EL PINGÜINO
Fecha: 02/08/2025 Audiencia: 15.600 Sección: OPINION
Vpe: $218.607 Tirada: 5.200 Frecuencia: 0
Vpe pág: $1.035.072 Difusión: 5.200
Vpe portada: $1.035.072 Ocupación: 21,12%

La columna de...

EIRINI CHATZIKOUMI,
ACADÉMICA FACULTAD DE COMUNICACIONES Y ARTES UDLA

Copuchateando

En este momento se hablan alrededor de 7.000 lenguas en el mundo. De ellas, aproximadamente en 20 se concentra la investigación para desarrollar tecnologías de chats de inteligencia artificial (IA) generativa, como Gemini o ChatGPT. Y de esas no podemos decir que se vean suficientemente reflejados todos sus dialectos. En otras palabras, ChatGPT ha sido mayoritariamente entrenado en inglés, mientras que el español comparte el pequeño porcentaje restante con otras lenguas, y en este espacio ¿qué tan representado será el chileno? y ¿el aymara?

Es ese el vacío que viene a llenar pronto el Latam-GPT, un chat que se está desarrollando en el marco de una colaboración internacional, en la que Chile tiene un papel primordial.

¿Por qué es tan importante eso? En primer lugar, debido a que se espera que su rendimiento en español y en sus distintos dialectos sea mucho mejor que los existentes hasta ahora, ya que estará entrenado con una gran cantidad de información textual de nuestra región, aunque contempla una limitada cantidad de datos en lenguas originarias, que esperemos puedan aumentar con el tiempo. En segundo lugar, porque los usuarios y usuarias de Latinoamérica no dependerán del desarrollo y posibles limitaciones de empresas muy lejanas de nuestra realidad con fines propios.

Sucede que lo que hace un chat de IA generativa es en su base una tarea de predicción probabilística. Pongamos como ejemplo la generación de una frase que empieza con la palabra “sopaipillas”. La herramienta necesita colocar después de ella la palabra más probable; busca, entonces, en el vasto volumen de textos con los que fue entrenada y ve que la mayoría de las veces la palabra que sigue es “pasadas”, y así sucesivamente. Si bien se puede argumentar que esta es una descripción demasiado simplista, permite dejar dos cosas claras: uno, el chat no conoce si las sopaipillas son pasadas o no, simplemente “sabe” que muchas veces la gente dice “sopaipillas pasadas” (y no “sopaipillas acarameladas” o “sopaipillas florecidos”); y dos, si en los textos con los que fue entrenado nadie habla de sopaipillas, el chat tampoco hablará de sopaipillas.

Así pues, seamos usuarios conscientes y ocupemos el chat adecuado según el idioma y el dialecto que queremos usar y acorde al propósito que tenemos (teniendo claro que el chat no “sabe” nada de sopaipillas). Y para reivindicar el derecho a las sopaipillas pasadas, podemos ser expectantes en cuanto al Latam-GPT e incluso aportar a su entrenamiento, interactuando con el “Copuchat”, que es como se llama la versión disponible en este momento para que la gente contribuya a su desarrollo.