El tercer régimen del conocimiento: la inteligencia artificial como respuesta emergente a los límites de la academia clásica y los repositorios abiertos

The Third Knowledge Regime: Artificial Intelligence as an Emergent Response to the Limits of Classical Academia and Open Repositories

Carlos Eduardo Ravello Joo
ORCID: 0009-0007-5631-7436
carlosravello.com
Trujillo — Lima, Perú

DOI: 10.5281/zenodo.24164493

Noviembre 2026 · Versión 2 (v1: mayo 2026)

Nota de versión

Versión 2 (noviembre de 2026). Esta versión corrige y matiza la de mayo de 2026, que sigue disponible en Zenodo.

Corregido:

- La cifra de SWE-bench, ahora con fuente en el AI Index 2025.
- La proporción de modelos notables producidos por la industria: casi el 90%, en 2024.
- El tiempo de doctorado: mediana de 7,2 años desde el posgrado y 5,7 desde el doctorado (NSF/NCSES).
- La encuesta de *Nature*: 52% de «crisis significativa», no 83%.
- Los años de Richard Smith al frente del *BMJ*: 13, no 25.
- El DOI y la revista de Carneiro et al. (2020).
- El título de Topaz et al. (2026), la descripción de su corpus y las causas que proponen.
- La fecha y la autoría de la propuesta de reforma de notabilidad de Wikidata.
- La atribución de las cifras de bloqueo de rastreadores: BuiltWith, Tollbit y Cloudflare.

Retirado: las cifras sin fuente verificable (PyPI, npm, uso de Wikidata en Wikipedia, integraciones comerciales) y la afirmación de que Wikidata es la fuente primaria del Knowledge Graph de Google.

Matizado: el uso de «fracaso», varias frases de tono de manifiesto y la lectura de Carneiro et al. y de Gartenberg et al.

Añadido:

- Una sección sobre los terrenos donde la revisión lenta debe mantenerse, con la distinción entre revisar por utilidad, por obligación o por defecto.
- Evidencia sobre el uso de IA por estudiantes y sobre cómo las IA rastrean, resumen y citan.
- El caso del propio autor como crítica al argumento del «candado circular».
- Una nota sobre el carácter de preprint del texto.

- Limitaciones ampliadas y las referencias que faltaban.
- Un anexo con el recálculo reproducible de las cifras del Reproducibility Project: Psychology.

Nota del autor

Este trabajo forma parte de una serie de preprints independientes del autor. Conecta con los marcos conceptuales desarrollados en Ravello Joo (2026a, 2026b, 2026c, 2026d), depositados en Zenodo. No ha recibido financiamiento externo. El autor declara ausencia de conflicto de intereses.

Resumen

La inteligencia artificial no supera a los humanos por capacidad cognitiva intrínseca. La ventaja operativa observable emerge tanto de capacidad computacional como de la eliminación de fricciones institucionales acumuladas que limitan la cognición humana colectiva. Este artículo examina ese diagnóstico desde tres frentes simultáneos: la lentitud estructural del sistema académico tradicional, el ruido epistémico producido por repositorios abiertos sin filtro adecuado, y el gatekeeping anónimo que opera sobre los grafos de conocimiento de los que se alimentan los sistemas de inteligencia artificial. Se introduce el concepto de costos de coordinación epistemológica como marco para entender por qué tanto la academia clásica como los repositorios abiertos fallan, cada uno a su manera, ante el escenario actual, y cómo la IA entra en esa grieta histórica. Cada régimen revisa por un motivo distinto: el abierto por utilidad, el clásico por obligación, y la IA no revisa, resume por defecto. Por eso el rigor que se le exige al conocimiento no puede depender del régimen por el que circula. Se delimita además dónde la tesis no aplica: en el conocimiento cuyo error no se puede revertir, como el clínico, la revisión lenta debe mantenerse. Se argumenta que el conocimiento ya no circula solo por las instituciones humanas, y que documentar con rigor antes de que los sistemas automatizados llenen ese vacío con imaginación sin método es una tarea que nos toca a quienes investigamos, no un adorno académico. La información, como cualquier sistema físico, tiende al desorden sin flujo de energía organizada. Las instituciones de conocimiento fueron los mecanismos que la humanidad construyó para reducir esa entropía localmente. Cuando esos mecanismos se quedan cortos, el desorden avanza más rápido que la corrección.

Palabras clave: fricción institucional, validación académica, repositorios abiertos, Wikidata, grafos de conocimiento, costos de coordinación epistemológica, inteligencia artificial, peer review, entropía informacional

Tabla 1. Tres regímenes epistemológicos

Sistema	Velocidad	Validación	Resultado
Academia clásica	Baja	Alta en lo formal, desigual en la práctica	Rigidez + invisibilidad
Repositorios abiertos	Alta	Mínima o nula	Ruido a escala
Inteligencia artificial	Alta	Compresión de señales	Ventaja operativa

1. Introducción

Al 24 de marzo de 2025, Hugging Face alojaba 1.533.973 modelos de inteligencia artificial (González et al., 2025). No como espectáculo. Como síntoma.

El síntoma apunta a una pregunta que las instituciones de conocimiento no están respondiendo con suficiente honestidad: ¿por qué una arquitectura sin metabolismo, sin carrera profesional,

sin miedo reputacional, sin tiempo biográfico, opera cognitivamente más rápido que los sistemas diseñados específicamente para producir y validar conocimiento?

La respuesta habitual —que la IA tiene más poder computacional— es verdadera pero incompleta. El poder computacional explica velocidad de procesamiento. No explica por qué una investigación clínica de cinco años puede quedar invisible para los sistemas que hoy construyen conocimiento a escala, mientras un repositorio mal estructurado en GitHub aparece indexado en horas.

Este artículo propone una respuesta más incómoda: la ventaja operativa de la IA emerge tanto de capacidad computacional como de la eliminación de fricciones institucionales acumuladas que limitan la cognición humana colectiva (Gartenberg et al., 2026; Maslej et al., 2025). No es un problema de inteligencia. Es un problema de arquitectura.

La información, como cualquier sistema físico, tiende al desorden sin flujo de energía organizada (Shannon, 1948; Schrödinger, 1944). Las instituciones de conocimiento fueron los mecanismos que la humanidad construyó para reducir esa entropía localmente. El desarrollo formal de ese marco conecta con trabajos previos del autor (Ravello Joo, 2026a, 2026b, 2026c). Lo que este artículo examina es qué sucede cuando esos mecanismos se quedan cortos al mismo tiempo, desde dos direcciones opuestas.

2. Dos fracasos simultáneos

Uso «fracaso» en un sentido preciso: un sistema que funcionó para su época y que no da abasto con la velocidad y el volumen de la actual.

2.1 El fracaso lento: academia como arquitectura de escasez

El peer review nació en un mundo donde el problema era la escasez informacional. Había poco conocimiento producido, y el desafío era protegerlo de la contaminación. La lentitud era una característica adaptativa, no un defecto: la validación cuidadosa era el precio razonable de la confiabilidad.

Ese mundo ya no existe. En 2024, casi el 90% de los modelos de IA notables vino de la industria (Maslej et al., 2025). En SWE-bench, un benchmark de problemas reales de programación presentado en 2023, los sistemas de IA resolvían el 4,4% de los problemas en 2023 y el 71,7% en 2024 (Maslej et al., 2025). En Estados Unidos, la mediana para obtener un doctorado de investigación fue de 7,2 años desde el ingreso al posgrado en 2022, y de 5,7 años desde el inicio del programa doctoral (NSF/NCSES, 2024).

El riesgo es que investigación relevante quede obsoleta durante su propio proceso de validación. Gartenberg et al. (2026), con el corpus completo de 6.957 envíos iniciales y 10.389 revisiones gestionados por *Organization Science* desde 2021, documentaron un aumento del 42% en los envíos desde la llegada de ChatGPT, con peor calidad de escritura, y concluyen que el sistema se está desplazando hacia más investigación, no mejor investigación.

Más grave: el sistema no produce con consistencia lo que promete. La Open Science Collaboration (2015) encontró que solo el 36% de las réplicas de estudios psicológicos obtuvo resultados estadísticamente significativos, y que el 39% fue juzgado como replicado por los propios equipos; los efectos de las réplicas fueron, en promedio, la mitad de los originales (recálculo en el Anexo). Camerer et al. (2018) replicaron 21 experimentos de ciencias sociales publicados en *Nature* y *Science*: 13 (62%) replicaron, con efectos de alrededor de la mitad del original. La base de datos de Retraction Watch superaba las 58.000 retractaciones en abril de 2025 (Retraction Watch, 2025). En las primeras siete semanas de 2026, en el subconjunto de acceso abierto de PubMed

Central, uno de cada 277 artículos tenía al menos una referencia inexistente, frente a uno en 458 en 2025 y uno en 2.828 en 2023. Los autores señalan como posibles orígenes las fábricas de papers, el fraude deliberado y el uso acrítico de herramientas de IA (Topaz et al., 2026).

Richard Smith (2006) escribió desde adentro del sistema: el peer review es un proceso fallido en el corazón de la ciencia. Y lo dijo alguien que lo administró, como editor del *BMJ*, durante trece años.

2.2 El fracaso rápido: repositorios como arquitectura de ruido

La respuesta al fracaso lento fue la velocidad sin filtro. Los repositorios abiertos —GitHub, Zenodo, Hugging Face, arXiv, OSF— resolvieron el problema de la invisibilidad técnica y la lentitud burocrática: un DOI en horas, visibilidad global, versiones y la posibilidad de ser citado.

Pero la velocidad sin metodología produce su propio fracaso. Con más de un millón y medio de modelos solo en Hugging Face ya en marzo de 2025, ningún filtro humano de fondo da abasto.

El dato más revelador viene de Carneiro et al. (2020): la diferencia en la calidad del reporte entre preprints y artículos revisados es de alrededor del 5%. Tras meses de proceso editorial, la mejora es pequeña. Lo que eso significa depende de cómo se lea: o el peer review no agrega tanto valor como proclama, o los repositorios abiertos no son tan caóticos como sus críticos arguyen. Si la diferencia en el reporte es de un 5%, la pregunta de cuánto aporta la revisión, frente a lo que cuesta, es legítima. Los propios autores concluyen que la mejora existe, aunque sea pequeña.

El ruido no es solo un problema de calidad. Es un problema de entrenamiento. Los modelos de lenguaje que alucinan más son los que operan sobre datos escasos, contradictorios o de baja calidad. Cuando los repositorios abiertos producen ruido a escala, ese ruido se convierte en señal para los sistemas que construyen conocimiento automatizado. El caos se escala.

2.3 El tercer régimen: la IA como compresión de abundancia

Academia clásica y repositorios abiertos parecen opuestos. Funcionalmente son mecanismos distintos para resolver el mismo problema: cómo filtrar información bajo límites humanos de atención y coordinación.

La academia gestiona incertidumbre mediante jerarquía, escasez, credenciales y exclusión. Los repositorios abiertos, mediante abundancia, iteración, exposición masiva y selección memética. Ambos son respuestas a condiciones históricas distintas. Ninguno fue diseñado para el escenario actual.

La IA entra como un tercer régimen epistemológico. No valida como la academia. No explora caóticamente como internet abierto. Sintetiza, recombina, comprime y acelera navegación cognitiva. Pero tiene un límite estructural crítico: su calidad depende de la calidad de lo que ingiere.

Un equipo de Stanford mostró que, al combinar un modelo de lenguaje con consultas verificables sobre Wikidata, se podía dar una respuesta útil al 96% de las preguntas del conjunto de desarrollo de un benchmark; en ese mismo conjunto la exactitud estricta fue de 76%, y de 65% en el de prueba (Xu et al., 2023). La IA no es autosuficiente epistémicamente. Es una máquina de compresión que amplifica lo que recibe, bueno o malo.

Esto introduce el concepto central de este artículo: costos de coordinación epistemológica. La crisis contemporánea del conocimiento no emerge simplemente de un exceso o ausencia de filtros, sino de la incapacidad institucional para reconciliar validación epistémica y velocidad cognitiva bajo condiciones de sobreproducción informacional. La IA reduce radicalmente esos costos de

coordinación. Las instituciones humanas, en cambio, los fueron acumulando hasta volverlos muy difíciles de sostener.

3. El candado circular

3.1 Wikidata como filtro operativo

En 2020, Google dijo que su Knowledge Graph reunía más de 500.000 millones de hechos sobre 5.000 millones de entidades, tomados de cientos de fuentes, con Wikipedia entre las más citadas (Sullivan, 2020). Wikidata es una de las bases abiertas que alimentan ese tipo de grafos.

Wikidata no es un repositorio de enciclopedia. Es parte de la arquitectura de grafos de conocimiento sobre la que operan los sistemas que construyen el mundo semántico que los LLMs ingieren. Sin nodo en Wikidata, a Google le cuesta más resolver la entidad. Sin entidad resuelta, las menciones dispersas en la web no se acumulan hacia ningún centro. El trabajo existe, pero para el sistema no es nadie.

Ese filtro está administrado en parte por editores anónimos con criterios de notabilidad del siglo XIX, aplicados con herramientas del XX, sobre fenómenos del XXI. Un editor sin nombre, sin expertise verificable en el dominio, puede borrar en minutos el trabajo técnico de un investigador independiente invocando criterios de notabilidad que no distinguen entre investigación sin filiación institucional y contenido sin valor, sin una vía real de réplica y sin que nadie responda por el impacto.

3.2 El mecanismo y sus consecuencias técnicas

La propuesta de reforma de la política de notabilidad de Wikidata, abierta en diciembre de 2025, lo reconoce: cierto conocimiento ha sido y sigue siendo marginado estructuralmente, lo que conduce a menor cobertura en fuentes confiables y, por tanto, a una barrera mayor para demostrar notabilidad (Wikidata, 2025).

Eso es el candado circular en su formulación más precisa: el sistema exige notabilidad para entrar, pero la notabilidad se construye con la visibilidad que el sistema niega al no entrar. El problema está en la estructura, no en un fallo puntual de gobernanza.

Las consecuencias técnicas son concretas y escalables. Sin nodo en Wikidata, entrar al Knowledge Graph de Google se vuelve más difícil, y con eso todo lo que viene después: el panel de conocimiento, que las menciones se acumulen en un mismo nodo, que los modelos reconozcan la entidad. El trabajo tiende a no aparecer en respuestas generadas por IA, y el círculo virtuoso de autoridad semántica no arranca.

Mi propio caso complica este argumento. Mi ítem en Wikidata fue borrado y, aun así, Google mantiene una entidad a mi nombre en su Knowledge Graph (comprobado el 12 de noviembre de 2026). No tengo registro de si esa entidad apareció mientras el ítem existía o después, así que no puedo saber si Wikidata fue la semilla y la entidad sobrevivió al borrado, o si Google la armó con otras señales. Lo que sí muestra es que Wikidata no es la única puerta. Por eso aquí digo que la ausencia del nodo dificulta la entrada, no que la impide.

El mismo periodo que vio expandirse este gatekeeping anónimo vio a ClaudeBot bloqueado en unos 5,8 millones de sitios según BuiltWith (Claburn, 2025), y una proporción estimada de 73.000 páginas rastreadas por cada visita devuelta en junio de 2025 (Cloudflare, 2025). Las máquinas tienen hambre de conocimiento verificado. Los sistemas que deberían alimentarlas las bloquean o las dejan sin nodos donde anclar lo que encuentran.

3.3 La invisibilidad técnica activa

Publicar en Scopus, en *Nature*, en Elsevier no es solo lento. En 2025 es, además, un acto de invisibilidad técnica progresiva. Las grandes editoriales académicas operan con hard paywalls que sirven al Googlebot solo metadatos o fragmentos iniciales. Según Tollbit, citado por *The Register*, los sitios que bloquean rastreadores de IA aumentaron un 336% en el año previo al segundo trimestre de 2025 (Claburn, 2025). Un estudio de Rutgers y Wharton encontró que los grandes editores de noticias que bloquearon los rastreadores de modelos de lenguaje perdieron alrededor de un 7% de visitas semanales en las seis semanas siguientes (Zhao y Berman, 2026).

Los repositorios universitarios completan el cuadro: URLs sin estructura, páginas huérfanas sin interlinks, sin schema markup, sin sitemaps optimizados, velocidades de carga del siglo anterior. El Googlebot tiene un presupuesto de rastreo. Es una economía de tiempo. No va a perderse en una arquitectura técnica de 2003. Se va donde la comida está servida: Zenodo, arXiv, OSF, GitHub Pages.

El investigador que publica solo en canales tradicionales no está siendo más riguroso. Está siendo invisible. Y en el ecosistema actual, invisible para las máquinas es invisible para el conocimiento que ya se construye a escala.

4. Los validadores ante el umbral

Este artículo no pide abolir la revisión de pares. Pide que los revisores estén a la altura del momento que guardan.

El día que la primera IA dijo hola cambió el umbral mínimo de lo que significa hacer ciencia con seriedad. Ese umbral subió para todos. Para los que investigan y para los que validan. El problema es que solo a unos se les está exigiendo notar que el piso se movió.

Un revisor de pares que no domina las herramientas con las que se produce el conocimiento que revisa no es un guardián del rigor. Es un cuello de botella con título. No por mala fe —la mayoría opera ad honorem, con carga académica propia y criterios heredados de un sistema que no actualizó sus estándares—. Por eso la exigencia no es individual sino estructural: los sistemas de validación deben incorporar criterios que incluyan si el trabajo puede confrontar la realidad que describe, si puede ser encontrado, si puede ser falsado por algo más que tres colegas en circuito cerrado.

Popper no pedía pares anónimos ni títulos institucionales. Pedía exposición real a la refutación. El sistema que hoy reclama esa herencia produce —según la encuesta de *Nature* a 1.576 investigadores (Baker, 2016)— que más del 70% haya intentado y fallado en reproducir el experimento de otro, y que más del 50% haya fallado en reproducir el suyo propio. El 52% considera que hay una crisis significativa de reproducibilidad. La exposición a la refutación ya se había debilitado antes de que llegara la IA, sin mucho ruido, dentro del mismo sistema que la reivindicaba. No desapareció: se desplazó.

Los repositorios abiertos, los benchmarks públicos, el código que o funciona o no funciona, aplican más presión popperiana real sobre una afirmación que ocho meses de peer review seguidos de un paywall que bloquea al Googlebot.

4.1 Donde la revisión lenta tiene que quedarse

Hay un terreno donde este argumento tiene que frenar. Lo dicho hasta aquí vale para el conocimiento que se puede revertir: un benchmark mal medido se corrige en la siguiente versión, un paquete roto se retira, un preprint equivocado se versiona. Para eso sirven la velocidad y las descargas: muchas veces lo prueba quien le encuentra una utilidad, lo descarta con la

misma velocidad y de paso deja un comentario. Esa es una exposición que la ciencia, en ciertos ámbitos, no puede ignorar. Pero un criterio diagnóstico, una intervención psicológica o un fármaco no funcionan igual. Un puente mal hecho o una nueva medida de seguridad tampoco. Ahí la revisión no es un lujo: exige que lo examinen, y lo destrocen si hace falta, quienes tienen la experiencia y las credenciales. Aquí las credenciales sí cuentan, porque quien firma con ellas responde por el daño; lo que fallaba en la sección 4 era el título sin dominio de las herramientas, no el título. El error no lo paga el sistema sino una persona, y para ella muchas veces no hay versión 2.

Conviene separar el conocimiento candidato, que puede circular rápido y exponerse a la refutación a la vista de todos, de la evidencia, que es lo que se usa para decidir sobre alguien. La pregunta que ordena esa frontera es quién absorbe el error cuando el error no se puede deshacer.

La psicología está en una posición incómoda en este punto. Es una de las disciplinas más golpeadas por la crisis de replicación (Open Science Collaboration, 2015) y, a la vez, de las que más necesitan una revisión cuidadosa, porque trabajan con personas. Para ella, la exigencia es otra: que la revisión que se hace sea real.

En los dos casos el rigor que se exige es el mismo: los estudios y las estadísticas tienen que ser igual de sólidos en el régimen lento y en el rápido. Lo que cambia en esta era es quién paga el error: en el terreno irreversible lo paga quien recibe la intervención; en el reversible, cada vez más, lo paga quien está recopilando esa información ahora mismo. Desde mi punto de vista, negarles la información a las inteligencias artificiales ya no es una opción realista; esa discusión está quedando atrás. Las IA nos buscan aunque intentemos cerrarles la puerta. En junio de 2025, Anthropic rastreaba unas 73.000 páginas por cada visita que devolvía (Cloudflare, 2025), y según Tollbit, citado por *The Register*, el 13,26% de las peticiones de bots de IA ignoró las instrucciones de robots.txt en el segundo trimestre de ese año (Claburn, 2025). Nos resumen y nos citan, pero casi nadie va a la fuente: en un análisis de 900 usuarios de Estados Unidos, solo en el 1% de las visitas a una página con resumen de IA alguien hizo clic en un enlace del propio resumen (Pew Research Center, 2025). Y no garantizan ni autorregulación ni una verificación previa de los datos. Así que la carga, con menos visibilidad, cae sobre los estudiantes que están aprendiendo a citar y que trabajan en sus primeras investigaciones, y que ya usan estas herramientas de forma masiva (Digital Education Council, 2024; Freeman, 2025). Las encuestas disponibles son del Reino Unido y de un estudio internacional; no tengo un dato equivalente para el Perú. Son ellos quienes van a heredar el problema.

No se trata de decir qué régimen es mejor. Se trata de por qué revisa cada uno: en el abierto se revisa por utilidad, en el clásico por obligación, y la IA no revisa: resume por defecto, sin motivo propio. Por eso el rigor no puede depender del régimen.

5. La obligación antes de que sea tarde

El peso algorítmico no distingue entre verdad y opinión. Distingue entre engagement y silencio. Un influencer con 10 millones de seguidores que opina sobre vacunas, economía o salud mental genera más señal en el grafo que un investigador con 40 años de trabajo clínico que publica en un repositorio que el Googlebot no visita. No porque el algoritmo sea malicioso. Porque el algoritmo optimiza lo que se le pidió que optimice: atención, tiempo en pantalla, interacción.

Si el conocimiento documentado con rigor no puede competir en ese grafo, el problema ya no es solo epistemológico. No estamos discutiendo si la academia es lenta. Estamos discutiendo quién va a entrenar a los sistemas que van a construir la realidad de las próximas generaciones. Si ese entrenamiento se hace con el peso algorítmico de opiniones sin método, sin falsabilidad, sin

registro verificable, el resultado no es un mundo más rápido. Es un mundo donde la diferencia entre verdad y narrativa desaparece no por ignorancia sino por arquitectura.

El conocimiento ya no circula solo por las instituciones humanas. Cada vez más pasa por algoritmos, y pesa más lo que esos algoritmos encuentran bien documentado. La IA tiene un tope real de comprensión y demasiada imaginación. Los LLMs alucinan más cuando las fuentes verificadas escasean o están bloqueadas. Si los humanos no dejan el registro bien documentado, las máquinas van a llenar ese vacío con la imaginación que les sobra. Y lo van a hacer con la velocidad que las instituciones nunca tuvieron.

Documentar bien no es vanidad académica: es la forma más segura que tenemos de que el conocimiento que producimos hoy sobreviva al algoritmo que decide mañana qué es real.

Limitaciones

Este trabajo constituye un ensayo analítico con revisión narrativa crítica de literatura verificable. No es un estudio empírico formal. La selección de fuentes introduce sesgo de selección inherente a toda revisión no sistemática. Algunos datos cuantitativos citados provienen de reportes de la industria cuya metodología exacta no es pública. El panorama tecnológico descrito evoluciona a velocidad que puede desactualizar aspectos específicos del análisis en el corto plazo.

El artículo es de posición, no una revisión neutral, y el tono lo refleja. Los ejemplos vienen sobre todo de la tecnología y el software, donde la velocidad juega a favor de la tesis; en otros campos el balance puede ser distinto. El concepto de costos de coordinación epistemológica todavía no está operacionalizado: aquí no propongo cómo medirlo, y mientras no se pueda medir tampoco se puede refutar.

Este texto es un preprint sin revisión por pares que critica la revisión por pares. La contradicción existe y no la escondo. La versión 2 corrige cifras y referencias que estaban mal en la primera; la versión 1 sigue disponible en Zenodo para quien quiera compararlas. Es el tipo de corrección a la vista que el artículo defiende.

Anexo. Verificación reproducible de las cifras del Reproducibility Project: Psychology

Qué es y qué no es. El cuerpo del artículo cita las tasas de replicación de la Open Science Collaboration (2015). Este anexo las recalcula desde los datos públicos del proyecto. No es un hallazgo nuevo, porque el RPP se ha reanalizado muchas veces. Es una comprobación hecha a la vista: lo que el artículo cita, el autor lo volvió a calcular, y el código y los datos permiten que cualquiera lo repita.

Datos. `rpp_data.csv`, componente «Analysis» del proyecto del RPP en OSF (https://osf.io/download/fgjvw/), licencia CC0 1.0. SHA-256 del fichero: `4d87a2636181d8e31230422854cf0fabeaecd2f1b47a59627482c05a2b56d54d`, que coincide con el que declara OSF. El fichero tiene 168 filas, de las cuales 100 son réplicas completas (`Completion (R) = 1`). Está codificado en latin-1, no en UTF-8.

Procedimiento. Dos programas en Python que solo usan la biblioteca estándar: `rpp_reanalisis.py` y `diag_p.py`. El primero se detiene si el SHA-256 de la entrada no coincide con el declarado y lista aparte las celdas vacías y los valores fuera de categoría, sin descartarlos en silencio. Se ejecutó dos veces, en dos entornos de Miniconda distintos: el entorno base, con Python 3.13.5, y uno nuevo con Python 3.15.2. El resumen de salida fue idéntico byte a byte en ambos (SHA-256 `07de851a5821c5ef798323c6d1211784c9e28e1866eff978a6070a700c36e8c6`).

Tabla A1. Resultados del recálculo frente a los valores publicados

Indicador	Recálculo	Publicado (OSC, 2015)
Réplicas con $p < .05$, entre originales con $p \leq .05$ recalculado	34 de 92 (37,0%)	36%
Ídem, con el denominador de OSC (originales con p reportado significativo)	34 de 97 (35,1%)	36% (35 de 97)
Juicio subjetivo de los equipos: replicó	39 de 100 (39,0%)	39%
Efecto original dentro del IC 95% de la réplica	44 de 94 (46,8%)	47%
r medio del original / de la réplica (97 pares)	0,396 / 0,197	0,403 (DT 0,188) / 0,197 (DT 0,257)
Mediana del cociente r réplica / r original, por estudio	0,43	—

Por disciplina (réplicas completas): la psicología cognitiva (n = 43) replicó entre 1,7 y 2,1 veces más que la social (n = 57), según el criterio:

- **Significancia:** 52,5% frente a 25,0%.
- **Juicio de los equipos:** 53,5% frente a 28,1%.
- **Efecto original dentro del IC:** 61,0% frente a 35,8%.

Lo que no coincide, explicado.

El denominador de significancia. El artículo cuenta 97 originales significativos; el criterio estricto (p recalculado $\leq .05$) cuenta 92. El diagnóstico identifica los 8 casos que quedan fuera:

- **Tres originales con p reportado no significativo:** $p > .1$, $p = .48$ y p_rep = .18. Son los que el artículo también excluye: $100 - 3 = 97$.
- **Cuatro originales reportados como significativos o marginales cuyo p recalculado queda apenas por encima de .05:** entre .0503 y .0516.
- **Un original reportado como $p < .001$** sin p recalculado en el fichero.

Ninguna de esas cinco réplicas tiene $p < .05$ en el fichero. Con el denominador de OSC, el recálculo da 34 de 97 (35,1%), frente al 36% publicado (35 de 97). Queda una réplica de diferencia. Hay una explicación probable, no verificada: el estudio 59, reportado como $p < .001$ en el original, tiene la réplica juzgada como exitosa por su equipo pero sin p en el fichero; si el artículo lo contó como réplica significativa, su numerador sería 35. Se reporta como hipótesis.

El estudio 135 muestra un p original recalculado de .91 y una réplica con $p = 3 \times 10^{-10}$. No afecta a los conteos, porque el original no fue significativo; no se investigó la causa.

La «mitad» del efecto. Comparando las medias, el efecto de las réplicas es el 50% del original, como dice el artículo. Por estudio, la mediana del cociente es 0,43. La «mitad» es, por tanto, la lectura más favorable.

El r medio del original. El recálculo da 0,396 con 97 pares, y el artículo publica 0,403 (verificado en el resumen estructurado del artículo en *Science*, consultado el 12-nov-2026). La diferencia probablemente viene de que no se usó el mismo conjunto de pares; no se verificó.

Disponibilidad. Los dos programas, la tabla de salida (`tabla_rpp.csv`) y el resumen (`resumen_rpp.txt`) acompañan a esta versión en el registro de Zenodo.

Referencia del conjunto de datos: Open Science Collaboration. (2015). *Reproducibility Project: Psychology* [conjunto de datos]. OSF. https://osf.io/ezcuj/

Referencias

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, *533*(7604), 452–454. https://doi.org/10.1038/533452a

Camerer, C. F., et al. (2018). Evaluating the replicability of social science experiments in *Nature* and *Science* between 2010 and 2015. *Nature Human Behaviour*, *2*(9), 637–644. https://doi.org/10.1038/s41562-018-0399-z

Carneiro, C. F. D., et al. (2020). Comparing quality of reporting between preprints and peer-reviewed articles in the biomedical literature. *Research Integrity and Peer Review*, *5*, 16. https://doi.org/10.1186/s41073-020-00101-3

Claburn, T. (2025, 8 de diciembre). Publishers say no to AI scrapers, block bots at server level. *The Register*. https://www.theregister.com/software/2025/12/08/publishers-say-no-to-ai-scrapers-block-bots-at-server-level/2064025

Cloudflare (Lee, J.-H., Gupta, D., Mitchell, B., Tatoris, R., y Clausen, H.). (2025, 1 de julio). *Control content use for AI training with Cloudflare's managed robots.txt and blocking for monetized content*. https://blog.cloudflare.com/control-content-use-for-ai-training/

Digital Education Council. (2024, 7 de agosto). *What students want: Key results from DEC Global AI Student Survey 2024*. https://www.digitaleducationcouncil.com/newsroom/what-students-want-key-results-from-dec-global-ai-student-survey-2024

Freeman, J. (2025). *Student generative AI survey 2025*. Higher Education Policy Institute y Kortext. https://www.hepi.ac.uk/reports/student-generative-ai-survey-2025/

Gartenberg, C., Hasan, S., Murray, A., y Pierce, L. (2026). More versus better: Artificial intelligence, incentives, and the emerging crisis in peer review. *Organization Science*, *37*(3), 795–812. https://doi.org/10.1287/orsc.2026.ed.v37.n3

González, A., et al. (2025). *How do pre-trained models support software engineering? An empirical study in Hugging Face*. arXiv:2506.03013. https://arxiv.org/abs/2506.03013

Maslej, N., et al. (2025). *AI Index Report 2025*. Stanford Institute for Human-Centered Artificial Intelligence. https://hai.stanford.edu/ai-index/2025-ai-index-report

NSF/NCSES. (2024). *Doctorate recipients from U.S. universities: 2022* (NSF 24-300). National Science Foundation. https://ncses.nsf.gov/pubs/nsf24300

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Pew Research Center. (2025, 22 de julio). *Google users are less likely to click on links when an AI summary appears in the results*. https://www.pewresearch.org/short-reads/2025/07/22/google-users-are-less-likely-to-click-on-links-when-an-ai-summary-appears-in-the-results/

Ravello Joo, C. E. (2026a). Metacognición 2.0: Diseño deliberado de identidad digital ante sistemas predictivos de inteligencia artificial. Zenodo. https://doi.org/10.5281/zenodo.20092009

Ravello Joo, C. E. (2026b). El perfil que no escribiste: Identidad digital, biometría ocular y autoridad semántica en sistemas algorítmicos predictivos. Zenodo. https://doi.org/10.5281/zenodo.20112409

Ravello Joo, C. E. (2026c). La década del influencer como colapso termodinámico de la información: Cierre operacional, alostasis del conocimiento y bifurcación del ecosistema digital (2012–2025). Zenodo. https://doi.org/10.5281/zenodo.20260383

Ravello Joo, C. E. (2026d). La inteligencia artificial como navaja suiza epistemológica: cuando la autoridad percibida reemplaza a la verificación. Zenodo. https://doi.org/10.5281/zenodo.24116983

Retraction Watch. (2025, 26 de abril). [Entrada del blog con el recuento de la base de datos]. https://retractionwatch.com/2025/04/26/

Schrödinger, E. (1944). *What is life?* Cambridge University Press.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Smith, R. (2006). Peer review: A flawed process at the heart of science and journals. *Journal of the Royal Society of Medicine*, *99*(4), 178–182. https://doi.org/10.1258/jrsm.99.4.178

Sullivan, D. (2020, 20 de mayo). *A reintroduction to our Knowledge Graph and knowledge panels*. Google. https://blog.google/products/search/about-knowledge-graph-and-knowledge-panels/

Topaz, M., Roguin, N., Gupta, P., Zhang, Z., y Peltonen, L.-M. (2026). Fabricated citations: An audit across $2 \cdot 5$ million biomedical papers [Correspondence]. *The Lancet*, *407*(10541), 1779–1781. https://doi.org/10.1016/S0140-6736(26)00603-3 (Fe de erratas: *The Lancet*, *408*(10551), 218.)

Wikidata. (2025, diciembre). *Requests for comment: Notability policy reform*. Wikimedia. https://www.wikidata.org/wiki/Wikidata:Requests_for_comment/Notability_policy_reform (consultado el 12 de noviembre de 2026)

Xu, S., Liu, S., Culhane, T., Pertseva, E., Wu, M.-H., Semnani, S. J., y Lam, M. S. (2023). Fine-tuned LLMs know more, hallucinate less with few-shot sequence-to-sequence semantic parsing over Wikidata. *Proceedings of EMNLP 2023*. arXiv:2305.14202. https://arxiv.org/abs/2305.14202

Zhao, H., y Berman, R. (2026). *Strategic response of news publishers to generative AI* (v4). arXiv:2512.24968. https://arxiv.org/abs/2512.24968