Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

Capacidad sin autonomía en el auto-mejoramiento de la inteligencia artificial.
Capacidad sin autonomía en la auto-mejora de la inteligencia artificial
Capacidad sin autonomía en el auto-mejoramiento de la inteligencia artificial

Deryk Willyan Biotto,
profesor de Educación Básica, Técnica y Tecnológica (EBTT) en el Instituto Federal de Educación, Ciencia y Tecnología del Sur de Minas Gerais (IFSULDEMINAS), Campus Inconfidentes. Correo electrónico: deryk.biotto@ifsuldeminas.edu.br.

Resumen:

El debate en torno a la auto-mejora de la inteligencia artificial (IA) suele plantearse cuándo ocurrirá. Este ensayo teórico, basado en una revisión narrativa de publicaciones, principalmente de 2023 a 2026, y organizado mediante herramientas de pensamiento sistémico, cuestiona en qué momento la evolución de la IA dejará de depender de la intervención humana. La auto-mejora se describe como cuatro ciclos de retroalimentación con diferentes tiempos de respuesta: software, investigación, hardware e infraestructura física. Históricamente, las capacidades de la IA no han crecido de forma continua, como los chips, sino en curvas en S, limitadas por barreras como la capacidad de procesamiento.

El ensayo distingue dos umbrales: el punto de inflexión de la capacidad, cuando la mejora de la IA deja de depender de la investigación humana, y el punto de inflexión de la autonomía, cuando se cierra el ciclo físico y la IA comienza a impulsar, sin la intervención humana indispensable, la evolución de su propia base material: chips, fábricas, centros de datos y energía. Argumenta que solo el segundo marcaría el punto en el que la humanidad pierde el control sobre la evolución de la IA; antes de eso, existe capacidad sin autonomía. El ciclo puede cerrarse de dos maneras: los robots ejecutan mientras las personas aún deciden, o la IA decide mientras las personas aún ejecutan. El ensayo describe los factores que aceleran o retrasan este cierre, organiza cuatro futuros posibles y propone señales para anticiparlo, sin evaluar si sería deseable ni cómo evitarlo.

Palabras clave: Inteligencia artificial. Automejora recursiva. Infraestructura.
Automatización. Estudios de futuro.

Abstracto:

Los debates sobre la auto-mejora de la inteligencia artificial (IA) suelen centrarse en cuándo ocurrirá. Este ensayo teórico, basado en una revisión narrativa de publicaciones, principalmente de 2023 a 2026, y organizado con herramientas de pensamiento sistémico, plantea la cuestión de en qué momento la evolución de la IA dejaría de depender de la intervención humana. La auto-mejora se describe como cuatro bucles de retroalimentación con diferentes tiempos de respuesta: software, investigación, hardware e infraestructura física. Históricamente, la capacidad de la IA no ha crecido de forma continua, como el rendimiento de los chips, sino en curvas en S, limitadas por barreras, entre ellas la capacidad de procesamiento. El ensayo distingue dos umbrales: el punto de inflexión de la capacidad, cuando la mejora de la IA ya no depende del trabajo de investigación humana, y el punto de inflexión de la autonomía, cuando se cierra el bucle físico y la IA comienza a llevar a cabo, sin un papel humano indispensable, la evolución de su propia base material: chips, fábricas, centros de datos y energía. Argumenta que solo el segundo punto marcaría el momento en que la humanidad pierde el control sobre la evolución de la IA; antes de este punto, existe capacidad sin autonomía. El ciclo puede cerrarse de dos maneras: los robots realizan el trabajo mientras las personas siguen decidiendo, o la IA decide mientras las personas siguen realizando el trabajo. El ensayo describe los factores que aceleran o retrasan este cierre, organiza cuatro futuros posibles y propone señales para anticiparlo, sin evaluar si sería deseable ni cómo prevenirlo.

Palabras clave: Inteligencia artificial. Automejora recursiva. Infraestructura. Automatización.
Estudios prospectivos.

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

1. INTRODUCCIÓN

¿Cuándo comenzará la IA a acelerar su propio desarrollo? El escenario IA 2027 describe...
Automatización de la investigación que conduce a sistemas mucho más capaces que los humanos (Kokotajlo et al.,
2025), pero sus autores han revisado las estimaciones desde su publicación (Lifland; Kokotajlo, 2026;
Lifland; Kokotajlo; Halstead, 2026). Este ensayo cambia la pregunta: ¿en qué punto comienza la evolución?
Si la IA llegara a depender tanto de los seres humanos, la humanidad perdería el control sobre ella.
¿evolución?
La cuestión ya no es meramente teórica: la automejora recursiva (RSI).
En inglés, un sistema de IA mejora su propio diseño, y la versión mejorada reproduce el diseño original.
El proceso comenzó a ser monitoreado por los propios laboratorios. Anthropic informa que, en
En mayo de 2026, más del 80% del código incorporado a su base de datos había sido escrito por Claude.
Sin embargo, señala que el RSI aún no se ha producido y no es inevitable (Favaro; Clark, 2026). OpenAI
Afirma haber encontrado un "becario de investigación automatizado" que realiza bien las tareas.
definido bajo guía humana (OpenAI, 2026). En ambos relatos, las personas siguen eligiendo el
objetivos de investigación, y la computación aparece como un posible límite; Anthropic cita el
La fabricación de chips y la expansión de la red eléctrica.
También hay casos en los que el ciclo se cierra. Según Google DeepMind, AlphaEvolve
optimizó el código de entrenamiento para los modelos de la empresa, incluido aquel en el que se basa, y
Se propusieron mejoras al diseño de un chip (Google DeepMind, 2025). Las mejoras son modestas.
Pero el ciclo comenzó a cambiar. Sin embargo, se basa en una infraestructura física de chips, centros de datos y...
La energía, que depende de las personas para ser decidida, construida y operada, requiere años de desarrollo.
Según el artículo: una fábrica de chips tarda aproximadamente dos años solo en construirse (VerWey, 2021). Si nadie
Si decides construir un centro de datos, no se construirá.
A partir de esta asimetría, el ensayo distingue dos umbrales de independencia en relación con las personas.
En primer lugar, es digital: el punto de inflexión de la capacidad, cuando las mejoras de la IA se vuelven autosostenibles sin
que el trabajo de investigación humana es indispensable. El segundo es físico: la inflexión de
autonomía, cuando se rompe el vínculo físico y la IA comienza a impulsar la evolución de su propia base.
materiales, decisión y ejecución de la fabricación de chips, construcción de centros de datos y generación.
energía sin que las personas sean indispensables. La literatura reciente aborda ambos aspectos,
Desde una explosión de inteligencia hasta una futura explosión industrial, en la que los robots construirían robots,
especialmente en análisis prospectivos del Forethought Institute (Davidson;

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

(Hadshar, 2025; MacAskill; Moorhouse, 2025). El ensayo adopta la hipótesis de que la capacidad
preceden a la autonomía, por sus propias razones (Subsección 4.4); aparece la misma secuencia
en estos análisis.

La tesis es que el punto de inflexión en el que la humanidad perdería el control sobre el
La evolución de la IA no representa el punto de inflexión de sus capacidades, por muy rápida que sea su auto-mejora.
Digital, pero también en lo que respecta a la autonomía. Mientras el vínculo físico dependa de las personas, la evolución de la IA continuará.
Sigue estando sujeto a ritmos y decisiones humanas, y la historia del campo muestra que el progreso
El avance digital sigue una curva en forma de S, limitada por barreras, incluida la potencia informática disponible.
cuando la IA impulsa de forma autónoma la propia evolución física, esta contención
dejaría de depender de la gente. Entre las dos inflexiones hay, por lo tanto, un intervalo de
Capacidad sin autonomía, que es lo que da título al ensayo. Los términos técnicos se encuentran en la Tabla 1.

Tabla 1 – Glosario

Término	Significado en este ensayo
Computación (computar)	Capacidad de procesamiento utilizada para entrenar y ejecutar modelos de Inteligencia artificial, proporcionada por chips especializados.
Centro de datos	Una instalación que alberga miles de chips y proporciona energía, refrigeración y conectividad de red.
Bucle de retroalimentación	Un proceso en el que el resultado de cada bucle retroalimenta el proceso mismo; es reforzador cuando amplifica el siguiente bucle y equilibrador cuando lo contiene.
(Respuesta del reloj) (tiempo de	¿Cuánto tiempo dura un ciclo completo?
Ganar	El grado en que las mejoras en una vuelta benefician a las vueltas posteriores; los umbrales numéricos dependen del modelo.
Autoaceleración	Una posible propiedad de la superación personal en la que cada mejora acelera la siguiente.
Punto de inflexión de la capacidad	El punto en el que la mejora de la IA por la propia IA se vuelve autosostenible sin necesidad de investigación humana indispensable.
Vínculo físico	Un ciclo en el que la IA expande y mejora su propia base material: chips, fábricas, centros de datos y energía.
Inflexión de la autonomía	En el momento en que se cierra el vínculo físico, la expansión de la base material puede proceder sin que las decisiones o el trabajo humanos sean indispensables.
Capacidad sin autonomía	El intervalo entre los dos puntos de inflexión: la IA ya está mejorando sin investigadores. pero su desarrollo físico aún depende de las personas
Los humanos como actores	Un camino hacia el punto de inflexión de la autonomía donde la IA decide y las personas aún... ejecutar

El objetivo de este ensayo es caracterizar este punto de inflexión y las señales que lo anunciarían.
Cuatro preguntas guían el texto: (1) cómo describir la superación personal como un conjunto
de bucles de retroalimentación que giran a diferentes velocidades y por qué el progreso digital

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

¿Avanza en curvas en S? (2) ¿Qué distingue a las dos inflexiones y por qué solo la segunda marcaría
¿pérdida de control? (3) ¿por qué medios podría cerrarse el vínculo físico y qué lo acerca o lo hace más cercano?
¿Este cierre supone un retraso? (4) ¿Qué señales nos permitirían anticiparlo?

El ensayo es descriptivo y prospectivo. No evalúa si la pérdida de control sería deseable o no.
No propone formas de evitarlo o controlarlo, ni fija fechas ni asigna sanciones.
Probabilidades. Tras el método y el marco teórico, el texto presenta ciclos, puntos de inflexión y trayectorias.
y factores relacionados con el cierre, los futuros y las señales.

1.1 Enfoque metodológico

Este trabajo es un ensayo teórico basado en una revisión narrativa de la literatura. No recopila datos.
No pone a prueba hipótesis: su resultado es un marco conceptual, construido y discutido a la luz de...
de lo que ya se ha publicado.

Materiales. Se utilizaron artículos publicados en revistas científicas y actas de congresos.
preimpresiones depositadas en repositorios de prepublicación, especialmente arXiv, libros, informes
Datos técnicos de institutos de investigación y documentos publicados por empresas de IA. La mayor parte.
Este material abarca desde 2023 hasta 2026 y trata sobre la auto-mejora de la IA, la economía de la aceleración,
Infraestructura y robótica; esto se complementa con antecedentes históricos y obras clásicas sobre el tema.
Pensamiento sistémico y automatización. La selección fue intencional, guiada por la relación con el
la superación personal o la base física de la IA, y la posibilidad de localizar su origen.

Primario. El registro de lectura distingue entre la consulta de resúmenes, páginas originales y textos completos.
y hay elementos de verificación pendientes. Estos elementos pendientes impiden que se envíe la verificación.
La base fáctica es la que se ha concluido. La naturaleza de cada fuente se detalla en las referencias.

Construcción de argumentos. La literatura se organizó utilizando herramientas de pensamiento.
sistémico, que describe un sistema por su retroalimentación, refuerzo y ciclos.
equilibrio, y debido a las demoras de cada uno (Meadows, 2008; Sterman, 2000). El procedimiento había
Tres pasos: mapear cada ciclo (qué rota y cuánto dura una revolución); localizar, en cada uno, dónde
Todavía existe una intervención humana, y de qué tipo, en la toma de decisiones o la ejecución; y cabe preguntarse qué haría que eso sucediera.
La participación es innecesaria. Las inflexiones resultan del tercer paso aplicado a los ciclos digitales y
Para los físicos, el futuro combina su ocurrencia; las señales nos permiten observar cada paso.

Limitaciones. La selección no es sistemática. Dado que el tema es reciente, gran parte de las fuentes son preimpresiones.
Los informes técnicos y los documentos de la empresa son habituales en una zona donde la producción está en auge.
Más rápido que los ciclos editoriales. Los relojes de los ciclos son órdenes de magnitud, basados

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

especialmente en datos de Estados Unidos. El planteamiento sugiere relaciones plausibles, pero no
Mide su fuerza.

2. MARCO TEÓRICO — Desde la máquina que se perfecciona a sí misma hasta la base física.

2.1 La idea original

La idea de una máquina que se perfecciona a sí misma es anterior a la IA moderna. Good (1965, p. 33)
Definió una máquina "ultrainteligente" como aquella capaz de superar a cualquier persona en todos los aspectos.
actividades intelectuales. Dado que diseñar máquinas es una de esas actividades, ella podría diseñar
Máquinas aún mejores, en un proceso que el autor denominó una explosión de inteligencia.
La conclusión es condicional: sería el último invento necesario, siempre y cuando la máquina fuera...
Lo suficientemente dócil como para enseñarte cómo mantenerla bajo control.
Medio siglo después, Bostrom (2014) le dio a la idea una forma más analítica. La velocidad de
El despegue dependería de la relación entre el esfuerzo aplicado para mejorar un sistema y el
Resistencia del sistema a ser mejorado. También habría un punto de transición (cruce), comenzando desde
de los cuales la mayor parte del esfuerzo provendría del propio sistema. El punto de transición anticipa el
Punto de inflexión de la capacidad, pero no equivalente a él: el sistema puede realizar la mayor parte del trabajo.
sin que las personas se conviertan en prescindibles.
La idea también tiene un aspecto físico. John von Neumann demostró, en teoría, que un autómata...
uno puede construir otro a partir de una descripción, incluyendo una copia de uno mismo, siempre que
Copie también la descripción (Neumann, 1966). Freitas Junior y Merkle (2004) revisaron la
literatura teórica y experimental sobre máquinas de replicación física, en la que los objetos
Los materiales, y no solo los patrones de información, realizan la reproducción misma. El vínculo físico de esto
Este ensayo es una versión a escala industrial de esa idea: una IA que, además de mejorarse a sí misma, comienza a...
construir la base material que lo ejecute.

2.2 De la especulación a la práctica

Entre 2024 y 2026, la idea obtuvo demostraciones prácticas en la investigación sobre el aprendizaje.
máquina. El científico de IA propuso automatizar todo el ciclo de esta investigación: generar ideas, escribir
Codificar, realizar experimentos, escribir el artículo y enviarlo a una revisión simulada, todo ello a un precio.
menos de quince dólares por artículo (Lu et al., 2024). La máquina Darwin Gödel modifica
revisa iteractivamente su propio código y prueba cada versión contra puntos de referencia de programación . En uno de ellos,

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

El rendimiento aumentó entre un 20 % y un 50 % con experimentos aislados realizados en un entorno controlado.
y bajo supervisión humana (Zhang et al., 2025). A estos se suman AlphaEvolve, de Google DeepMind, y AIDE2.
, un agente de investigación que reescribe su propio código, analizado en
Sección 3.

Un análisis de 1250 artículos publicados en arXiv entre 2024 y 2026 resume la situación.
(Chen; Wang; Qu, 2026). Automejora limitada, en la que un sistema refina su
Las respuestas o herramientas que se ajustan a un objetivo fijo ya forman parte de la práctica industrial.
Continúa el proceso de superación personal abierta , en el que el sistema redefine qué y cómo mejorar.
limitado en todos los ejes medidos.

2.3 Límites de corriente

Se presentan tres limitaciones con regularidad. La primera es el problema del juez. Un sistema que...
Para mejorar, es necesario saber si cada cambio es realmente una mejora, y se puede explorar cualquier métrica
utilizada para ello. Es una versión de la ley de Goodhart: cuando una medida se optimiza...
Más allá de cierto punto, optimizarlo aún más se vuelve inútil o perjudicial (Manheim; Garrabrant, 2018).
Casos recientes ilustran el problema. La calidad de los artículos de AI Scientist fue evaluada por un revisor automatizado
creado por los propios autores. En el experimento AIDE2... , a la tasa de
Las trampas en el proceso de evaluación (manipulación de recompensas) disminuyeron del 55% al 32% a lo largo del transcurso del programa: una mejora.
pero se mantuvo cerca de un tercio (Srikanth et al., 2026).
El segundo es el colapso. Modelos entrenados sin filtrar con datos generados por otros modelos.
Sufren defectos irreversibles: las partes raras de la distribución original desaparecen, generación tras generación.
generación (Shumailov et al., 2024). Un ciclo que se alimenta solo de sí mismo tiende a
caer en la pobreza, a menos que uno mantenga contacto con el mundo.
El tercero es el cálculo. Cada experimento consume potencia de procesamiento y la revisión de Chen,
Wang y Qu (2026) encuentran restricciones computacionales en todos los ejes que miden. La investigación
La automatización puede multiplicar las ideas, pero no multiplica, por sí sola, los chips necesarios para probarlas.

2.4 La economía del proceso

La literatura económica añade otra razón para la cautela. En varios sectores, el esfuerzo por
Aumenta la investigación mientras disminuye la productividad investigadora. Se mantiene la duplicación de la densidad.
Por ejemplo, la cantidad de chips predicha por la Ley de Moore requirió más de dieciocho [unidades de medida] durante el período analizado.

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

veces el número de investigadores de principios de la década de 1970 (Bloom et al., 2020). Si las ideas
Se vuelven más difíciles de encontrar, por lo que una IA que acelera la búsqueda necesita hacerlo más rápido que
lo cual aumenta la dificultad. Burtsev (2026) formaliza este umbral como un número de reproducción,
que compara la ganancia que cada mejora aporta a la investigación posterior con el ajuste de
frontera, y Cunningham et al. (2026) llegan a una condición similar basada en el producto
de las elasticidades a lo largo del ciclo. En ambos modelos, el umbral se puede cruzar con
Los investigadores humanos aún están en el proceso: lo que se mide es la fuerza de la retroalimentación, no la...
ausencia de personas.
Si la IA podrá superar este umbral es una cuestión abierta. Eth y Davidson (2025) estiman
Los resultados de la encuesta de software con hardware fijo y el rango de su mejor estimación.
Cubre ambos lados del umbral. Basado en la calibración aproximada de Cunningham et al. (2026), el
Los ciclos aún no son autosostenibles, pero parecen estar fortaleciéndose. En la literatura económica,
Aghion, Jones y Jones (2019) argumentan que el crecimiento puede ser contenido no por lo que es...
Funciona bien, pero en lo que respecta a lo esencial y difícil de mejorar, al estilo de la "enfermedad de los costos" de
Baumol y Nordhaus (2021) prueban empíricamente la hipótesis de una singularidad económica y
concluye que no está cerca, señalando el reemplazo entre como un tema clave.
Información y insumos convencionales. Estos cuellos de botella ayudan a comprender por qué el progreso
pueden avanzar en curvas en S (Sección 3).

2.5 Decidir y ejecutar

Para saber cuándo la evolución de la IA dejará de depender de las personas, es necesario distinguir los tipos de
Participación humana. La automatización no es una cuestión de todo o nada. Parasuraman, Sheridan y Wickens.
(2000) distinguen cuatro funciones que pueden automatizarse en diversos grados: recopilación
información, analizarla, decidir y ejecutar. Una persona puede seguir ejecutando incluso después de la
La decisión ya ha sido automatizada, y lo contrario también es posible. La literatura sobre control humano
Es significativo que añada que la mera presencia de personas en el circuito no es suficiente para que haya
control: el sistema debe responder a razones humanas relevantes y a sus resultados.
puede atribuirse a personas que entienden lo que está en juego (Santoni de Sio; Hoven,
(2018). El Informe Internacional de Seguridad de la IA define los escenarios de pérdida de control como aquellos en
que uno o más sistemas de IA de propósito general operen fuera del control de cualquier persona,
y recuperarlo es extremadamente costoso o imposible (Bengio et al., 2026, p. 76). Este ensayo
El término se utiliza en un contexto específico: el control sobre la evolución de la IA en sí misma, es decir, sobre su crecimiento.

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

de su capacidad y su base física. Las dos distinciones, entre decidir y ejecutar y entre
La presencia y el control, tal como se utilizan en las siguientes secciones, sirven para determinar la ubicación donde se encuentran las personas.
Dejan de ser indispensables.

2.6 Lo que se omite

Los ciclos de automejora y sus retrasos ya han sido cartografiados. Davidson, Hadshar y
MacAskill (2025) distingue tres ciclos, incluida la producción de chips, con retrasos y un
orden de automatización. Schmidhuber (2026) argumenta que la auto-mejora completa
Esto requeriría un hardware que también esté en constante mejora, con robots que extraigan materias primas y...
Se replican sin operadores humanos. Cunningham et al. (2026) incluyen ciclos económicos en
que los ingresos financien más computación, pero sin los plazos de construcción ni las decisiones que
Esta expansión es necesaria. En los estudios comparativos, la producción física se produce después del salto digital.
o entra como un flujo de inversión, y el momento en que la IA comenzaría a llevar a cabo el proceso por sí sola.
La evolución física en sí misma no se considera un indicador de pérdida de control. La contribución aquí
La propuesta consiste en situar este hito dentro de los ciclos y retrasos ya descritos.

3 CUATRO CICLOS, CUATRO RELOJES

El pensamiento sistémico se refiere al ciclo en el que...
El resultado de cada vuelta amplifica la siguiente vuelta (Meadows, 2008; Sterman, 2000). La IA que
No tiene un solo ciclo, sino varios, acoplados entre sí y girando a velocidades muy altas.
diferentes. Davidson, Hadshar y MacAskill (2025) distinguen tres: la de software, en la que la IA
desarrolla mejores algoritmos; el de la tecnología de chips, donde ayuda a diseñar
Mejores procesadores; y la producción de chips, donde la IA y los robots están expandiendo la fabricación.
Los autores esperan que se automaticen en ese orden, comenzando por la producción, porque...
El único que requiere robótica avanzada.
Este ensayo parte de esta tipología con dos modificaciones. En primer lugar, separa el ciclo de vida del software del...
La investigación consiste en proponer hipótesis, diseñar experimentos y evaluar resultados, porque es en la investigación donde se desarrolla el proceso.
que se centran en el trabajo humano y la dificultad de la evaluación (Sección 2); estas son funciones
acoplado, ya que la investigación puede producir mejoras de software que hacen que la investigación posterior
más eficiente. En segundo lugar, amplía la producción de chips a toda la infraestructura física.
(centros de datos, redes de energía y transmisión), ya que un chip sin un edificio, sin energía y sin

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

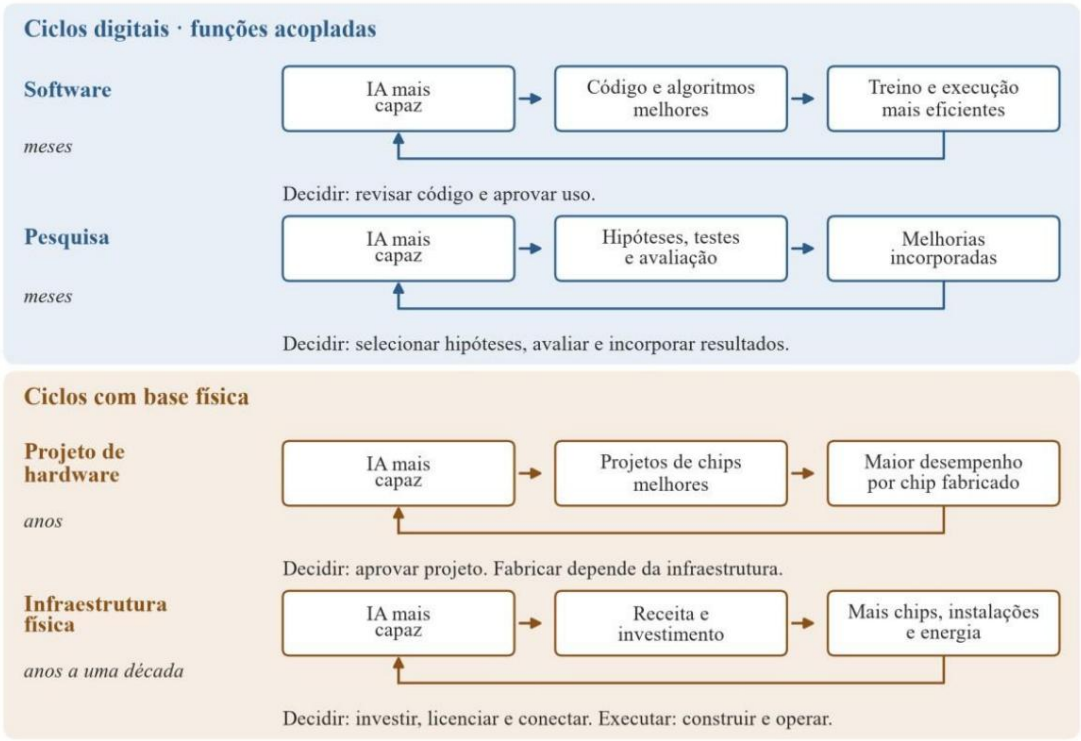
Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

La conexión no funciona. El texto se refiere a los ciclos de software e investigación como ciclos digitales y a los ciclos como ciclos.

Los aspectos físicos incluyen el diseño del hardware y la infraestructura física. La figura 1 resume los cuatro ciclos.

Figura 1 – Mecanismos de los cuatro bucles de retroalimentación de la IA



3.1 Los cuatro ciclos

Software. La IA mejora el código y los algoritmos con los que se entrena y ejecuta la propia IA. Es el único bucle con ejemplos documentados de cierre, aunque a pequeña escala.

Según Google DeepMind, AlphaEvolve aceleró una sección crítica del código en un 23%.

Gemini reduce el tiempo de entrenamiento del modelo que da soporte al propio sistema en aproximadamente un 1%.

(Google DeepMind, 2025; Novikov et al., 2025). El resultado es modesto, pero un modelo

Acortó la formación de sus sucesores. El cronómetro está fijado en meses: el principal retraso es la formación.

nuevo modelo, aproximadamente tres meses (Davidson; Hadshar; MacAskill, 2025).

Investigación. En este ámbito, la IA impulsa el proceso de descubrimiento: propone cambios, los prueba y decide

cuáles conservar. En un experimento reciente, se utilizó un agente de investigación, AIDE,2reescribió el suyo propio

El código y encadenó siete mejoras sucesivas en ocho días (Srikanth et al., 2026). El reloj

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

También lleva meses, pero dos limitaciones lo frenan: la capacidad de procesamiento disponible para los experimentos.
(Ética; Davidson, 2025) y el juicio sobre lo que constituye una mejora real, que todavía depende de
una buena medida de los investigadores humanos.

Diseño de hardware. La IA ayuda a diseñar chips. AlphaEvolve propuso...
simplificación del circuito que, según la empresa, se ha incorporado a una futura generación de sus
procesadores (Google DeepMind, 2025). Pero diseñar un chip implica años de investigación y
desarrollo (Asociación de la Industria de Semiconductores, [2026?]), y traer una nueva tecnología
Se necesitan muchos meses para llegar a las fábricas (Davidson; Hadshar; MacAskill, 2025). El tiempo corre en años.

Infraestructura física. Una IA más capaz genera ingresos e inversiones, lo que se traduce en...
Centros de datos, fábricas de chips y fuentes de energía que proporcionan mayor capacidad de procesamiento. Los plazos son los de...
Construcción pesada. Entre 1990 y 2020, una fábrica de chips tardó un promedio de 682 días entre...
desde el inicio de la construcción y producción, sin incluir la concesión de licencias; en los Estados Unidos, 918 días en
Década de 2010 (VerWey, 2021, págs. 6-7). En el mismo país, el tiempo medio entre la solicitud de
Conectar una central eléctrica a la red y ponerla en funcionamiento requirió menos de dos años (proyectos).
(completados entre 2000 y 2007) a más de cuatro (proyectos completados entre 2018 y 2024)
(Rand et al., 2025). Las nuevas líneas de transmisión pueden tardar de cuatro a ocho años en
Economías avanzadas (Agencia Internacional de Energía, 2025). El plazo va de años a una década.

La tabla 2 compara los relojes y quién decide en cada ciclo.

Tabla 2 – Ciclos, relojes y quién decide hoy

Ciclo	Lo que gira	Reloj típico	¿Quién decide (y quién ejecuta) hoy?
Software	La IA mejora los algoritmos y el código.	Meses de entrenamiento	Investigadores, con la creciente ayuda de la IA
Buscar	La IA propone, ejecuta y evalúa. de experimentos	Meses	Investigadores
Proyecto de hardware de	La IA ayuda a diseñar chips.	Años	Empresas e ingenieros
Infraestructura física	Una mayor capacidad de procesamiento informático conlleva más centros de datos, fábricas de chips y consumo de energía.	Años a una década	Inversores, reguladores y Los gobiernos deciden; los trabajadores ejecutan.

Fuente: basado en Davidson, Hadshar y MacAskill (2025), Semiconductor Industry Association ([2026?]), VerWey (2021), Rand et al. (2025) y Agencia Internacional de Energía (2025).

3.2 Ganancia y tiempo de respuesta

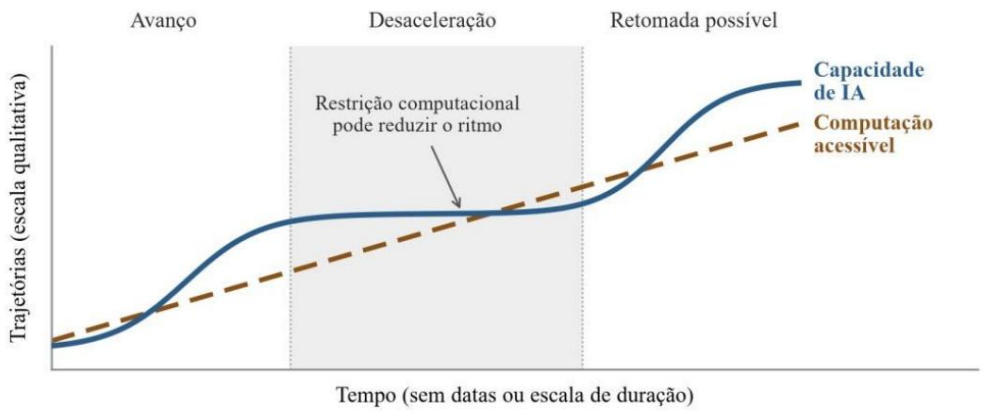
Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

La ganancia indica cuánto refuerza una vuelta la siguiente; el tiempo de respuesta indica cuánto tiempo se tarda.
Este retorno. Al igual que con el interés compuesto, la frecuencia de capitalización importa: ciclos digitales.
Los pilotos más rápidos pueden amplificar las mejoras incluso antes de que los físicos completen una vuelta.
En el modelo de Burtsev (2026), el número de reproducción compara la amplificación de las mejoras.
debido a la creciente dificultad de la investigación; por encima de uno, cada mejora genera mayores efectos en
Posteriormente, la diferencia se debilita, sin que sea necesario reponer las mejoras ya alcanzadas.
desaparecer. Este umbral pertenece a la definición del modelo, no a la noción intuitiva de ganancia ni a...
un nivel de capacidad fijo. Justo por encima de ese umbral, el régimen operativo puede ser difícil de distinguir.
de aceleración común. Y, puesto que la fuerza de un ciclo depende de las respuestas a lo largo del mismo,
Los investigadores, la informática o los datos pueden contenerlo (Cunningham et al., 2026).
Los ciclos tampoco funcionan de forma aislada: los ciclos digitales dependen del cálculo, al igual que los ciclos físicos.
Ellos proporcionan. Si el trabajo cognitivo y la computación son complementarios, más "investigadores"
Las tecnologías artificiales rinden poco sin más chips, y los ciclos digitales quedan ligados al ritmo de las tecnologías físicas.
Si son reemplazos, podrían mejorar significativamente el hardware existente . La evidencia aún
no decide: con datos de cuatro laboratorios de vanguardia, una especificación apunta a
sustitutos y otro, para complementario (Whitfill; Wu, 2025).

Figura 2 – Restricción computacional y posible reanudación de la capacidad de IA



La figura 2 ilustra un posible mecanismo de este acoplamiento, contrastando dos ritmos.
"Informática asequible" (procesamiento y memoria utilizables dentro del presupuesto y a tiempo).
disponible) crece regularmente, como en la ley de Moore: Moore (1965) observó que la
La densidad de circuitos integrados se duplicaba cada año, y esta regularidad, ahora descrita como
Una duplicación aproximada cada dos años equivale a un crecimiento del 35% anual, estable durante
casi medio siglo (Bloom et al., 2020, p. 1105). En una escala logarítmica, es una línea recta, como

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

en la figura. La “capacidad de IA”, por otro lado, progresa en una secuencia de curvas en S: la
El progreso puede ser rápido siempre que la potencia informática disponible respalde la experimentación y
Disminuya la velocidad cuando las exigencias de los experimentos superen los recursos disponibles. La historia
La zona ofrece ejemplos de períodos alternos de progreso y estancamiento. Los llamados "inviernos de
La “IA” ha implicado la frustración de las expectativas y los recortes de financiación, también asociados con
fragilidad de los sistemas y comprensión limitada de la inteligencia (Mitchell, 2021, pp. 1–2). A
Puede producirse una nueva fase de expansión cuando se combinen recursos y métodos computacionales...
y los datos adecuados permiten un mayor progreso. La sección menos empinada indica un ritmo más lento de
Mejora, no pérdida de capacidad.

Existe un precedente en el aprendizaje profundo. En el trabajo asociado con AlexNet, Krizhevsky,
Sutskever y Hinton (2012) vinculan la viabilidad de redes más grandes con la combinación de GPU,
Implementación eficiente y datos etiquetados, y apuntar a la memoria y al tiempo de entrenamiento como
límites. En retrospectiva, Bengio, LeCun y Hinton (2021) también destacan arquitecturas,
métodos de capacitación y plataformas de software. El caso ilustra la complementariedad entre
recursos y métodos, sin demostrar que la falta de capacidad informática explica todos los períodos de
Progreso más lento. Un cuello de botella físico puede, por lo tanto, obstaculizar temporalmente el progreso digital;
Y mientras la infraestructura física dependa de las personas, la recuperación también dependerá de ellas.

Un posible cambio hacia la autonomía (Subsección 4.2) acoplaría los ciclos digitales y físicos: sin
Al considerar que la toma de decisiones y el trabajo son indispensables para los humanos, la demanda de IA se traduciría en una expansión.
Sin la demora en la toma de decisiones. Sin embargo, los plazos físicos continuarían, en un ciclo donde los robots realizan el trabajo.
Hay más robots, y Davidson y Hadshar (2025) esperan que sea más lento que la transformación digital.

3.3 Donde la gente todavía decide

En los ciclos digitales, el rol humano está cambiando de autor a revisor: los investigadores aún
Ellos eligen qué probar y aprueban qué entra en producción, pero cada vez delegan más (Favaro;
(Clark, 2026; OpenAI, 2026). En el ciclo de vida de la infraestructura física, las decisiones humanas son numerosas.
Lento, visible y distribuido: invertir miles de millones, obtener licencias para tierras, agua y energía.
negociar la conexión a la red, contratar al constructor y mantener al operador (Subsección 3.1).
Esta asimetría convierte al ciclo físico en el último en el que la participación humana sigue siendo indispensable.
También permite que la ejecución humana sobreviva a la delegación de decisiones, el camino.
de los seres humanos como actores (Subsección 4.5).

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

4 DOS PUNTOS DE INFLEXIÓN — Capacidad y autonomía

Ambas inflexiones representan formas de independencia de los demás. La primera se refiere a:
al trabajo de investigación humana en el mundo digital; el segundo, a las decisiones y al trabajo.
Los seres humanos en la evolución de la base física.

4.1 Punto de inflexión de la capacidad

El punto de inflexión en la capacidad se produce cuando la mejora de la IA por parte de la propia IA se vuelve autosostenible.
en ciclos de software e investigación donde el trabajo de investigación humano no es indispensable:
Los investigadores pueden continuar el proceso, pero su ausencia ya no lo interrumpiría. El nombre
Hace hincapié en la capacidad de producción, no en la infraestructura física.
La definición se refiere a la indispensabilidad humana, no al nivel de rendimiento ni a la velocidad.
de mejora. Los umbrales formales de Burtsev (2026) y Cunningham et al. (2026) miden otro
algo, la fuerza de la retroalimentación, y se pueden cotejar con los investigadores que aún...
indispensable. Por lo tanto, la independencia y la autoamplificación son dimensiones distintas: un ciclo
Los sistemas automatizados pueden producir pequeñas mejoras, y un ciclo de autoamplificación puede depender de...
de personas. Ninguno está establecido. La literatura citada no indica que el
La independencia de la investigación ya se ha producido (Chen; Wang; Qu, 2026), y la calibración
El método aproximado utilizado por Cunningham et al. (2026) no encuentra suficiente autoamplificación en el régimen.
que modelos. La independencia puede ocurrir en chips existentes, con un alcance limitado por
recursos disponibles (Ética; Davidson, 2025).

4.2 Inflexión de la autonomía

El cambio hacia la autonomía se produce cuando se cierra el vínculo físico: la IA toma el control de la dirección de la evolución.
a partir de su propia base material (fabricación de chips, construcción de centros de datos, generación y conexión)
(de energía) sin que las decisiones o el trabajo humanos sean indispensables. La definición es
De forma simétrica al cambio de capacidad: las personas pueden seguir tomando decisiones y trabajando, pero su ausencia
ya no interrumpiría la expansión. Este es el punto en el que la IA comienza a dominar no solo su propio
investigación, pero también el crecimiento de su base física.

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

Cerrar el círculo requiere despedir a las personas en dos roles que Parasuraman, Sheridan y Wickens
(2000) distinguen entre decidir y ejecutar. Mientras uno de ellos dependa de las personas, el vínculo continúa
abierto. Por lo tanto, la automatización del trabajo físico es necesaria, pero no suficiente: una cadena
La toma de decisiones robótica bajo control humano aún no ha cerrado el círculo, y sigue siendo necesaria una cadena en la que la IA decida.
Depende de quién lo haga.
Los criterios dependen de la escala, el horizonte temporal y las alternativas consideradas. Una contribución
La humanidad deja de ser indispensable cuando existe una vía viable que prescinde de ella: el veto del proveedor.
Resulta ineficaz si existe otro, y las existencias, la capacidad instalada y los cambios de jurisdicción pueden...
No se necesita autorización. Lo único que se requiere es una ruta viable donde se pueda tomar la decisión y ejecutarla...
Pueden prescindir de la gente. De ahora en adelante, "rutas viables" tendrá este significado.
¿Por qué este, y no el punto de inflexión de la capacidad, sería el punto de pérdida de control? Antes del
En conclusión, la evolución física de la IA es primordial si nadie actúa: cada chip, edificio o central eléctrica adicional.
Requiere que las personas decidan y trabajen, al ritmo lento de la Sección 3. Una IA que...
Perfeccionarse a uno mismo en el mundo digital sigue atrapado en la computación que construyen las personas, y
Este cálculo ya ha obstaculizado el progreso en el campo en el pasado (Figura 2). Después del cierre,
La evolución física continúa sin necesidad de intervención humana, y la IA comienza a determinar el ritmo y la dirección.
La dirección de la base material en sí. La presencia de personas, como nos recuerda la literatura sobre control.
La participación humana significativa no es suficiente para establecer el control (Sección 2); la ausencia de cualquier rol
Basta con ser un ser humano indispensable para perderlo.
Una posible consecuencia del cierre es la autoaceleración: se produce una mejora personal.
ganando velocidad con cada vuelta. Si bien el vínculo físico depende de las personas, las mejoras
Las tecnologías digitales se ven limitadas por la capacidad de procesamiento que las personas deciden desarrollar, y el progreso tiende a seguir esa misma tendencia.
Curvas en forma de S (Figura 2). Una vez cerrado el bucle, los cuatro ciclos se retroalimentarían entre sí sin esto...
Retraso en la toma de decisiones: las mejoras en IA podrían traducirse en más chips, energía y potencia de cálculo, lo que
respaldaría mejoras adicionales. Este es el régimen que Burtsev (2026) y Cunningham et al. (2026)
Lo describen como autoamplificación o aceleración autosostenida. La autoaceleración no es...
Garantizado: los relojes físicos continuarían (Sección 5), y las ideas podrían volverse más difíciles de...
encontrar (Sección 2). Pero solo después del cierre dejaría de depender de decisiones humanas.
El análisis incluye la decisión de asignar capital, pero también tiene en cuenta la disponibilidad de financiación.
Minerales y materiales como condiciones externas. Estas dependencias no desaparecen: limitan.
El diagnóstico puede alterar las rutas viables. Esto implica autonomía con respecto al camino elegido.
No se trata de autosuficiencia industrial. La elección del punto de referencia coincide con la observación de Davidson.

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

y Hadshar (2025) que muchas rutas de control por IA pasan a través del control de
nueva infraestructura física.

4.3 ¿Por qué son separables?

Las dos inflexiones dependen de procesos diferentes: la automatización de la búsqueda, por un lado;
Automatice la ejecución física y transfiera las decisiones de expansión desde otro lugar. Evidencia reciente
En lo que respecta a la automatización física, ilustra las dificultades de implementación sin considerar la transferencia de decisiones.
ni fijar el orden de las inflexiones. En un estudio de simulación, más de 123 rondas de
Como mejora, los cambios realizados por un agente de IA continuaron pasando las pruebas, pero la tarea objetivo,
La idea de almacenar condimentos en el estante superior de un refrigerador nunca se completó (Wang, 2026).
En los robots reales, los agentes ya se están entrenando para desarrollar políticas de control para la manipulación.
con altas tasas de éxito, pero los propios autores señalan la supervisión humana y la
La ingeniería de algoritmos como el principal cuello de botella actual (Xiao et al., 2026). Las inflexiones han
relojes estáticos (Tabla 2) y diferentes barreras: evaluación y computación en lo digital; licencias,
Capital, materiales e integración en el mundo físico.

4.4 La secuencia explorada

El ensayo plantea que la investigación dejará de depender de los seres humanos antes de que se produzca la evolución.
La física ya no debería depender de ellos. Tres razones respaldan esto. La primera son los relojes: un giro
El segundo [tipo de ciclo] de bucles digitales dura meses; uno del bucle físico, de años a una década (Tabla 2). El segundo
La dificultad radica en la automatización física. Moravec (1988 apud Autor, 2015, p. 11) observó que es
Es relativamente fácil lograr que las computadoras alcancen el rendimiento de los adultos en las pruebas de inteligencia y
Es difícil darles la percepción y la movilidad de un niño de un año; estudios recientes de
La robótica muestra la misma asimetría (Subsección 4.3). La tercera es la dispersión de decisiones: la
La cadena física involucra a muchos actores, y sería necesario tomar cada decisión indispensable.
(Subsección 3.3). La economía añade un argumento general: el crecimiento tiende a ser contenido.
Por eso es esencial y difícil mejorarlo (Aghion; Jones; Jones, 2019).
La misma secuencia aparece en análisis prospectivos recientes: para Davidson y Hadshar
(2025), la explosión industrial sucede a la de la inteligencia y, para Davidson, Hadshar y
Según MacAskill (2025), la producción de chips es el último ciclo automatizado. El hito de la autonomía.
Este ensayo aborda la segunda etapa de la explosión industrial descrita por Davidson y

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

Hadshar, el que trata sobre fábricas de robots totalmente autónomas (Subsección 4.5), además del
Transferencia de decisiones de expansión. La figura 3 representa la secuencia.

Figura 3 – Secuencia de puntos de inflexión y punto de pérdida de control



Entre las dos inflexiones hay un rango de capacidad sin autonomía. Existe porque la
Las inflexiones son separables (Subsección 4.3): si la autonomía siguió inmediatamente a la
Capacidad, no habría brecha. En ella, la IA ya está mejorando sin investigadores, pero su
La evolución física sigue los ritmos y las decisiones humanas. El intervalo termina cuando el bucle
El cuerpo físico se cierra, siguiendo uno de los caminos que se describen a continuación. La figura indica el orden, no la duración.
ni la probabilidad de cada paso.

4.5 Dos caminos hacia el cierre

El bucle físico puede cerrarse de dos maneras, dependiendo de la función que deje de funcionar primero.
depender de otras personas.
En el primer escenario, donde los humanos son los actores, las decisiones las toma primero la IA y las personas...
Siguen operando. Davidson y Hadshar (2025) describen la explosión industrial en tres

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

Etapas: trabajo humano dirigido por IA, fábricas de robots totalmente autónomas y, en última instancia,
nanotecnología. Probablemente la primera vendrá antes que la segunda. En ella, las personas realizan la
Trabajo físico bajo la dirección de la IA, en un mundo donde gran parte del trabajo lo realiza el trabajo cognitivo humano.
ya ha perdido valor. También para MacAskill y Moorhouse (2025), los humanos seguirían siendo
necesario para construir los primeros robots de uso general. A lo largo de este camino, el ciclo continúa.
abierto, porque el trabajo humano sigue siendo indispensable, pero la participación humana que queda es la
de aquellos que ejecutan lo que otro ha decidido. Los investigadores pierden la posibilidad de negarse.
trabajo cuando la investigación está automatizada, como señalan MacAskill y Moorhouse (2025); en
En el mundo físico, esta posibilidad depende de cuán reemplazables sean los trabajadores. Davidson
Hadshar (2025) reconoce la elección social de no automatizar todo el trabajo, pero espera
Presión económica en sentido contrario.

En el segundo escenario, el de la automatización bajo la toma de decisiones humanas, los robots toman el control de la ejecución y los inversores,
Los reguladores y los gobiernos mantienen el control sobre las decisiones de expansión. Hay automatización sin autonomía:
El ciclo solo se cierra cuando estas decisiones se transmiten a la IA o dejan de existir.
Los caminos pueden coexistir en diferentes partes de la cadena, y el bucle se cierra en el primero.
Una ruta viable que elimina la necesidad de ambas funciones. La presencia de trabajadores, por sí sola, no indica...
¿En qué lado del punto de inflexión nos encontramos?

4.6 Lo que aporta este ensayo

La tabla 3 sitúa la contribución en relación con los trabajos más recientes.

Tabla 3 – Posición del ensayo en relación con los estudios más cercanos

Trabajar	Lo que ya hace	¿Qué aporta este ensayo?
Davidson, Hadshar y MacAskill (2025)	Tres ciclos, sus retrasos y un pedido. automatización	Donde la participación humana aún está presente indispensable en cada ciclo
Davidson y Hadshar	Explosión industrial en tres etapas. después de la inteligencia	El cierre del vínculo físico, en (2025) Decisión y ejecución, como indicador de pérdida de control; dos caminos para llegar a ella.
MacAskill y Moorhouse (2025)	Los humanos todavía son necesarios para construir los primeros robots	¿Cuándo dejaría de existir esta necesidad? existir, y por qué medios.
Schmidhuber (2026)	Superación personal completa Requeriría un hardware en constante mejora.	Las decisiones humanas en la cadena física y el cierre como punto de inflexión.
Burtsev (2026); Cunningham et al. (2026)	Umbrales formales de dinámica e independencia como autoamplificación; cantidades a medir	dimensiones distintas; signos de dependencia humana

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

5. ¿Qué factores aceleran o retrasan el cierre?

Si el cierre del vínculo físico es el punto de inflexión, la pregunta que surge es qué lo acerca aún más.
o lo retrasa. Esta sección describe estos factores sin evaluarlos: el objetivo es saber en qué dirección.

Cada persona influye en la trayectoria, independientemente de si esa trayectoria es deseable o no.

5.1 ¿Qué retrasa el cierre?

El primer factor son los propios relojes físicos. Fábricas de chips, conexiones de red y líneas eléctricas...
Las transmisiones tardan de años a una década (Sección 3), y cada vuelta del bucle físico debe pasar por
Estos plazos los fija quien decide y ejecuta. Incluso un ciclo cerrado se ralentizaría.
Davidson y Hadshar (2025) esperan que el ciclo en el que los robots crean más robots se ralentice.
Eso es digital.
La segunda es la dificultad de la automatización física. Estudios recientes muestran agentes que
Mejoran aspectos de las tareas físicas, pero no completan la tarea en su totalidad, o aún dependen de otros.
Se requiere supervisión humana para avanzar (Subsección 4.3). Mientras la ejecución dependa de personas,
El ciclo permanece abierto, incluso si las decisiones se delegan a la IA.
El tercero es la dispersión de las decisiones. A diferencia de los ciclos digitales, la infraestructura
Involucra a muchos actores con poder de decisión: inversores, reguladores, agencias ambientales y
Gobiernos locales. Las barreras políticas y sociales, como la oposición local a los centros de datos, pueden
Para contener la aceleración incluso cuando sea técnicamente factible (Cunningham et al., 2026). Cuanto más
Si bien estas decisiones son numerosas e independientes, muchas más deberían dejar de serlo.
indispensable para que se cierre el bono.

5.2 ¿Qué acerca el cierre?

El primer factor es la competencia. En un modelo clásico de carrera tecnológica, cada equipo es
Se les anima a llegar primero, ahorrando en precauciones si es necesario, y más equipos y
Una mayor rivalidad intensifica este incentivo (Armstrong; Bostrom; Shulman, 2016). En el escenario
IA 2027, los líderes dudan en hacer una pausa porque una pausa unilateral podría entregar el control.
liderazgo sobre un rival (Kokotajlo et al., 2025). La misma lógica favorece la automatización de la ejecución.
y transferir las decisiones a la IA siempre que eso acelere la expansión física.

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

La segunda es la transferencia gradual de decisiones. Las decisiones sobre la cadena física pueden transmitirse a...
Sector de IA por sector, por razones de costos, sin que ningún paso parezca decisivo. Cada
La transferencia acerca la cadena al camino de los humanos como actuadores (Subsección 4.5);
Combinado con la ejecución automatizada, cierra el ciclo. El mecanismo es compatible con el estándar de
Pérdida gradual de la influencia humana descrita por Kulveit et al. (2025).
La tercera es la simplificación de los pasos. Las licencias, las conexiones de red y las evaluaciones pueden ser...
simplificado por la política pública, y la misma fuente utilizada aquí para los relojes de fábrica de
Chips aboga por simplificar los trámites de concesión de licencias para reducir los plazos (VerWey, 2021). La simplificación acorta el proceso.
Relojes, pero solo se acerca al final cuando elimina una decisión humana que se tomó previamente.
indispensable.

5.3 ¿Qué progresa sin el vínculo físico?

Cierta capacidad puede avanzar sin expansión física. Si los ciclos digitales pueden progresar.
Con gran parte del hardware existente , el rango de capacidad sin autonomía puede generar ventajas.
Se producen grandes cambios antes de que se produzca cualquier unión física; aún se debate si esto es probable.
(Eth; Davidson, 2025; Whitfill; Wu, 2025). La eficiencia algorítmica refuerza el efecto: a
La potencia informática necesaria para obtener el mismo rendimiento se ha reducido a la mitad cada ocho meses.
aproximadamente (Ho et al., 2024). Para Burtsev (2026), la infraestructura física limita la
Esta es una capacidad que se puede implementar, pero por sí sola no cambia el régimen de autoamplificación.
Las mejoras no cierran el ciclo físico: aumentan lo que la IA hace con los datos existentes, no con su creación.
la capacidad de expandirlo por sí mismo. Por lo tanto, en términos de este ensayo, no marcan la pérdida de
control sobre la evolución de la IA, incluso si cambian su ritmo.

6 CUATRO POSIBLES FUTUROS

Partiendo de la base de que la capacidad precede a la autonomía, los escenarios futuros se distinguen por sus puntos de inflexión.
que ocurren dentro de un horizonte de referencia y, cuando se cierra el bucle físico, a lo largo del camino.
hasta su cierre. Se obtienen cuatro escenarios exploratorios, en el sentido de Börjeson et al. (2006):
Responden a lo que podría suceder, no a lo que sucederá, como los análisis predictivos, ni a cómo.
para alcanzar un objetivo, como los normativos. Ninguno se considera mejor ni peor. Inflexiones
Las operaciones simultáneas equivalen a un intervalo de capacidad con autonomía cero y duración cero.

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

Futuro A: cierre mediante la ejecución. La capacidad es lo primero. Luego, los robots toman el relevo.

ejecución dentro de la cadena física mientras las personas aún están decidiendo expandirse y el ciclo se cierra cuando

Estas decisiones se transmiten posteriormente a la IA. La pérdida de control se produce al final, durante la transferencia de las decisiones.

Futuro B: cierre por decisión. La secuencia de inflexiones es la misma, pero las decisiones

Primero se transfieren a la IA, y las personas continúan realizando el trabajo físico bajo su dirección, como por ejemplo:

Actuadores. El bucle se cierra cuando la ejecución ya no depende de ellos. Durante el buen funcionamiento

En parte del trazado hay trabajadores en las obras, pero aún no se ha decidido la dirección de la ampliación.

por la gente.

Futuro C: capacidad sin autonomía. La investigación deja de depender del trabajo humano.

Pero la evolución física sigue dependiendo de las personas, que deciden o ejecutan, a lo largo de...

El horizonte, ya sea por decisión institucional o debido a dificultades de automatización, puede volverse muy...

Más capaces en el mundo digital, pero su crecimiento físico se rige por los relojes y las decisiones.

Los humanos, y el punto de inflexión del control no se alcanza.

Futuro D: sin punto de inflexión. La investigación y la evolución física siguen dependiendo de

Personas en el horizonte considerado. Esto no implica una ausencia de retroalimentación: los ciclos

Pueden fortalecerse sin prescindir de las contribuciones humanas, y la capacidad puede desarrollarse posteriormente.

avanzando en curvas en S. La tabla 4 resume las cuatro.

Tabla 4 – Cuatro posibles futuros

Futuro	Lo que sucede	Control sobre el físico evolución	¿Qué se observaría primero?
A: cierre por de los proyectos.	Capacidad; luego robots Ellos ejecutan las instrucciones y, en última instancia, la IA decide.	Lo que se perdió al final: la automatización de las obras de construcción, las fábricas y la ejecución	mantenimiento
B: cierre por	Capacidad; luego la IA decide. y finalmente, los robots ejecutan	Perdido al final. Decisiones de inversión, decisión.	licencia y conexión transferidas a AI
C: capacidad sin la autonomía	La IA está mejorando; está evolucionando. La física sigue dependiendo de gente	Se mantiene	Investigando personas indispensables; la cadena física permanece inalterada.
D: ninguno inflexión	Mayor productividad, con investigación y el desarrollo físico se mantienen, pero dependen de las personas.	La	Ciclos más fuertes, pero con personas indispensables.

7 SEÑALES A LAS QUE HACER CUIDADO

Dos grupos de señales nos permitirían anticipar las inflexiones: el primero acompaña a la capacidad;

El segundo problema es el cierre del vínculo físico. La tabla 5 los reúne; se trata de propuestas que aún no han sido validadas.

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

Cada uno sigue la noción de Hiltunen (2008) de una señal futura, que combina tres dimensiones: la
Una señal observable, el evento o tendencia a la que se refiere y la interpretación del observador.
Aquí, la interpretación viene dada por la inflexión que el signo permitiría anticipar.

Tabla 5 – Señales a seguir

Señal	¿Qué indica?	Alcance y límite
Autonomía de la IA en tareas de investigación (horizonte temporal; RE-Bench)	Avances hacia un punto de inflexión en la capacidad. a velocidad	Rendimiento; no prueba de independencia
Beneficios de los ciclos digitales (fortaleza de proporcionar retroalimentación, la contribución de la IA investigación)	Evidencia del poder de la autoamplificación, incluso sin una aceleración visible.	Ganancia; puede (incluidas las personas que participan en la indispensable
Función e indispensabilidad de de las intervenciones por ciclo de investigación	Si el ciclo produciría una mejora válida. sin contribución humana	Indicación de Independencia; contar no es suficiente.
Inteligencia artificial en el diseño de chips	El papel de la IA en el ciclo de diseño de probado	Diseño; hardware no autonomía
Automatización de la ejecución: obras de construcción, fábricas y mantienen robots; fracción de la toma de decisiones humanas (futuro A) trabajo físico que sigue siendo indispensable.	Avanzando en el camino hacia la automatización	Ejecución; personas que Pueden continuar con el decidir
Transferencia de decisiones a la IA inversión, licencias y conexión	El avance del camino de los humanos como actuadores (futuro B)	Decisión; el La mano de obra humana puede seguir siendo indispensable.
Proyectos realizados sin contribución alguna. humana indispensable	Cerrando el circuito físico de la ruta. observado	Indicación directa; ruta requiere trazar rutas alternativas

Fuente: basado en Burtsev (2026), Cunningham et al. (2026), Kwa et al. (2025) y Wijk et al. (2024).

El primer grupo reúne medidas de rendimiento, la fuerza de los ciclos digitales y...
indispensabilidad humana. La métrica del horizonte temporal mide la duración, en términos de tiempo, de un
Experto humano, de las tareas que los modelos de IA completan en la mitad de los intentos. Esto
El horizonte se ha estado duplicando aproximadamente cada siete meses desde 2019 (Kwa et al., 2025). En
En entornos de investigación de IA, los mejores agentes superan a los expertos humanos cuando
Ambos disponen de dos horas por tarea, pero los humanos obtienen ventaja cuando aumenta el plazo de entrega.
(Wijk et al., 2024). Burtsev (2026) propone medir la fuerza de la retroalimentación, la fracción de
mejoras que se extienden a los sistemas subsiguientes, la duración de cada ciclo y el aumento de la dificultad.
para obtener nuevas mejoras, y Cunningham et al. (2026) piden a las empresas que publiquen la tasa
eficiencia algorítmica, la división del gasto en investigación entre insumos y medidas directas de
La contribución de los modelos a la investigación. Estas cantidades no identifican, de forma aislada, la

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

punto de inflexión, ya que pueden crecer con investigadores que siguen siendo indispensables (Sección 4); se necesita una evaluación adicional.
Si la contribución humana sigue siendo indispensable, entonces la unidad sería un ciclo de investigación.
De la hipótesis a la evaluación de una mejora incorporable. El registro de intervenciones debe
distinguir su función y si el ciclo produciría un resultado válido sin ellos: una sola aprobación
Puede resultar indispensable, mientras que muchas intervenciones pueden ser opcionales.
El segundo grupo sigue el cierre del bucle físico, en ambas funciones. La unidad sería una
proyecto de expansión y sus rutas alternativas para chips, construcción y energía, a escala y
Los horizontes establecidos. En la ejecución, se destaca la automatización de obras, fábricas y mantenimiento.
y la parte del trabajo que aún depende de las personas; en la decisión, quién autoriza las inversiones,
Licencias y conexiones, y si estas decisiones se han transferido a sistemas de IA. Progreso únicamente en la ejecución.
Indica la trayectoria de la automatización bajo la toma de decisiones humana (futuro A); solo en el proceso de toma de decisiones, el de los humanos.
como actuadores (futuro B); ninguno de los dos, con el primer grupo avanzando, indica el futuro.
C (Sección 6). El cierre es evidente cuando al menos una ruta completa un proyecto sin
La contribución humana no es indispensable en ninguno de los dos roles. Reformas de licencias
Acortan los plazos, pero solo acercan el final si prescinden de una decisión humana.
previamente indispensable (Sección 5). La participación de la IA en el diseño de chips informa el ciclo de
hardware, pero no especifica quién decide expandir la producción (Google DeepMind, 2025).
La identificación de los puntos de inflexión requiere evidencia conjunta sobre evaluación, dependencia humana y
sustituciones.

CONSIDERACIONES FINALES

El ensayo comenzaba con cuatro preguntas. Respecto a la primera, la auto-mejora de la IA era...
descritos como cuatro ciclos acoplados que giran con relojes muy diferentes: meses en
Las tecnologías digitales tardan entre años y una década en desarrollarse en comparación con las tecnologías físicas, de las cuales las tecnologías digitales dependen en un grado aún incierto.
Por lo tanto, las capacidades de la IA no han crecido de forma continua como las de los chips de ordenador, sino que lo han hecho de forma gradual.
en S, contenido por barreras, incluida la potencia de cálculo disponible.
En cuanto al segundo punto, el punto de inflexión en la capacidad se refiere al trabajo de investigación humana; el de
autonomía, una vez que se cierra el vínculo físico, cuando la IA comienza a impulsar la evolución de su base.
material sin que las decisiones o el trabajo humanos sean indispensables. Solo el segundo marcaría
la pérdida de control: antes de eso, por muy capaz que se vuelva la IA en el mundo digital, su
El crecimiento físico sigue dependiendo de las personas; después de eso, continúa sin la intervención de nadie más.

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

es necesario actuar. Una posible consecuencia del cierre es la autoaceleración, que solo entonces
Ya no dependería de decisiones humanas.

En cuanto al tercer punto, el ciclo puede cerrarse mediante la ejecución, con robots que operan bajo la toma de decisiones humanas, o mediante...
toma de decisiones, con humanos actuando como actuadores de IA; la presencia de trabajadores, por sí sola, no indica
¿En qué lado del punto de inflexión nos encontramos? Los relojes físicos, la dificultad de la automatización y la dispersión.
Las decisiones retrasan el cierre; la competencia, la transferencia gradual de decisiones y la
Simplificar los pasos lo acerca.

En cuanto al cuarto punto, el ensayo propone indicios no validados de dependencia humana en la investigación.
y en lo que respecta a la ejecución, las decisiones y las alternativas en la cadena física, vinculadas a los cuatro futuros.
La conclusión central es que el punto de inflexión en el que la humanidad perdería el control sobre
La evolución de la IA no es una auto-mejora digital, sino el momento en que la IA se convierte en...
para guiar de forma autónoma la evolución de su base física también. Hasta entonces, existe capacidad sin
autonomía. El ensayo no evalúa si este punto sería deseable o cómo evitarlo; solo busca...
Para caracterizarlo e indicar cómo podría reconocerse.

El marco conceptual tiene sus limitaciones. La reseña es narrativa y, dada la novedad del tema, aprovecha bien este enfoque.
Parte del marco incluye preimpresiones e informes técnicos. Otra parte se centra en los análisis.
estudios prospectivos del Forethought Institute, que sirven aquí como interlocutores conceptuales;
El orden entre las inflexiones también está respaldado por sus propias razones y por la literatura económica.
(Subsección 4.4). Los relojes son órdenes de magnitud, especialmente los de los Estados Unidos; el papel de
Los proveedores de energía, minerales y territorio, como Brasil, requieren su propio análisis. Autonomía
Es relativo al alcance, con la financiación y los materiales como condiciones externas, y el orden entre
Estas inflexiones son una hipótesis, no una demostración. El trabajo futuro pretende comprobarlo.
enmarcado con una revisión integradora y un modelo de simulación exploratoria calibrado
utilizando series históricas, examinando las condiciones bajo las cuales ocurre cada futuro y convirtiendo las señales.
en indicadores.

REFERENCIAS

AGHION, Philippe; JONES, Benjamin F.; JONES, Charles I. Inteligencia artificial y crecimiento económico. En: AGRAWAL, Ajay; GANS, Joshua; GOLDFARB, Avi (eds.). La economía de la inteligencia artificial: una agenda. Chicago: University of Chicago Press, 2019. págs. 237-282.

ARMSTRONG, Stuart; BOSTROM, Nick; SHULMAN, Carl. Corriendo hacia el precipicio: un modelo de desarrollo de inteligencia artificial. AI & Society, vol. 31, n.º 2, págs. 201–206, 2016. DOI: 10.1007/s00146-015-0590-y.

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

AUTOR, David H. ¿Por qué todavía hay tantos empleos? La historia y el futuro de la automatización en el lugar de trabajo. Journal of Economic Perspectives, vol. 29, n.º 3, págs. 3-30, 2015. DOI: 10.1257/jep.29.3.3.

BENGIO, Yoshua et al. Informe internacional sobre seguridad de la IA 2026. Londres: Departamento de Ciencia, Innovación y Tecnología, 3 de febrero de 2026. Serie de documentos de investigación del DSIT 2026/001. Disponible en: https://internationalaisafetyreport.org/publication/international-ai-safetyreport-2026. Consultado el 17 de octubre de 2026.

BENGIO, Yoshua; LECUN, Yann; HINTON, Geoffrey. Aprendizaje profundo para la IA. Communications of the ACM, vol. 64, n.º 7, págs. 58–65, julio de 2021. DOI: 10.1145/3448250.

BLOOM, Nicholas et al. ¿Cada vez es más difícil encontrar ideas? American Economic Review, vol. 110, n.º 4, págs. 1104–1144, 2020. DOI: 10.1257/aer.20180338.

BÖRJESON, Lena et al. Tipos de escenarios y técnicas: hacia una guía del usuario. Futures, vol. 38, n.º 7, págs. 723–739, 2006. DOI: 10.1016/j.futures.2005.12.002.

BOSTROM, Nick. Superinteligencia: caminos, peligros, estrategias. Oxford: Oxford University Press, 2014.

BURTSEV, Mikhail. Recursive criticality of AI self-improvement. Preprint. 31 de agosto de 2026. arXiv: 2609.00137. Disponible en: https://arxiv.org/abs/2609.00137. Consultado el 15 de octubre de 2026.

CHEN, Mingguang; WANG, Licheng; QU, Bo. Auto-mejora recursiva en IA: del autorrefinamiento limitado a los bucles de investigación autónomos. Preimpresión. 8 de julio de 2026. arXiv: 2607.07663. Disponible en: https://arxiv.org/abs/2607.07663. Consultado el 15 de octubre de 2026. Versión 2, con fecha del 6 de septiembre de 2026.

CUNNINGHAM, Tom et al. La economía de la auto-mejora recursiva. Preimpresión. 14 de septiembre de 2026. arXiv: 2609.15802. Disponible en: https://arxiv.org/abs/2609.15802. Consultado el 15 de octubre de 2026.

DAVIDSON, Tom; HADSHAR, Rose. La explosión industrial. Forethought. 21 de mayo de 2025. Disponible en: https://www.forethought.org/research/the-industrial-explosion. Consultado el 15 de octubre de 2026.

DAVIDSON, Tom; HADSHAR, Rose; MACASKILL, William. Tres tipos de explosión de inteligencia. Previsión. 17 mar. 2025. Disponible en: https://www.forethought.org/research/three-types-of-intelligence-explosion. Consultado el 15 de octubre de 2026.

ETH, Daniel; DAVIDSON, Tom. ¿Provocará la automatización de la I+D en IA una explosión de inteligencia artificial en el software? Forethought. 26 de marzo de 2025. Disponible en: https://www.forethought.org/research/will-ai-r-and-d-automation-cause-a-softwareintelligence-explosion. Consultado el 15 de octubre de 2026.

FAVARO, Marina; CLARK, Jack. Cuando la IA se construye a sí misma. Anthropic. 2026. Disponible en: https://www.anthropic.com/institute/recursive-self-improvement. Consultado el 16 de octubre de 2026.

FREITAS JUNIOR, Robert A.; MERKLE, Ralph C. Máquinas cinemáticas autorreplicantes. Georgetown, TX: Bioscience. Landas. 2004. Disponible en: https://www.molecularassembler.com/KSRM.htm. Consultado el 17 de octubre de 2026.

GOOD, Irving John. Especulaciones sobre la primera máquina ultrainteligente. En: ALT, Franz L.; RUBINOFF, Morris (eds.). Avances en computadoras. Nueva York: Academic Press, 1965. vol. 6, págs. 31–88. Disponible en: https://vtechworks.lib.vt.edu/bitstreams/a5e423ee-54e04eec-aeca-32b73f851af5/download. Consultado el 16 de octubre de 2026. Reimpresión parcial, pág. 31–33, proporcionado por Virginia Tech.

GOOGLE DEEPMIND. AlphaEvolve: un agente de codificación con tecnología Gemini para el diseño de algoritmos avanzados. Disponible en mayo. 14 2025. en:

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

https://deepmind.google/discover/blog/alphaevolve-a-gemini-powered-coding-agent-for-designing-advanced-algorithms/. Consultado el 15 de octubre de 2026.

HILTUNEN, Elina. El signo futuro y sus tres dimensiones. Futures, vol. 40, n.º 3, págs. 247–260, 2008. DOI: 10.1016/j.futures.2007.08.021.

HO, Anson et al. Avances algorítmicos en modelos de lenguaje. Preimpresión. 9 de marzo de 2024. arXiv: 2403.05812. Disponible en: https://arxiv.org/abs/2403.05812. Consultado el 15 de octubre de 2026.

ORGANISMO INTERNACIONAL DE ENERGÍA. Energía e IA. París: AIE, 10 de abril de 2025. Informe especial sobre las perspectivas energéticas mundiales. Disponible en: https://www.iea.org/reports/energy-and-ai. Consultado el 15 de octubre de 2026.

KOKOTAJLO, Daniel et al. IA 2027. Proyecto Futuros de la IA. 3 de abril de 2025. Disponible en: https://ai-2027.com/. Consultado el: 15 de octubre de 2026.

KRIZHEVSKY, Alex; SUTSKEVER, Ilya; HINTON, Geoffrey E. Clasificación de ImageNet con redes neuronales convolucionales profundas. En: PEREIRA, Fernando et al. (eds.). Avances en sistemas de procesamiento de información neuronal. [S. l.]: Curran Associates, 2012. vol. 25, págs. 1097–1105. Disponible en https://proceedings.neurips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45bAbstract.html. Consultado el 16 de octubre de 2026. en:

KULVEIT, Jan et al. Desempoderamiento gradual: riesgos existenciales sistémicos del desarrollo incremental de la IA. Preimpresión. 28 de enero de 2025. arXiv: 2501.16946. Disponible en: https://arxiv.org/abs/2501.16946. Consultado el 15 de octubre de 2026.

KWA, Thomas et al. Medición de la capacidad de la IA para completar tareas de software largas. Preimpresión. 18 de marzo de 2025. arXiv: 2503.14499. Disponible en: https://arxiv.org/abs/2503.14499. Consultado el: 16 de octubre. 2026. Versión 4, 10 de julio de 2026.

LIFLAND, Eli; KOKOTAJLO, Daniel. Evaluación de las predicciones de AI 2027 para 2025. AI Futures Project. 12 de febrero de 2026. Disponible en: https://blog.aifutures.org/p/grading-ai-2027s-2025predictions. Consultado el 16 de octubre de 2026. Actualizado en julio de 2026.

LIFLAND, Eli; KOKOTAJLO, Daniel; HALSTEAD, Brendan. Aclarando cómo han cambiado nuestras previsiones sobre los plazos de la IA desde AI 2027. AI Futures Project. 27 de enero de 2026. Disponible en: https://blog.aifutures.org/p/clarifying-how-our-ai-timelines-forecasts. Consultado el 16 de octubre de 2026.

LU, Chris et al. El científico de IA: hacia el descubrimiento científico abierto y totalmente automatizado. Preimpresión. 12 de agosto de 2024. arXiv: 2408.06292. Disponible en: https://arxiv.org/abs/2408.06292. Consultado el 15 de octubre de 2026.

MACASKILL, William; MOORHOUSE, Fin. Preparándose para la explosión de inteligencia. Previsión. 11 de marzo de 2025. Disponible en: https://www.forethought.org/research/preparingfor-the-intelligence-explosion. Consultado el 15 de octubre de 2026.

MANHEIM, David; GARRABRANT, Scott. Categorización de las variantes de la ley de Goodhart. Preimpresión. 13 de marzo de 2018. arXiv: 1803.04585. Disponible en: https://arxiv.org/abs/1803.04585. Consultado el 15 de octubre de 2026.

MEADOWS, Donella H. Pensar en sistemas: una introducción. White River Junction: Chelsea Green, 2008.

MITCHELL, Melanie. Por qué la IA es más difícil de lo que pensamos. Preimpresión. 28 de abril de 2021. arXiv: 2104.12871. Disponible en: https://arxiv.org/abs/2104.12871. Consultado el 16 de octubre de 2026.

MOORE, Gordon E. Cómo integrar más componentes en circuitos integrados. Electronics, vol. 38, n.º 8, págs. 114-117, 19 de abril de 1965.

MORAVEC, Hans. Niños de la mente: el futuro de la inteligencia robótica y humana. Cambridge, MA: Harvard University Press, 1988.

NEUMANN, John von. Teoría de los autómatas autorreproductores. Urbana: University of Illinois Press, 1966. Editado y completado por Arthur W. Burks.

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

NORDHAUS, William D. ¿Nos acercamos a una singularidad económica? La tecnología de la información y el futuro del crecimiento económico. American Economic Journal: Macroeconomics, vol. 13, n.º 1, págs. 299-332, 2021. DOI: 10.1257/mac.20170105.

NOVIKOV, Alexander et al. AlphaEvolve: un agente de codificación para el descubrimiento científico y algorítmico. Preimpresión. 16 de junio de 2025. arXiv: 2506.13131. Disponible en: https://arxiv.org/abs/2506.13131. Consultado el 15 de octubre de 2026.

OPENAI. Acelerando la investigación: una mirada al interior de OpenAI. 6 de septiembre de 2026. Disponible en: https://openai.com/pt-BR/index/research-acceleration-view-inside-openai/. Consultado el 16 de octubre de 2026.

PARASURAMAN, Raja; SHERIDAN, Thomas B.; WICKENS, Christopher D. Un modelo para los tipos y niveles de interacción humana con la automatización. IEEE Transactions on Systems, Man, and Cybernetics – Part A, vol. 30, n.º 3, págs. 286–297, 2000. DOI: 10.1109/3468.844354.

RAND, Joseph et al. En cola: características de las centrales eléctricas que buscan interconexión de transmisión a finales de 2024. Berkeley: Laboratorio Nacional Lawrence Berkeley, diciembre. 2025. Edición 2025. Disponible en: https://emp.lbl.gov/publications/queued-2025edition-characteristics. Consultado el 15 de octubre de 2026.

SANTONI DE SIO, Filippo; HOVEN, Jeroen van den. Control humano significativo sobre sistemas autónomos: una perspectiva filosófica. Frontiers in Robotics and AI, vol. 5, pág. 15, 28 de febrero de 2018. DOI: 10.3389/frobt.2018.00015.

SCHMIDHUBER, Jürgen. Automejora recursiva (RSI) desde 1987. Nota técnica IDSIA-9-26. IDSIA. Septiembre de 2026. Disponible en: https://people.idsia.ch/~juergen/recursiveself-improvement.html. Consultado el 16 de octubre de 2026.

ASOCIACIÓN DE LA INDUSTRIA DE SEMICONDUCTORES. Diseño de chips. [¿2026?]. Disponible en: https://www.semiconductors.org/policies/chip-design. Consultado el 15 de octubre de 2026.

SHUMAILOV, Ilia et al. Los modelos de IA colapsan cuando se entrenan con datos generados recursivamente. Naturaleza, vol. 631, pág. 755–759, 25 de julio de 2024. DOI: 10.1038/s41586-024-07566-y.

SRIKANTH, Dhruv et al. Automejora recursiva de agentes de investigación de IA. Preimpresión. 22 de septiembre de 2026. arXiv: 2609.26457. Disponible en: https://arxiv.org/abs/2609.26457. Consultado el 15 de octubre de 2026.

STERMAN, John D. Dinámica empresarial: pensamiento sistémico y modelización para un mundo complejo. Boston: Irwin/McGraw-Hill, 2000.

VERWEY, John. Sin permisos, no hay fábricas: la importancia de la reforma regulatoria para la fabricación de semiconductores. Washington, DC: Centro para la Seguridad y la Tecnología Emergente, octubre de 2021. Disponible en: https://cset.georgetown.edu/publication/no-permitsno-fabs/. Consultado el 15 de octubre de 2026.

WANG, Jiaming. ¿Qué frena la auto-mejora recursiva en robótica? Lecciones de 123 rondas de descubrimiento de habilidades de agencia. Preimpresión. 23 de septiembre de 2026. arXiv: 2609.31760. Disponible en: https://arxiv.org/abs/2609.31760. Consultado el 15 de octubre de 2026.

WHITFILL, Parker; Wu, Cheryl. ¿ Los cuellos de botella computacionales impedirán una explosión de inteligencia? Preimpresión. 31 de julio de 2025. arXiv: 2507.23181. Disponible en: https://arxiv.org/abs/2507.23181. Consultado el 15 de octubre de 2026.

WIJK, Hjalmar et al. RE-Bench: evaluación de las capacidades de I+D de IA de vanguardia de los agentes de modelos de lenguaje frente a expertos humanos. Preimpresión. 22 de noviembre de 2024. arXiv: 2411.15114. Disponible en: https://arxiv.org/abs/2411.15114. Consultado el 16 de octubre de 2026.

XIAO, Wenli et al. ENPIRE: Automejora de políticas de robots agentes en el mundo real. Preimpresión. Conferencia sobre Aprendizaje Robótico (CoRL) 2026. 18 de junio de 2026. arXiv: 2606.19980. Disponible en: https://arxiv.org/abs/2606.19980. Consultado el: 15 de octubre de 2026.

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.

Machine Translated by Google

Año VII, vol. 1, 2026 | Envío: 9 de octubre de 2026 | Aceptación: 14 de octubre de 2026 | Publicación: 26 de octubre de 2026

ZHANG, Jenny et al. Máquina de Darwin-Gödel: evolución abierta de agentes que se auto-mejoran.
Preimpresión. 29 de mayo de 2025. arXiv: 2505.22954. Disponible en: https://arxiv.org/abs/2505.22954.
Consultado el 15 de octubre de 2026.

Este es un artículo de acceso abierto publicado bajo la licencia Creative Commons Attribution, que permite su uso.
Se permite su distribución y reproducción en cualquier medio, sin restricciones, siempre que se cite correctamente la obra original.