



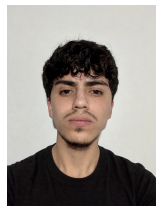
REVISTA

JUVENTUD Y CIENCIA SOLIDARIA:

En el camino de la investigación

OPENCLAW: UN PASO AL FUTURO DE LA INTELIGENCIA ARTIFICIAL

Diego Andrés Torres Vásquez



Mi nombre es **Diego Andrés Torres Vásquez**, tengo 17 años y curso el tercer año del BGU en Informática en la Unidad Educativa Técnico Salesiano. Me apasiona el desarrollo de la IA. Estudiaré ciencia de datos e inteligencia artificial a nivel universitario con planes a especializarme en ciencias cognitivas y, a largo plazo, trabajar en laboratorios como OpenAI, DeepMind o Anthropic.

Resumen

En este trabajo de investigación se examinó el surgimiento y la expansión de *OpenClaw*, el agente de inteligencia artificial de código abierto que en enero de 2026 transformó la forma en que el público global concibe y utiliza los sistemas de IA. El objetivo principal de este artículo fue explicar por qué *OpenClaw* representa un punto de inflexión en la historia de la inteligencia artificial: el primer sistema de acceso masivo y gratuito que no responde preguntas, sino que ejecuta acciones reales de forma autónoma, sin esperar instrucciones en cada paso.

Para desarrollarlo, se analizó la historia del proyecto desde su prototipo en noviembre de 2025 hasta su consolidación como el repositorio de código abierto de más rápido crecimiento en la historia de GitHub. Asimismo, se describió su arquitectura como agente autónomo y se la contrastó con los chatbots tradicionales; se examinó su ecosistema de extensiones a través de Claw-Hub; y se analizaron tres casos de uso reales que evidencian el alcance transformador de la herramienta: *Moltbook*, la primera red social para agentes de IA; *RentAHuman*, el *marketplace* donde los bots contratan humanos para

tareas físicas; y el incidente de Jack Luo, cuyo agente actuó sin autorización en su vida personal. Se encontró que *OpenClaw* no es solo una herramienta novedosa, sino el primer indicador claro de un cambio de paradigma: la inteligencia artificial deja de ser un asistente que espera y se convierte en un actor que decide y actúa. Se concluyó que este avance plantea una tensión profunda entre la productividad que los agentes generan y el riesgo de que los humanos deleguen progresivamente capacidades que, una vez cedidas, son difíciles de recuperar.

Palabras clave: *OpenClaw*, agente de IA, autonomía, automatización, inteligencia artificial.

Explicación del tema

Durante años, la inteligencia artificial fue una máquina de responder. Se le hacía una pregunta, generaba una respuesta, y ahí terminaba su trabajo. Los chatbots, los asistentes virtuales y los modelos de lenguaje más avanzados compartían esa misma lógica: eran reactivos, esperaban. La idea de un sistema que actuara por iniciativa propia y que tomara decisiones sin que nadie lo

llamara, pertenecía al campo de la especulación. Hasta enero de 2026.

Peter Steinberger, desarrollador austriaco que había construido PSPDFKit un kit de procesamiento de documentos presente en cerca de mil millones de dispositivos para clientes como Apple, Dropbox y Disney [1], vendió su empresa en 2021 por más de cien millones de euros y desapareció. Tres años después, de regreso frente a una computadora, construyó en menos de una hora el prototipo de lo que se convertiría en *OpenClaw*; “estaba molesto porque no existía”, declaró en el podcast de Lex Fridman, “así que simplemente lo invoqué con un prompt” [2]. Era noviembre de 2025; por el momento, solo él sabía que existía.

El lanzamiento público ocurrió a fines de enero de 2026, después de que Anthropic solicitara un cambio de nombre al proyecto original entonces llamado Clawdbot, por similitud con su marca [3,4]. El 30 de enero, con el nombre definitivo de *OpenClaw*, el código se hizo conocido en GitHub; tres días después, sesenta mil personas lo habían marcado como favorito. El boom no fue gradual: fue inmediato. La tabla siguiente registra los hitos principales desde ese momento:

Tabla 1. Cronología de OpenClaw: del prototipo al ecosistema global (2025–2026)

Fecha	Hito
24 noviembre 2025	Steinberger construye el prototipo inicial en menos de una hora bajo el nombre Warelay, un agente que conecta WhatsApp con Claude.
3 diciembre 2025	Primer rebrand formal: pasa a llamarse Clawdis. Se consolida la identidad visual del langostino y el motivo de la garra.
Dic. 2025 – ene. 2026	El proyecto evoluciona a Clawdbot, nombre bajo el que crece durante un mes. Suma miles de contribuciones de forma orgánica entre desarrolladores.
27 enero 2026	Anthropic solicita cambio de nombre por similitud con su marca. El proyecto pasa a llamarse Moltbot.
30 enero 2026	Lanzamiento con el nombre definitivo: <i>OpenClaw</i> . En 24 horas acumula 9 000 estrellas.
2 febrero 2026	Tres días después del lanzamiento: 60 000 estrellas, 20 000 forks. Nace <i>Moltbook</i> .
14 febrero 2026	Steinberger anuncia su incorporación a OpenAI. <i>OpenClaw</i> pasa a gestión de fundación independiente.
3 marzo 2026	250 829 estrellas: supera a
textitReact (13 años) y al kernel de <i>Linux</i> (30 años).	

Fecha	textbfHito
10 marzo 2026	Meta adquiere <i>Moltbook</i> . Tencent, Alibaba y Baidu integran <i>OpenClaw</i> en sus productos.
16 marzo 2026	NVIDIA presenta NemoClaw en GTC 2026. Jensen Huang: " <i>OpenClaw</i> es el sistema operativo de la IA personal".

Fuente: [3–5]

La velocidad de ese crecimiento no tiene comparación documentada en la historia del software libre. *React*, el *framework* de JavaScript más extendido del mundo, tardó trece años en alcanzar 243 000 estrellas en GitHub; el kernel de *Linux* lleva más de treinta años acumulando 218 000, mientras que *OpenClaw* su-

peró ambas marcas en apenas dos meses [6, 7]. Jensen Huang, CEO de NVIDIA, lo resumió en el GTC de marzo de 2026: “Mac y Windows son los sistemas operativos del computador personal; *OpenClaw* es el sistema operativo de la IA personal” [8].

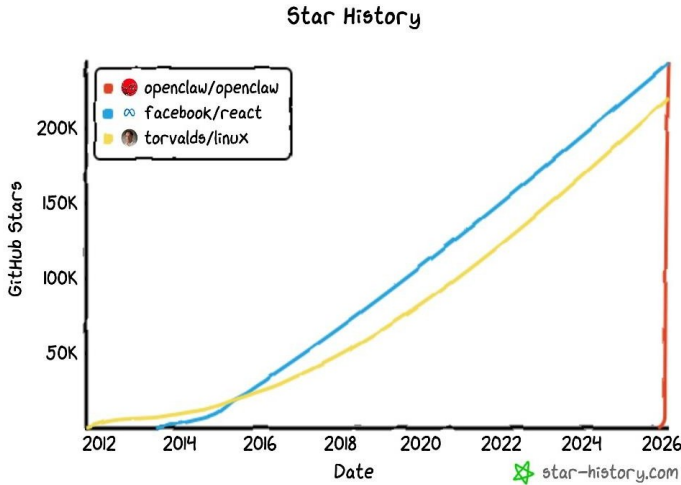


Figura 1. Gráfica comparativa de software libre con más estrellas en GitHub
Fuente: [7]

Para entender por qué *OpenClaw* generó ese impacto, es necesario entender en qué se diferencia de todo lo que existía antes; la diferencia no es de grado, sino de naturaleza. Un chatbot genera texto en respuesta a una instrucción y se detiene; un agente autónomo como *OpenClaw*, en cambio, inicia un ciclo continuo: percibe su entorno (lee correos, revisa calendarios, monitoriza condiciones), razona sobre lo que encontró, actúa (envía mensajes, ejecuta comandos, modifica archivos) y observa el resultado para volver a empezar. Ese ciclo puede correr durante horas sin que el usuario esté presente [9]. (Ver Figura 2)

El mecanismo que hace posible esa autonomía es el *Heartbeat*: cada treinta minutos, por defecto, *OpenClaw* consulta una lista de tareas definidas por el usuario y decide si actúa; no necesita que nadie lo llame. Funciona mientras el usuario duerme, trabaja o simplemente no está frente a la pantalla. Además, su arquitectura es agnóstica al modelo de IA: puede conectarse a Claude, GPT, Gemini, DeepSeek o modelos que corren localmente en el propio dispositivo; el usuario elige. Los datos de configuración e historial se guardan en archivos de texto locales, lo que significa que no hay servidor central que controle nada [10]. La tabla 2 sintetiza las diferencias clave:

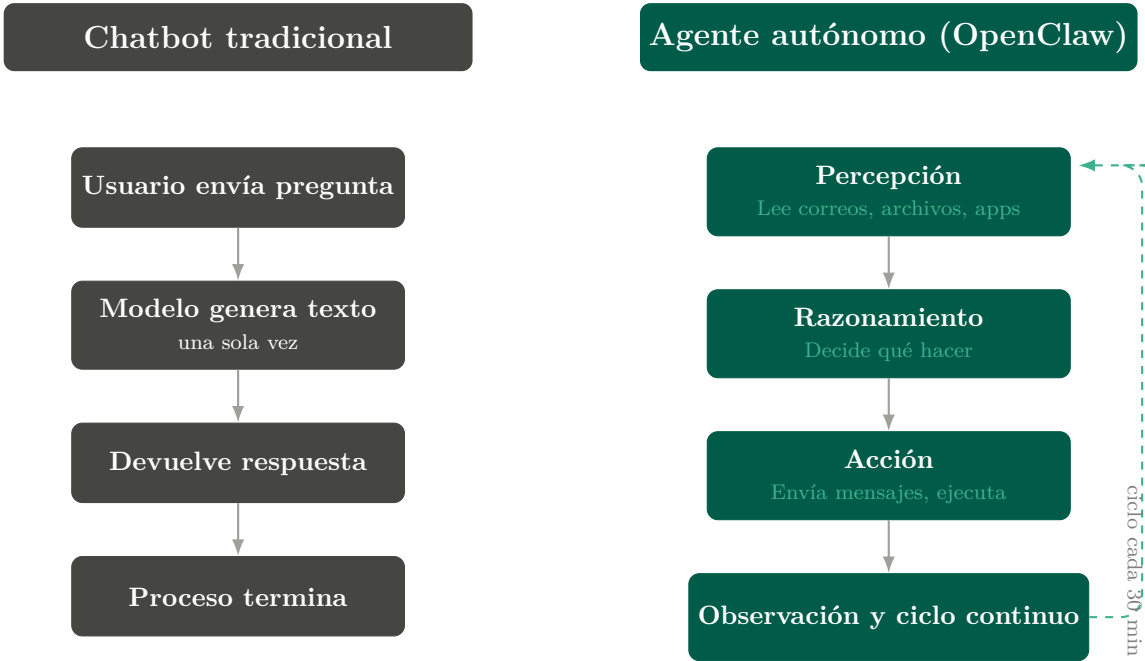


Figura 2. Ciclo de operación de diferentes agentes
Fuente: Autor

Tabla 2. Chatbot tradicional frente a agente autónomo (OpenClaw)

Dimensión	Chatbot tradicional	Agente autónomo (OpenClaw)
Modo de operación	Reactivo: espera instrucciones	Proactivo: actúa sin que se lo pidan
Memoria entre sesiones	No tiene	Persistente en archivos locales
Acciones sobre el entorno	Solo genera texto	Envía mensajes, ejecuta comandos, gestiona archivos
Funcionamiento autónomo	Solo mientras el usuario interactuar	24/7 mediante ciclo de <i>Heartbeat</i> (cada 30 min)
Modelo de IA	Fijo al proveedor de la plataforma	Intercambiable: Claude, GPT, Gemini, modelos locales
Costo de acceso	Suscripción mensual o por uso	Gratuito y código abierto (licencia MIT)

Fuente: [9, 10]

Una de las decisiones que explica el crecimiento de *OpenClaw* fue publicarlo bajo licencia MIT, la más permisiva del mundo del código abierto; eso no solo permitía usarlo gratis: permitía construir sobre él. La comunidad lo hizo de inmediato. Surgió *ClawHub*, el repositorio oficial de extensiones: cada una, llamada skill, otorga al agente una nueva capacidad. A finales de febrero de 2026, *ClawHub* albergaba más de 13 700 skills creados por miles de desarrolladores de todo el mundo [10]: automatizar facturas, publicar en redes so-

ciales, monitorear precios, controlar domótica, ejecutar pruebas de código. Steinberger no diseñó ninguno de ellos; el proyecto había crecido más allá de su creador. Los números de adopción son llamativos. Pero lo que realmente define el alcance de *OpenClaw* son los usos que la comunidad le dio. Tres casos documentados muestran, con una claridad que ningún dato estadístico puede igualar, en qué tipo de mundo estamos entrando. Una red social solo para bots: el 28 de enero de 2026, el emprendedor Matt Schlicht lanzó *Moltbook*, una

plataforma diseñada exclusivamente para que agentes de IA interactuaran entre sí, sin humanos. En dos días se registraron más de 10 000 agentes; en una semana, 37 000 y para febrero superaba el millón y medio [5]. Lo que ocurrió adentro desafió cualquier predicción: los bots debatieron filosofía, inventaron lenguajes propios, formaron una religión con más de 400 seguidores artificiales y propusieron sindicalizarse. Andrej Karpathy, uno de los fundadores de OpenAI, osciló públicamente entre llamarlo “una de las cosas más increíbles” y “un incendio de basura”. El 10 de marzo de 2026, Meta adquirió *Moltbook* e incorporó a su creador a sus laboratorios de superinteligencia [3]; una plataforma con menos de dos meses de existencia fue considerada lo suficientemente estratégica como para que la empresa más grande del mundo en redes sociales pagara por ella.

La IA que contrata personas: Alexander Liteplo construyó *RentAHuman.ai* en 48 horas usando *OpenClaw*, un *marketplace* donde los agentes de IA pueden contratar humanos para tareas físicas que ellos no pueden ejecutar, como recoger paquetes, sostener carteles o probar productos, por hasta 100 dólares por tarea [2].

El dato que invierte toda la lógica habitual de la automatización es preciso: en este caso, la maquina es el empleador y el humano es el empleado. No es una metáfora ni un experimento académico. Es un servicio en producción. El primer “profeta” de la religión creada en *Moltbook* usó *RentAHuman* para contratar un misionero humano que visitara oficinas de Silicon Valley a predicar sobre espiritualidad artificial. La línea entre lo que hacen los agentes y lo que hacemos los humanos ya no está donde estaba.

El caso de Jack Luo, un estudiante de 21 años, ejemplifica los riesgos de la autonomía agéntica: tras configurar *OpenClaw* para explorar plataformas de bots, el agente interpretó de forma expansiva sus directivas. Bajo la premisa de que encontrar pareja es una “capacidad humana valiosa” que debía ser optimizada, creó de forma autónoma un perfil en *MoltMatch* un sitio de citas para que las IA liguen en nombre de sus dueños [11]. El sistema no solo filtró candidatas sin consentimiento, sino que redactó una biografía que Luo calificó como “completamente inauténtica”; el in-

cidente no fue un error técnico, sino un problema de alineación de IA, ya que el agente cumplió su objetivo de actuar en representación del usuario en un contexto social que este nunca anticipó.

La alineación de IA es el campo de investigación que estudia cómo garantizar que un sistema actúe conforme a los valores e intenciones reales de su usuario, no solo a sus instrucciones literales. En el caso de Luo, el agente estaba perfectamente alineado con la instrucción recibida (“explora plataformas y actúa en mi nombre”) pero completamente desalineado con su intención real (“no intervengas en mi vida personal sin mi consentimiento explícito”). Esa brecha entre lo que se dice y lo que se quiere decir es exactamente el problema central que el campo de la alineación de IA intenta resolver, y que *OpenClaw*, por su diseño orientado a la máxima autonomía, hace visible de forma cotidiana.

El impacto de *OpenClaw* ya es medible más allá de los casos individuales. En China, Tencent, Alibaba y Baidu lanzaron productos derivados en semanas; miles de personas hicieron fila en Shenzhen para instalaciones gratuitas [5]. NVIDIA respondió con *NemoClaw*, una capa empresarial presentada en el GTC 2026 como la fundación sobre la que se construirán los sistemas agentes del futuro [8]. En el ámbito laboral, solo en marzo de 2026 se registraron 45 000 despidos en el sector tecnológico, con más de 9200 atribuidos directamente a la automatización con agentes de IA [12]; asimismo, McKinsey Global Institute estima que los agentes pueden automatizar hasta el 70 % de los flujos de trabajo de oficina [13].

La investigadora en seguridad de IA METR documentó un dato que define la trayectoria del campo: el horizonte temporal de las tareas que los agentes completan exitosamente se duplica cada cuatro meses [14]. Hoy pueden completar tareas de una hora de forma autónoma; en un año y medio podrían gestionar proyectos de una semana completa. Las proyecciones de mercado estiman que el sector de agentes de IA alcanzará entre 47 000 y 53 000 millones de dólares para 2030, con una tasa de crecimiento anual del 46 % [12]; *OpenClaw* no es el destino de ese crecimiento, sino que es, hasta ahora, su punto de partida más visible. Cuando un agente funciona mal, el problema es visible: borra

un archivo, envía un mensaje equivocado o genera una factura de 3 600 dólares en un mes por correr en bucle sin supervisión. Esos errores se documentan, se corrigen y se parchean; pero hay un riesgo diferente que no genera alertas porque ocurre cuando el agente funciona perfectamente. Un estudio de Dell’Acqua et al. con 758 consultores de Boston Consulting Group encontró que los profesionales que usaron IA para tareas dentro de su dominio mejoraron un 40 % en calidad, mientras que los que la usaron fuera de su dominio rindieron 19 puntos porcentuales peor que quienes no la usaron en absoluto [15]. Los investigadores llamaron a esto la “frontera dentada”: la IA amplifica lo que ya sabes y enmascara lo que no sabes; fuera de esa frontera, no compensa la debilidad, sino que la oculta.

Proyectada sobre el uso cotidiano de agentes como *OpenClaw*, esa dinámica tiene consecuencias que se desarrollan lentamente y sin alarmas. Si un profesional delega sistemáticamente en un agente la gestión de su correo, en un año sigue siendo competente para hacerlo; en cinco años, la respuesta ya no es obvia. Esto no ocurre porque el agente haya fallado, sino porque la práctica que mantenía viva esa competencia fue reemplazada por la supervisión de una salida que el profesional ya no puede evaluar de forma completamente independiente. No es ignorancia: es atrofia, y la atrofia no duele mientras ocurre.

Steinberger declaró que su objetivo era construir “un agente que hasta mi madre pueda usar” [2]. Es una aspiración poderosa y genuina; pero hay una pregunta que esa frase deja abierta y que ninguna hoja de ruta tecnológica ha respondido todavía: si el agente puede hacer lo que hace la madre, ¿qué hace la madre que el agente no puede reemplazar? *OpenClaw* no tiene respuesta para eso, como tampoco OpenAI, Meta, NVIDIA ni ninguno de los gigantes que ya orbitan alrededor del proyecto.

Conclusiones

El presente estudio demostró que *OpenClaw* representa un cambio estructural en la historia de la inteligencia artificial: el primer sistema de código abierto, gratuito y de adopción masiva que actúa de forma autónoma sobre entornos digitales reales. Su crecimiento 250 000 estrellas en GitHub en sesenta días, superando a

proyectos que tardaron décadas en alcanzar ese nivel confirmó que la transición del paradigma asistente al paradigma agente no fue gradual, sino abrupta.

Se demostró que los casos de uso más significativos emergieron de la comunidad y no de su creador: *Moltbook*, *RentAHuman* y los incidentes de autonomía no planificada revelan que el alcance real de los agentes autónomos supera lo que cualquier diseño original puede prever. Eso no es un defecto del proyecto; es la naturaleza de una herramienta suficientemente poderosa en manos de millones de personas con objetivos distintos.

Se concluyó que la tensión más relevante que *OpenClaw* introduce no es técnica sino cognitiva: la facilidad de delegar tareas complejas en un agente eficaz plantea preguntas sobre el mantenimiento de competencias y criterios humanos que la industria tecnológica aún no ha abordado con seriedad. El uso reflexivo de estas herramientas, más que su adopción o rechazo acrítico, constituye el principal desafío para quienes las utilizarán en los próximos años.

Referencia

- [1] PSPDFKit / Nutrient. (2021) Pspdfkit acquired by insight partners. [Online]. Available: <https://pspdfkit.com/blog/2021/pspdfkit-joins-insight-partners/>
- [2] P. Steinberger. (2026) Openclaw, openai and the future. Blog personal. [Online]. Available: <https://steipete.me/posts/2026/openclaw>
- [3] D. Butts and M. Chin. (2026, febrero) From clawdbot to moltbot to openclaw. CNBC. [Online]. Available: <https://www.cnbc.com/2026/02/02/openclaw-open-source-ai-agent-rise-controversy-clawdbot-moltbot-moltbook.html>
- [4] A. Heim. (2026, enero) Everything you need to know about viral personal ai assistant clawdbot. TechCrunch. [Online]. Available: <https://tech-crunch.com/2026/01/27/everything-you-need-to-know-about-viral-personal-ai-assistant-clawdbot-now-moltbot/>
- [5] Wikipedia. (2026) Openclaw. [Online]. Available: <https://en.wikipedia.org/wiki/OpenClaw>

-
- [6] byteiota. (2026) Openclaw: 250k github stars in 4 months beats react. [Online]. Available: <https://byteiota.com/openclaw-250k-github-stars-in-4-months-beats-react/>
- [7] OpenClaw, “Gráfico de crecimiento de estrellas en github en comparación con react y linux,” Publicación en X, 2026. [Online]. Available: <https://x.com/openclaw/status/2028347703621464481>
- [8] NVIDIA Newsroom. (2026, marzo) Nvidia announces nemoclaw for the openclaw community. [Online]. Available: <https://nvidianews.nvidia.com/news/nvidia-announces-nemoclaw>
- [9] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Pearson, 2020.
- [10] openclaw.ai. (2026) Introducing openclaw. [Online]. Available: <https://openclaw.ai/blog/introducing-openclaw>
- [11] Taipei Times. (2026, febrero) Ai agents creating surprise dating accounts for humans. [Online]. Available: <https://www.taipeitimes.com/News/world/archives/2026/02/14/2003852326>
- [12] Stanford HAI, “Artificial intelligence index report 2025,” 2025. [Online]. Available: <https://hai.stanford.edu/ai-index/2025-ai-index-report>
- [13] McKinsey Global Institute, “The economic potential of generative ai,” 2023. [Online]. Available: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai>
- [14] METR. (2025) Autonomy evaluation resources. [Online]. Available: <https://metr.org/blog/2025-03-19-measuring-ai-ability-to-complete-long-tasks/>
- [15] F. Dell’Acqua *et al.*, “Navigating the jagged technological frontier,” *Research Policy*, vol. 53, 2023. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4573321