

Irene Unceta, Paula Subías-Beltrán, Oriol Pujol

La deuda epistémica de la Inteligencia Artificial generativa

nature machine intelligence, 26 de agosto de 2026.

Los sistemas de IA generativa comprimen en segundos tareas cognitivas que antes requerían un esfuerzo sostenido. Es decir, para muchos fines, es tiempo bien ahorrado. Sin embargo, cuando se trata de la creación de conocimiento, el tiempo hace más que retrasar los resultados. También desempeña una función cognitiva: es el medio a través del cual las ideas pueden ser interrogadas, reconstruidas y apropiadas por quien las desarrolla [1]. Cuando la presión por producir desplaza esa función, la relación entre una persona y lo que produjo se debilita. El resultado puede parecer convincente mientras que quien lo produjo en realidad no lo ha elaborado plenamente. Lo que surge es una brecha entre lo que se presenta y lo que se comprende y se hace propio: una deuda epistémica. Este artículo caracteriza en qué consiste esa deuda en el caso de la IA generativa y cómo puede amortizarse.

DESCARGA COGNITIVA EN LA ERA DE LA IA GENERATIVA

La cognición humana siempre ha dependido de apoyos externos. La escritura, los diagramas, los mapas y las calculadoras nos han ayudado durante mucho tiempo a formular ideas y problemas, mantener un registro de ellos y avanzar gradualmente hacia sus soluciones. El cuaderno nos ayuda a recordar; el diagrama nos ayuda a ver estructuras que no habíamos advertido; la calculadora libera atención para otras partes de un problema. En cada caso, realizamos parte del trabajo cognitivo a través de la propia herramienta.

Esta es la intuición central de la tesis de la «mente extendida» propuesta por Clark y Chalmers [2]. Su ejemplo más conocido es Otto, un paciente con deterioro de la memoria que depende de un cuaderno para almacenar aquello que no puede retener internamente. Como el cuaderno está disponible de forma constante, se consulta de manera habitual y se considera fiable, funciona de un modo muy similar a la memoria biológica. Entendida de este modo, la externalización constituye una extensión de la capacidad cognitiva.

La IA generativa invita al mismo marco interpretativo [3]. La integramos en nuestros flujos de trabajo, descargamos en ella partes de tareas que de otro modo realizaríamos nosotros mismos y dirigimos nuestra atención hacia otros aspectos. Al igual que el cuaderno de Otto, puede convertirse en parte de nuestro sistema cognitivo. Pero existe una diferencia importante. El cuaderno de Otto almacena aquello que Otto escribe en él. Un sistema de IA generativa produce contenido que Otto nunca formuló: argumentos que no construyó y estructuras que no diseñó. Además, ese contenido no está determinado únicamente por el *prompt* de Otto. Está configurado por múltiples capas arquitectónicas que operan bajo la superficie: instrucciones a nivel de sistema e historial de conversación, ciclos de refinamiento, conexiones externas y barreras de protección basadas en políticas. Cada capa incorpora supuestos y restricciones que el usuario no puede controlar. Lo que moldea el pensamiento del usuario está ahora, en parte, creado por otros y, en parte, oculto para él. Que esto constituya una forma de extensión de la cognición depende de lo que

el usuario haga con ello, y hacerlo verdaderamente propio requiere algo más que aceptarlo.

EL TRABAJO QUE SIGUE A LA GENERACIÓN

Para que una afirmación llegue a ser realmente propia, debe exponerse a la duda y a la revisión mediante la interrogación [4]. Cuando esa afirmación se origina en otro lugar, esa exigencia se vuelve aún más estricta. Interrogar una afirmación exige reconstruir el razonamiento que hay detrás de ella, rastrear cómo se relacionan sus distintas partes entre sí, qué supuestos la sostienen y en qué aspectos concuerda o entra en conflicto con los propios valores y creencias. Este es un proceso de apropiación [5]. En el sentido de Ricoeur, la apropiación es lo que ocurre cuando un significado inicialmente ajeno es asumido activamente y convertido en propio mediante la interpretación. Es comprenderlo lo suficientemente bien como para poder revisarlo, defenderlo o rechazarlo y, por tanto, habitarlo.

Sin apropiación, un resultado aparentemente coherente puede permanecer cognitivamente inerte. La persona puede repetir el argumento, pero no puede «ver» su arquitectura interna [6]. Puede no advertir si una premisa contradice sus otras creencias o si la lógica se desmorona ante un ligero cambio de contexto. El resultado responde a la solicitud, pero el individuo sigue siendo un extraño respecto de sus propias afirmaciones.

La apropiación requiere tiempo. La distancia entre producir una afirmación y comprenderla en profundidad no nace con la IA generativa. Lo que estos sistemas cambian es su escala. Crean una asimetría temporal: comprimen el tiempo necesario para generar contenido, pero el tiempo necesario para comprender ese contenido no disminuye en la misma proporción. Cuando la presión por producir es elevada, el tiempo ahorrado mediante una generación rápida tiende a ser absorbido por la demanda de producir todavía más, en lugar de invertirse en comprender lo que acaba de producirse. En este contexto, el tiempo se trata como un coste que debe minimizarse, y el esfuerzo que requiere la interrogación acaba percibiéndose como un gasto que compite con otras exigencias más urgentes [7].

Cuando una persona sopesa esas exigencias frente a la importancia percibida de una determinada cuestión, puede llegar a la conclusión de que esta ya ha sido suficientemente abordada y que un esfuerzo adicional no está justificado [7]. Los entornos tecnológicos en los que estamos cada vez más inmersos hacen más probable esta forma de «pereza epistémica». Cuando los sistemas de IA presentan determinadas creencias como más accesibles, más destacadas o más coherentes con aquello que se acepta ampliamente [8], la información disponible parece suficiente y disminuye la presión por someterla a examen. La abundancia de información debilita aún más el incentivo para someter las ideas a escrutinio, y el coste de la apropiación acaba pareciendo desproporcionado respecto a cualquier posible beneficio.

Sin embargo, no toda descarga cognitiva pone en riesgo la apropiación. Externalizar tareas de escasa importancia, como dar formato a una lista de referencias o corregir la gramática, puede reducir la carga cognitiva sin menoscabar la apropiación. El problema no es la delegación en sí misma, ni el uso de herramientas generativas como tal. El umbral crítico se alcanza cuando la delegación desplaza a la interrogación como práctica sostenida: cuando los resultados generados se aceptan sin haber sido realmente trabajados. Lo importante es si la persona que utilizó la herramienta todavía puede reconstruir lo que produjo y respaldarlo. Esa es la diferencia entre un uso de la herramienta que amplía la mente y uno que la sustituye [3].

Cuando una persona renuncia al esfuerzo de hacer suyo un resultado generado, se abre una brecha entre aquello que puede presentar y aquello que puede reconstruir de forma independiente. Es esa brecha la que constituye la deuda epistémica: una obligación incumplida de comprender, revisar y respaldar aquello que se ha producido.

LA DEUDA EPISTÉMICA Y CÓMO SE ACUMULA

La deuda epistémica se manifiesta al menos en cuatro dimensiones de la relación entre una persona y aquello que produce.

La primera se refiere a la autoría. Una persona puede iniciar una tarea, proporcionar *prompts* y aprobar el resultado final, pero una parte sustancial del trabajo no ha sido realizada por ella misma. El trabajo lleva su nombre, pero la afirmación pública que ese nombre implica solo está parcialmente justificada [9].

La segunda se refiere a la autenticidad: el grado en que lo que uno produce ha sido contrastado con sus propios valores y principios. Cuando comienza a acumularse deuda epistémica, ya no puede garantizarse esa alineación con los propios valores [10].

La tercera se refiere a la propiedad en un sentido más profundo y personal: experimentar una afirmación como verdaderamente propia requiere haberla trabajado, haberse enfrentado a su estructura interna y haber seguido su razonamiento. Ese sentimiento de propiedad se gana mediante el compromiso con el resultado. Sin ese esfuerzo, lo producido permanece a cierta distancia de la persona que lo produjo [9].

La cuarta dimensión se refiere a la responsabilidad. Respalda una afirmación implica haber identificado sus límites y anticipado sus implicaciones. La propiedad concierne a la relación con uno mismo; la responsabilidad concierne a lo que uno debe a los demás. Cuando partes del proceso de razonamiento han sido delegadas y posteriormente no se han reelaborado, esa capacidad se debilita. Se firma el texto, pero no se puede responder plenamente por él.

Cada una de estas dimensiones implica un trabajo que ha sido aplazado y sigue pendiente. De ahí el término deuda epistémica, inspirado en el concepto de deuda técnica de la ingeniería del software [11]. Cuando los desarrolladores

toman atajos para acelerar la entrega de un producto, contraen una deuda: el código funciona, pero su fragilidad subyacente hace que los cambios futuros sean más difíciles y más propensos a fallar. La ganancia inicial en velocidad genera un coste diferido que aumenta si no se aborda. La deuda epistémica funciona del mismo modo. Cada vez que se acepta un resultado generado por IA sin apropiación, se introduce una brecha. Del mismo modo que un código funcional puede seguir teniendo deuda técnica, un resultado correcto puede seguir teniendo deuda epistémica. Si no se aborda, esa deuda puede acumularse con el tiempo.

La acumulación es posible porque el conocimiento posee una estructura diacrónica y estratificada. Dado que esa estructura es parcialmente modular y está conectada de manera desigual, algunas lagunas permanecen locales, mientras que otras adquieren importancia cuando razonamientos y prácticas posteriores dependen de ellas. Las afirmaciones se evalúan en relación con conjuntos más amplios de creencias [12], por lo que una laguna que no se examina no siempre permanece donde comenzó. Que llegue a acumularse depende de dos condiciones: que el trabajo posterior termine apoyándose en ella y que el razonamiento que la sustenta permanezca sin ser interrogado en lugar de ser reconstruido durante el proceso. Cuando ambas condiciones se cumplen, la deuda epistémica puede acumularse a medida que aumentan las dependencias y se incrementa el coste de volver sobre la laguna original.

La metáfora del barco de Neurath ilustra esta idea. Revisamos nuestro conocimiento tablón a tablón mientras seguimos a flote, sin disponer de un dique seco desde el que reconstruirlo desde cero. El barco no es un todo indiferenciado: tiene compartimentos y piezas con funciones estructurales distintas. Un tablón con un defecto visible puede seguir siendo un problema localizado, identificable y reparable sin afectar al resto de la embarcación. El problema surge cuando un tablón se instala sin una inspección adecuada y más adelante se convierte en soporte de trabajos posteriores. Una vez que otros tablonés, uniones y la carga descansan sobre él, el barco puede seguir navegando e incluso parecer cada vez más seguro, mientras el coste de corregir la reparación original aumenta, porque hacerlo ahora implica desmontar parte de lo construido encima.

Un argumento que nunca fue interrogado actúa como ese tablón: se acepta sin inspección y se construye sobre él. Entonces operan conjuntamente tres factores. En primer lugar, las afirmaciones posteriores pasan a apoyarse en la anterior, de modo que reconstruir significaría reabrir no solo el resultado más reciente, sino también varias capas inferiores. En segundo lugar, a medida que esas dependencias se acumulan, el esfuerzo necesario para revisar el razonamiento original puede llegar a ser tan elevado que delegar nuevamente se convierta en el camino de menor resistencia. En tercer lugar, cuando la dependencia de resultados no examinados se convierte en una rutina, disminuye la necesidad percibida de volver sobre ellos. Cada uno de estos factores refuerza a los demás: cuanto más se acumulan las dependencias, más costoso resulta abordar la deuda y mayor es el incentivo para aplazar una vez más la investigación. La deuda se refuerza a sí misma cuando las dependencias dejan de ser locales y sus efectos se acumulan.

UNA PREOCUPACIÓN MÁS AMPLIA

Existe una versión de este problema que va más allá de cualquier tarea o resultado concreto. Las capacidades cognitivas que no se ejercitan tienden a debilitarse. El hábito de la interrogación independiente —seguir el rastro de un argumento, poner a prueba sus supuestos y relacionarlo con aquello en lo que uno ya cree— se construye mediante la práctica. Cuando la interrogación se evita sistemáticamente porque la parte más exigente del razonamiento se descarga en otros, la propia capacidad puede deteriorarse. Los resultados preliminares apuntan en esa dirección: el uso habitual de asistencia mediante IA se ha asociado con una menor conectividad neuronal durante tareas de escritura [13], un resultado que, si se confirma, podría ser un indicador temprano de este tipo de cambio cognitivo.

Las consecuencias a largo plazo son más difíciles de medir, pero no difíciles de imaginar. Con el tiempo, las personas pueden seguir siendo capaces de producir resultados sofisticados mientras pierden la capacidad de deconstruirlos. Si seguimos siendo hábiles formulando *prompts* y editando, pero cada vez menos capaces de generar el razonamiento subyacente, nuestra relación con el conocimiento se volverá progresivamente más indirecta.

La deuda epistémica también puede acumularse a escala de un campo de conocimiento o de una institución. Las ideas circulan, son citadas y sirven de base unas para otras sin que nadie las haya trabajado realmente. Las afirmaciones no examinadas acumulan autoridad simplemente por el hecho de circular, y cada referencia otorga una legitimidad aparente a la siguiente [14]. Esta dinámica es anterior a la IA generativa. Lo que estos sistemas modifican es la velocidad con la que tales afirmaciones pueden producirse y adoptarse, lo que puede acelerar esa acumulación. Lo que parece un cuerpo creciente de conocimiento puede ser, en realidad, un cuerpo creciente de supuestos heredados cuyos fundamentos nadie ha considerado necesario inspeccionar.

Parte de lo que permite que esta herencia pase desapercibida es que los sistemas de IA generativa no dejan rastro visible de su contribución. Esa invisibilidad facilita presentar el trabajo como propio y que los demás lo reciban como tal. Esta facilidad corre el riesgo de normalizar una práctica que plantea cuestiones de integridad, el llamado «problema de la procedencia» (*provenance problem*) [15]. Cuando esto se normaliza, la brecha que crea la deuda epistémica se convierte en una cuestión de honestidad respecto al propio trabajo.

AMORTIZAR LA DEUDA

Si la deuda epistémica se acumula, también puede reducirse. Amortizarla significa recuperar el trabajo cognitivo que fue aplazado. En la práctica, esto adopta la forma de reconstrucción. Los argumentos generados externamente deben reformularse en los propios términos. El razonamiento de una afirmación debe seguirse y ponerse a prueba hasta que pueda comprenderse y

defenderse de forma autónoma. El esfuerzo que no se invirtió en la generación debe invertirse en la apropiación.

Esto no constituye un argumento a favor de la ineficiencia. La rapidez que ofrece la IA generativa es una ganancia real. Pero, debido a la lógica acumulativa descrita anteriormente, el coste de la reconstrucción puede superar cualquier eficiencia obtenida. Esa ganancia genera, por tanto, una obligación: el tiempo ahorrado debe reinvertirse, al menos en parte, en el trabajo más lento de la apropiación. Lo que dificulta esto es la misma presión cultural que impulsa en primer lugar la pereza epistémica: cuando el resultado ya hace lo que se le pide, la amortización parece innecesaria y el esfuerzo que requiere se percibe como un coste sin retorno.

La preocupación radica precisamente en que la IA generativa facilita evitar el trabajo de apropiación y, al hacerlo, permite que la deuda epistémica se acumule. Las condiciones en las que está siendo adoptada, marcadas por la presión para producir rápidamente y en grandes cantidades, son precisamente las menos favorables para la reinversión que mantendría esa deuda bajo control. Los sistemas de IA generativa pueden ampliar nuestra capacidad para crear conocimiento. La velocidad que ofrecen y el tiempo que requiere la apropiación no están en conflicto por naturaleza. Solo entran en conflicto cuando ese tiempo se trata como un coste que debe eliminarse en lugar de como el medio a través del cual se crea el conocimiento.

Esto exige reconocer que el tiempo que requiere la apropiación está bien invertido y preservar espacios para la interrogación allí donde los incentivos premian la rapidez y el volumen. De ello se derivan dos prácticas. Antes de construir nuevo trabajo sobre un resultado generado, su razonamiento debe reconstruirse hasta poder defenderlo sin ayuda. Y cuando se comparta un resultado generado, el uso de la IA debe hacerse visible, para que los resultados no examinados no se hereden inadvertidamente. Considerar el tiempo necesario para la apropiación como una inversión es lo que permite que la velocidad de la IA generativa se traduzca en un conocimiento verdaderamente propio.

Declaración sobre el uso de IA.

Durante la preparación del manuscrito utilizamos Claude Sonnet 4.6 y GPT 5.5 para apoyar la búsqueda de referencias, ayudar a mejorar la claridad de la expresión y condensar el texto para ajustarlo al límite de palabras. El texto final fue redactado originalmente por los autores. Asumimos nuestra plena responsabilidad por todas las afirmaciones.

Referencias

- [1] Dewey, J. **How We Think**. D.C. Heath & Co., Lexington (1910)
- [2] Clark A, Chalmers D. **Analysis**. 58, 7 (1998).
- [3] Clark A. **Nat Commun.** 16, 4627 (2025).
- [4] Peirce CS. **Pop Sci Mon.** 12, 1 (1877).

- [5] Ricoeur P. *Interpretation Theory: Discourse and the Surplus of Meaning*. Texas Christian University Press, Fort Worth (1976).
- [6] Ananny M, Crawford K. *New Media Soc.* 20, 3 (2018).
- [7] Jarczewski D. *Soc Epistemol.* (2024)
- [8] Coeckelbergh M. *Soc Epistemol.* 40, 59 (2026).[9] Draxler F, Werner A, Lehmann F, Hoppe M, Schmidt A, Buschek D, Welsch R. 2024. *ACM Trans. Comput.-Hum. Interact.* 31, 2 (2024).
- [10] Sharma M, McCain M, Douglas R, Duvenaud D. *arXiv:2601.19062* (2026).
- [11] Sculley D, Holt G, Golovin D, Davydov E, Phillips T, Ebner D, Chaudhary V, Young M, Crespo JF, Dennison D. *Adv Neural Inf Process Syst.* 28, 2503 (2015).
- [12] Quine WV. *Philos Rev.* 60, 1 (1951).
- [13] Kosmyna N, Hauptmann E, Yuan YT, Situ J, Liao XH, Beresnitzky AV, Braunstein I, Maes P. *arXiv:2506.08872* (2025).
- [14] Messeri L, Crockett MJ. *Nature*. 627, 49 (2024).
- [15] Earp BD, Yuan H, Koplin J, Porsdam Mann, S. *Nat M*

Autores con igual contribución: Irene Unceta, ESADE Business School, Universidad Ramon Llull; Paula Subías-Beltrán, Observatorio de Bioética y Derecho - Cátedra UNESCO de Bioética, Universitat de Barcelona; Oriol Pujol, Departamento de Matemáticas e Informàtica, Universitat de Barcelona.
Correspondencia: Correo electrónico: paulasubias@ub.edu