tarea del próximo Gobierno definir el alcance estratégico de esta iniciativa, en términos de inversión soberanía tecnológica, instalación de capacidades y generación de valor en IA.

Hasta ahora, solo entre el 2% y el 3% de la información con que son entrenados los modelos globales proviene de América Latina.