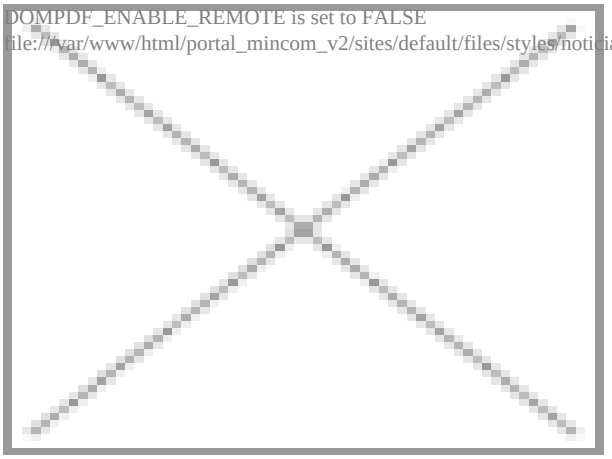
DOMPDF_ENABLE_REMOTE is set to FALSE
file:///var/www/html/portal_mincom_v2/sites/all/themes/mincom/logo.png

DOMPDF_ENABLE_REMOTE is set to FALSE
file:///var/www/html/portal_mincom_v2/sites/default/files/styles/noticias/public/img-20251007-wa0068-1_0.jpg



Fuente:
Tomado del sitio web de la Unión de Informáticos de Cuba

Con una Sesión de Inteligencia Artificial inició este martes su segunda jornada el IV Congreso Internacional Cibersociedad 2025.

Yudivián Almedia Cruz, profesor de la Facultad de Matemática y Computación de la Universidad de La Habana, presentó a CecilIA – Modelo de Lenguaje Cubano.

En sus palabras, repasó cómo han evolucionado los grandes modelos de lenguaje, como ChatGPT, y resaltó que estos están limitados por sus capacidades computacionales, a pesar de brindar la impresión de tener “memoria”.

Asimismo, tienen sesgos en función de la información con que han sido entrenado.

Hay muchas maneras de trabajar al respecto, indicó, como crearles una memoria externa, lo cual tiene sus propios desafíos, y resaltó que es necesario contar con modelos donde se atienda de manera especial la eliminación de sesgos.

Para crear un modelo de lenguaje, apuntó luego, no solo se necesitan especialista en inteligencia artificial, sino en todos los campos de la computación. También se necesita hardware especializado, que no es barato, y energía.

En resumen, crear un LLM cuesta cientos de millones de dólares por sus requerimientos.

Sin embargo, se pueden crear modelos más pequeños, que aunque tengan capacidad limitada, son más baratos. En esta línea de pensamiento, se desarrollan modelos especializados, por tanto, inquirió, ¿tendría sentido crear un modelo nacional?

Evidentemente es necesario, en tanto se estima que para 2030 el 50 por ciento de los niños emplearán modelos de lenguaje en sus interacciones digitales. Dicho comportamiento, señaló el experto, hará que aprendan a comunicarse bajo los códigos de esas IA, lo cual lleva a nuevos retos comunicacionales.

Empero, subrayó que para entrenar un modelo en lenguaje español, no existe suficiente data en ese sentido.

No obstante han impulsado la idea de crear un modelo de lenguaje cubano, y para ello se basaron en modelos previos, no iniciados desde cero, como Salamandra, empleada para crear CecilIA, que rinde homenaje a la novela homónima de Cirilo Villaverde.

El modelo actual cuenta con mil millones de tokens, a partir de unos 207 GB de datos recopilados de obras

literarias y otros sitios web como la Gaceta Oficial, sitios de canciones, discursos, etcétera.

La filosofía es trabajar con un corpus pequeño, el cual será traducido al español cubano, y el siguiente será instruir un corpus cubano propiamente.

Por otro lado, en este día también se lleva a cabo el Taller ERP: un pilar de la Transformación digital", organizado por el Grupo Empresarial de Informática y Comunicaciones (GEIC).

Bajo el lema: "De la necesidad operativa al impulso estratégico", se analizan temas relacionados con los desafíos de este tipo de tecnología y el impacto que puede tener en Cuba.

El GEIC organiza también el Taller Programa Sectorial de Ciencia, Tecnología e Innovación de la Industria del Software.