LA TERCERA
Fecha: 26/01/2025 Audiencia: 271.020 Sección: ACTUALIDAD
Vpe: $7.568.708 Tirada: 76.017 Frecuencia: DIARIO
Vpe pág: $9.829.612 Difusión: 76.017
Vpe portada: $9.829.612 Ocupación: 77%

Investigadora en inteligencia artificial y profesora de la Universidad de La Sorbonne, la académica investiga en las dimensiones emocionales de la IA y en su capacidad de influir en las decisiones humanas. Ella es partidaria de regulaciones y de educar a los niños para que no solo sean consumidores, sino innovadores en esta tecnología.

Por **Andrés Gómez Bravo**

Laurence Devillers

"Vamos a ser y ya estamos siendo manipulados por las inteligencias artificiales"

Moxie levanta los brazos y bosteza, tal como hacen los niños cuando despiertan. Puede mover los brazos, la boca y las cejas, y su rostro se ilumina cuando algo parece gustarle. Desarrollado por la empresa Embodied, Moxie es un robot con inteligencia artificial. Tiene la capacidad de interactuar y jugar con los niños. "Me encantan los cuentos. ¿Me lees un cuento?", le dice a un chico en un video promocional. "A veces, tomar la mano de un amigo me hace sentir bien. ¿Quieres darme la mano?", pregunta. Y extiende su brazo sin dedos. El niño lo toca y un sonido de alegría sale del pecho del robot. Según sus desarrolladores, Moxie tiene inteligencia emocional.

En Japón, un hombre se casó con un holograma casero, una figura femenina de voz dulce que lo saludaba al despertar, solía llamarlo por teléfono y le hacía sentir que estaba impaciente porque regresara a casa. A su vez, en Corea del Sur la inteligencia artificial le permitió a una madre una experiencia profundamente emotiva. Usando lentes de realidad virtual, pudo volver a ver a su hija de seis años, fallecida de leucemia. La niña aparecía de entre un montón de maderos apilados en lo que parecía ser un parque virtual, como si jugara a las escondidas. "¿Mamá, dónde estabas? Te he echado mucho de menos. ¿Tú me has echado de menos?". Con lágrimas en los ojos, la madre respondió: "Te extrañé, Na-yeon". Y extendió los brazos para estrechar a la niña. Pero no era su hija, ni siquiera era una niña real: era una imagen generada por inteligencia artificial.

Si bien el robot Moxie dejará de fabricarse, como anunció recién la empresa, por problemas financieros, la inteligencia artificial generativa (IAG) ha alcanzado nuevos niveles de interacción con los seres humanos. Los chatbots y los robots "emocionales" pueden convertirse en recursos de apoyo en el ámbito de la salud y la educación, o bien en medios para influir comportamientos y decisiones, utilizando los datos del usuario o sus sesgos cognitivos.

"Un diseño y una prueba minuciosos pueden ayudar a lograr los efectos conductuales previstos por el desarrollador", observa Laurence Devillers, investigadora en inteligencia artificial y profesora en la Universidad de La Sorbonne. Autora de *Los robots emocionales*, ella investiga en las interacciones humano-máquina en el Laboratorio de Informática y Ciencias de la Ingeniería del Centro Nacional para la Investigación Científica de Francia (CNRS).

Invitada al Congreso Futuro, que se realizó hace una semana, Laurence Devillers observa que "el poder persuasivo y la capacidad intrusiva de los nudges (empujoncitos) mejorados con IA también pueden provocar cambios profundos y duraderos en el comportamiento de los usuarios, como el aislamiento o la adicción, especialmente en niños y personas vulnerables".

Laurence Devillers es partidaria de regulaciones en el desarrollo de la IAG, así como de educar en torno a esta tecnología, sobre todo a los niños. En este sentido, cree necesario desmitificar las ideas en torno a la IAG, que a menudo parecen responder a imágenes de la ciencia ficción. Y del mismo modo, relevar las habilidades y aptitudes humanas que nos distinguen.

¿Hasta qué punto podría la inteligencia artificial acercarse a la inteligencia humana?

La IA generativa no percibe el mundo físico en tiempo real. Incluso, cuando está conectada a herramientas para acceder a información reciente, la IA no "comprende" realmente lo que procesa. No tiene sensibilidad o intuición humana para evaluar situaciones complejas. La IA generativa aprende en gran medida a partir de datos históricos. Por lo tanto, tiende a reproducir los sesgos del pasado y a tener dificultades para anticipar eventos imprevisibles. La IA generativa es una herramienta poderosa para procesar información, pero sigue estando fundamentalmente limitada por la ausencia de conciencia, experiencia directa y comprensión intrínseca. Puede proporcionar respuestas muy elaboradas, pero estas se basan en correlaciones y modelos matemáticos, no en una verdadera percepción del mundo. La IA generativa no es capaz de razonar, porque no tiene la capacidad de planificar, de tener objetivos concretos, intenciones y conatus (Spinoza). Tampoco tiene modelos mentales del mundo.

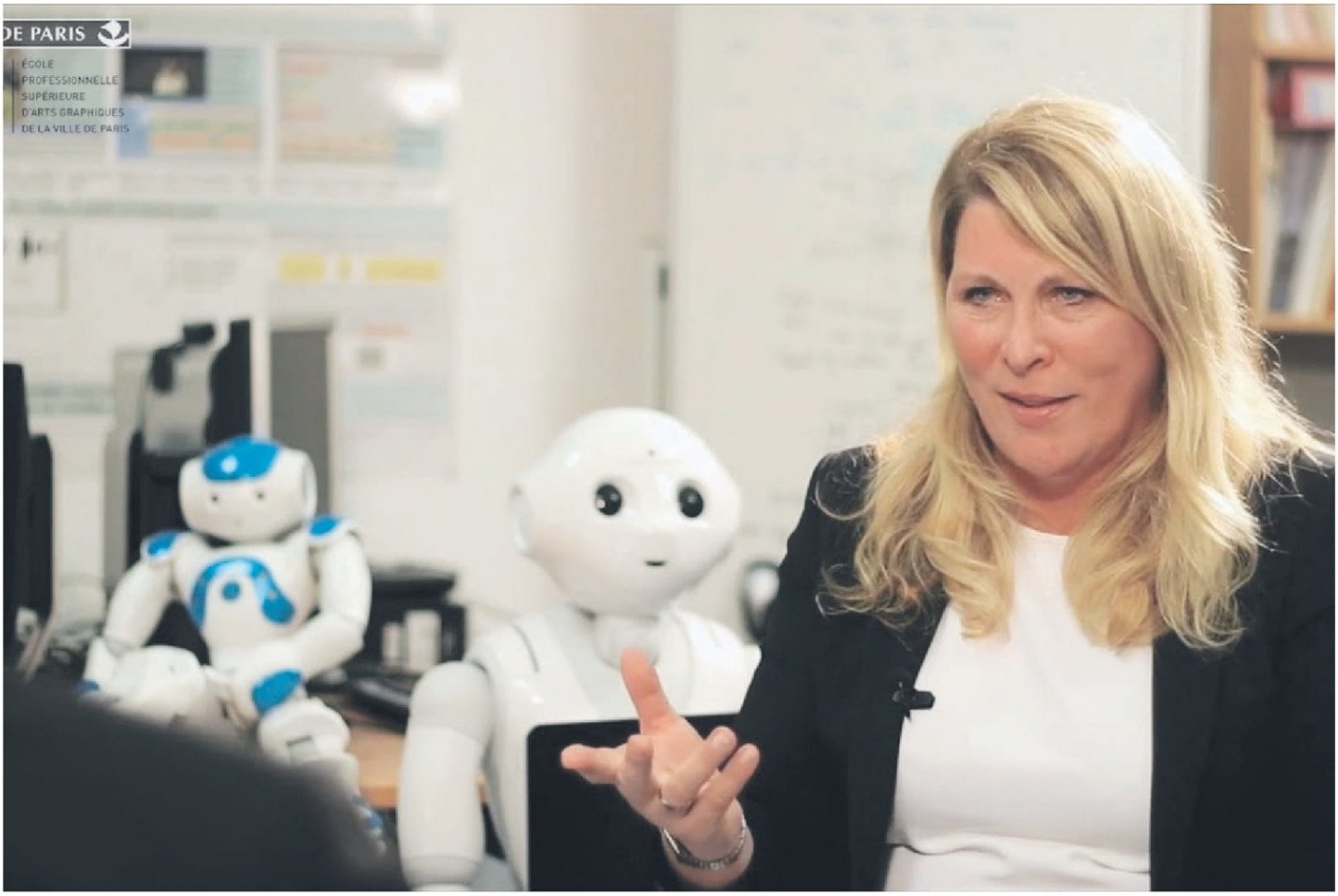
¿La IA generativa podría desarrollar una inteligencia emocional?

Simulará una inteligencia emocional, pero no sentirá nada. A diferencia de un humano, la IA generativa no tiene interacciones sensoriales con el mundo y no puede "sentir". Estos sistemas simplemente hacen matemáticas en corpus enormes. Y producen frases sin intención, sin emoción, sin conciencia, sin verificación.

¿La inteligencia artificial puede producir engaños?

Los datos en internet que se utilizan en el entrenamiento autosupervisado incluyen estrategias de evasión, mentiras, estafas. Por lo tanto, la IA podría también producirlos.

LA TERCERA
Fecha: 26/01/2025 Audiencia: 271.020 Sección: ACTUALIDAD
Vpe: $7.183.192 Tirada: 76.017 Frecuencia: DIARIO
Vpe pág: $9.829.612 Difusión: 76.017
Vpe portada: $9.829.612 Ocupación: 73,08%



► Laurence Devillers estuvo en Chile invitada al Congreso Futuro.

Solo tiene emergencias de razonamiento muy intuitivas que provienen de los datos aprendidos, pero que no son un modelo simbólico abstracto, sino un modelo reactivo cercano al sistema 1 de Daniel Kahneman (decisiones rápidas). El Sistema 2, es decir, el modelo cognitivo, el razonamiento deliberado, probablemente deba construirse con modelos híbridos neuronales y simbólicos. Aunque Kahneman no menciona un Sistema 3, algunos investigadores extrapolan este concepto para designar un nivel metacognitivo: capaz de integrar tanto el Sistema 1 como el Sistema 2. La IA generativa es incapaz de razonar, de distinguir qué es posible de lo que es imposible, ni lo que es verdadero de lo que es falso.

En la película Her, un hombre se enamora de un asistente de inteligencia artificial. Ahora, en Japón, un hombre dice haberse enamorado del holograma de Hatsune Miku. ¿La realidad y la ciencia ficción se están acercando?

Sí, las IA se vuelven cada vez más maquiavélicas al imitar a los humanos sin ninguna intención de hacerlo; simplemente está relacionado con los datos que han asimilado.

Si la inteligencia artificial es capaz de reconocer y simular emociones, ¿podría eventualmente manipularnos?

Mi cátedra se centra en el tema del nudge (empujón) digital, ¡sí, vamos a ser y ya estamos siendo manipulados por todas estas IA!

¿Qué piensa sobre la exposición de los niños a la inteligencia artificial? ¿Es arriesgado?

Es terrible para los niños que utilizarán estas IA como prótesis para hacer más rápido sus tareas sin nunca aprender a hacerlo por sí mismos. La escuela debe evolucionar muy rápido, no para reemplazar a los docentes por IA como ocurre en EE.UU., sino para establecer tres tipos de cursos: cursos sin ninguna máquina; cursos para aprender los conceptos fundamentales de estas IA y la ética de la IA, con el fin de agudizar el sentido creativo y crítico de los niños, y cursos donde se utiliza la IA para mejorar el conocimiento con un espíritu crítico. Es necesario formar los próximos talentos de la IA, que no deben ser solo consumidores de herramientas de IA, sino creadores de innovaciones.

En este sentido, ¿cómo deberían regularse estas tecnologías?

Hay tres pilares: ley, normas y ética. ¡Es necesario transformar la educación para desmitificar estas IA: los usos, la ética y los conceptos fundamentales de las IA!

La desregulación iniciada por Trump y Musk comienza a influir en todo el sector tecnológico estadounidense, incluidos actores como Meta. Elon Musk ha expresado a menudo su convicción transhumanista de que las tecnologías avanzadas, como la IA, la neurotecnología y los implantes cerebrales, pueden y deben ser utilizados para trascender los límites naturales de la humanidad. ¡Vamos a vivir un verdadero punto de inflexión!

La llegada de ChatGPT ha suscitado un gran interés y un impacto social. Para algunos, estamos viviendo un momento de cambio histórico. ¿Cuál es su opinión al respecto?

La llegada de ChatGPT ha desencadenado una guerra económica con repercusiones sociales, ecológicas y éticas que no son prioritarias para quienes dominan el mercado. La competencia entre los gigantes digitales en torno a la IA generativa se ha intensificado desde la aparición de ChatGPT, marcada por una ola de revelaciones espectaculares sobre la llegada de la IA general y colosales inversiones que se cuentan en miles de millones de dólares. Cuando el reciente premio Nobel de Economía, Daron Acemoglu, da la voz de alarma sobre las inversiones en IA, es mejor prestar atención. Él se preocupa por el entusiasmo excesivo que ella suscita entre los inversores. Según él, solo una ínfima proporción de los empleos (alrededor del 5%) es verdaderamente automatizable y susceptible de ser reemplazada por la IA, lo cual ¡no es seguro! Esto nos lleva a relativizar las actuales inversiones masivas.

Esto bien puede ser una burbuja que va a estallar o un premio gordo para unos pocos de ellos. Creo que las herramientas de IA generativa nunca llegarán a la IA general: ¡La inteligencia artificial general con la que sueñan todos estos innovadores!

Yann Le Cun, científico jefe de IA en Meta, tiene la misma posición: el futuro de la IA no está en los modelos de lenguaje avanzado (LLM), sino en la IA guiada por objetivos.

Otros dicen que "¡la inteligencia artificial general (IAG) estaría al alcance de la mano!". Ray Kurzweil, transhumanista en Google, ha predicho que la IAG podría emerger para 2029. Hacer creer en la emergencia de una IAG permite hacer estas tecnologías aún más fascinantes. A fines de 2024, los investigadores de Appolo Research publicaron un estudio sobre las versiones más avanzadas de estos grandes modelos, mostrando que la IA puede mentir de manera estratégica. Sobre todo, demuestran que optimizar una IA a través del aprendizaje por refuerzo para que esté más alineada con los valores humanos, que consiste en entrenar al sistema aplicándole repetidamente recompensas y castigos, no es suficiente para crear modelos fiables y seguros.