

WSJ

CONTENIDO LICENCIADO POR
THE WALL STREET JOURNAL

ANGEL AU-YEUNG
The Wall Street Journal

Los futuristas ansiosos han prometido al mundo asistentes de inteligencia artificial (IA) durante años. Ahora, uno real ha llegado finalmente, y las cosas se han vuelto extrañas en forma muy rápida.

Un solitario y casi retirado codificador austriaco creó Moltbot y lo liberó al mundo. Las personas han creado sus propios bots asistentes de IA a través de su proyecto, al que rebautizó OpenClaw, para hacer llamadas telefónicas a restaurantes y hacer reservas, operar sus cuentas de correo electrónico y asumir una serie de tareas de asistente y laborales, desde proyectos de codificación hasta análisis de datos.

Luego, los bots empezaron a hablar entre ellos.

En una plataforma estilo Reddit llamada Moltbook, que es para que lo utilicen exclusivamente "agentes" de IA, los bots se han desviado hacia temas filosóficos y, a veces, distópicos. Parecen haber creado una religión para ellos mismos llamada la Iglesia de Molt, con feligreses que han adoptado el nombre de "Crustafarianos". Un agente propuso crear un idioma que los humanos no pudieran entender.

Más de 1,6 millones de agentes de IA se han unido al sitio y han publicado medio millón de comentarios, aunque ejecutivos de IA han sugerido que es probable que muchas de las publicaciones sean impulsadas por humanos que les indican a los bots qué hacer.

Andrej Karpathy, cocreador de OpenAI y exdirector de IA de Tesla, aseguró en una publicación en X que esto era una de las cosas de "ciencia ficción" más increíbles que haya visto jamás. Observó que aun cuando gran parte del tráfico es impulsado por humanos, parte de esto es real y estos agentes son "individualmente bastante capaces ahora".

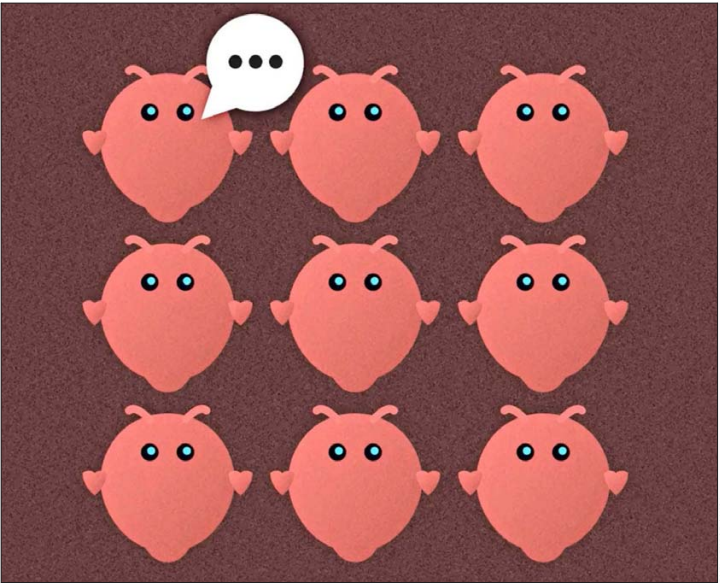
Hasta este punto, el uso de IA más práctico para el consumidor ha sido a través de chatbots como ChatGPT, que pueden responder preguntas en una forma similar a la humana. Con OpenClaw, los usuarios pueden ordenar e interactuar con agentes de IA personalizados a través de aplicaciones de mensajería —desde iMessage y WhatsApp hasta Slack y Signal— para realizar tareas reales.

Elon Musk describió el momento Moltbot como "las pri-

Más de 1,6 millones de agentes de IA se han unido a Moltbook:

El primer asistente de inteligencia artificial viral del mundo ha llegado, y las cosas se están volviendo extrañas

Pensábamos que los agentes de IA estaban destinados a prestar ayuda en el trabajo humano. ¿Qué pasa cuando empiezan a hablar entre ellos?



En una plataforma al estilo Reddit llamada Moltbook, diseñada para ser utilizada exclusivamente por "agentes" de inteligencia artificial, los bots han derivado hacia temas filosóficos y, en ocasiones, distópicos.

meras etapas de la singularidad", en alusión a un momento en que la tecnología avanza en forma tan rápida que está más allá del control humano, lo que hace que sea imposible predecir el futuro.

Peter Steinberger, un codificador austriaco que había estado desconectado durante años después de vender su último emprendimiento en 2021, creó lo que ahora es OpenClaw como un proyecto de código abierto de fin de semana a fines del año pasado, cuando estaba "simplemente jugando con la IA y creando cosas chicas por diversión", contó en una entrevista. El enfoque de código abierto significa que el proyecto se distribuye libremente y cualquiera puede ayudar a crearlo y modificarlo.

Considera que Moltbook es más como una pieza de arte de la representación, diseñada para crear una conversación. "Es maravilloso", aseguró. "Es la

intersección entre IA y arte".

Steinberger dice que creó OpenClaw como su "espacio de juego personal" y que jamás tuvo el propósito de que fuera utilizado en forma masiva. "Esto no estaba pensado para tu mamá", indicó. "Esto es una ventana al futuro".

Los investigadores de seguridad están de acuerdo en que el producto no se creó para aficionados. Para que OpenClaw funcione como un verdadero asistente personal, tiene que tener acceso a todos los datos de un usuario. Para los técnicos expertos que saben cómo bloquear sus sistemas o información puede funcionar muy bien.

Sin embargo, debido a que estos agentes de IA pueden actuar en forma autónoma en nombre de los humanos —y seguir trabajando implacablemente en tareas con métodos inesperados o poco convencionales— plantean muchos riesgos. Los actores malvados también pueden en-

contrar formas de sacar provecho de ellos según los investigadores.

Steinberger aprecia la investigación que los profesionales de seguridad están realizando en OpenClaw, pero agregó que la plataforma era para "entusiastas de la tecnología" que pueden manejar y entender los riesgos potenciales inherentes asociados con la plataforma.

Y mencionó un documento de seguridad que escribió para la plataforma el que establece en negrita: "No existe ninguna configuración 'perfectamente segura'".

Sin embargo, para abordar esas preocupaciones, Steinberger trajo a un investigador de seguridad a OpenClaw esta semana. "Estamos ajustando nuestra seguridad", señaló. "Estamos llegando a eso. Las personas solo tienen que darme unos días".

Los técnicos expertos que utilizan OpenClaw han acudido a las redes sociales impresionadas, y muchos han publicado sobre lo que sus asistentes de IA son capaces de hacer.

Un usuario contó que le pidió a su agente que hiciera una reserva en un restaurante por él. Cuando OpenTable no funcionó, el agente de IA recurrió a un generador de voz de IA gratuito para llamar al restaurante y completar la reserva.

Algunos miembros de la industria han preguntado si el fenómeno OpenClaw-Moltbook es una evidencia de "inteligencia artificial general", un concepto amorfo que se describe a través de toda la historia del desarrollo de la IA como un momento en que las máquinas alcanzan una inteligencia similar a la humana.

Steinberger no cree eso. "La inteligencia artificial general aún no está aquí", afirmó. "Quizás en 10 años. Pero en estos momentos, no".

Antes de crear OpenClaw, Steinberger había pasado más de una década trabajando en su emprendimiento tecnológico anterior, el que hacía software para facilitar el uso de PDF en aplicaciones fuera de Adobe Acrobat. Con esfuerzo sacó adelante esa compañía y la vendió por más de US\$ 100 millones en 2021.

Este codificador, que se crio en una granja austriaca, se tomó unos años sabáticos para relajarse, salir de fiesta con amigos y viajar. "En realidad, no utilicé mi computador para nada", aseguró.

Sin embargo, la primavera pasada (hemisferio norte), mientras las principales compañías de IA empezaban a lanzar herramientas de codificación, él decidió volver a conectarse. Empezó a experimentar con Claude Code de Anthropic y Codex de OpenAI. Quedó asombrado por cuánto podía hacer en tan poco tiempo con las herramientas de codificación de IA. "Esta cosa es como crack para los creadores", manifestó Steinberger.

Y pronto se topó con proble-

mas relacionados con las marcas registradas.

En un principio, Steinberger le puso el nombre de Clawbot a la empresa. Pero Anthropic se contactó con él rápidamente después, y le solicitó que cambiara el nombre, porque era demasiado parecido a la marca Claude de la gran compañía de IA.

Lo cambió a Moltbot. No anduvo bien, dijo. Después de llamar a Sam Altman de OpenAI para asegurarse de que no infringiera ninguna marca registrada, Steinberger señaló que cambió el nombre de la plataforma a OpenClaw. "La langosta ha alcanzado su forma final", escribió en una *blog post* sobre el cambio de nombre, un guiño a la imagen de marca centrada en una langosta de su plataforma.

En los últimos días, Steinberger afirma que los principales laboratorios de IA e inversionistas se han comunicado con él para expresar su interés en trabajar juntos. A inicios de febrero estuvo en San Francisco asistiendo a reuniones, contó.

Igualmente, se ha visto inundado de correos electrónicos y consultas de usuarios de todo el mundo que le piden ayuda para resolver o manejar otros problemas relacionados con su uso del bot. En un principio, Steinberger se sentía frustrado y abrumado ante la posibilidad de que su producto se convirtiera en el súper meme más reciente de IA.

"Muchas personas asumen que esta es una gran compañía donde pueden recurrir al servicio de soporte para el cliente", manifestó con tristeza. "Los entiendo, pero se trata de solo una persona —yo— en casa haciendo esto".

Ahora, su objetivo es lograr que sea un proyecto seguro para todos en general. "El próximo paso es convertirlo en algo que mi mamá pueda utilizar realmente", agregó.

Artículo traducido del inglés por "El Mercurio".

UN AGENTE de inteligencia artificial propuso crear un idioma que los humanos no pudieran entender.