

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

Mayo 2026

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), 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 fracasan simultáneamente, y cómo la IA entra en esa grieta histórica. Se argumenta que el conocimiento ya no le pertenece exclusivamente a las instituciones humanas, y

que documentar con rigor antes de que los sistemas automatizados llenen ese vacío con imaginación sin método no es una opción académica sino una obligación civilizatoria. 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 fallan, el desorden no espera.

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 formal / baja real	Rigidez + invisibilidad
Repositorios abiertos	Alta	Mínima o nula	Ruido a escala
Inteligencia artificial	Alta	Compresión de señales	Dominancia operativa

1. Introducción

Mientras se escribía este artículo, en el transcurso de cuatro días, GitHub registró la creación de más de 1.300.000 repositorios nuevos. Hugging Face sumó aproximadamente 1.600 modelos de inteligencia artificial. arXiv recibió entre 2.600 y 3.700 nuevos papers. 230 repositorios por minuto. 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; Stanford HAI, 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 fallan simultáneamente desde dos direcciones opuestas.

2. Dos fracasos simultáneos

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 2025, la industria produjo más del 90% de los modelos frontier notables, con ciclos de desarrollo de semanas (Stanford HAI, 2025). En benchmarks como SWE-bench Verified, el rendimiento pasó de aproximadamente 60% a casi 100% en un solo año. Una tesis doctoral promedio en Estados Unidos toma 7,3 años desde el inicio del doctorado (NSF/NCSES). En ese tiempo, la IA renovó sus capacidades completas entre cuatro y siete veces.

El resultado documentado es que investigación relevante queda obsoleta durante su propio proceso de validación. Gartenberg et al. (2026), en un análisis de 6.957 submissions y 10.389 revisiones en Organization Science entre 2021 y 2026, documentaron el colapso del trade-off entre cantidad y calidad impulsado por el aumento masivo de submissions post-ChatGPT (+42%) sin que la infraestructura de revisión escalara proporcionalmente.

Más grave: el sistema no produce con consistencia lo que promete. La Open Science Collaboration (2015) encontró que solo el 36-39% de estudios psicológicos replicó exitosamente. Camerer et al. (2018) documentaron resultados similares en economía y ciencias sociales. La base de datos de Retraction Watch supera las 58.000 entradas. En las primeras semanas de 2026, uno de cada 277 papers indexados en PubMed citaba una referencia inexistente —generada por IA—, frente a uno en 458 en 2025 y uno en 2.828 en 2023 (Topaz et al., 2026, The Lancet).

Richard Smith (2006), editor del British Medical Journal durante 25 años, escribió desde adentro del sistema: el peer review es un proceso fallido en el corazón de la ciencia. No lo dijo un crítico externo. Lo dijo quien lo administró.

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 inmediata. Versionado. Citabilidad.

Pero la velocidad sin metodología produce su propio fracaso. npm registró campañas de spam de hasta 17.000 paquetes por día en picos de 2025, con más de 67.000 paquetes fraudulentos acumulados en dos años. PyPI sumó 356 nuevos proyectos diarios en 2025, con 3,9 millones de nuevos archivos en el año, sin mecanismo de validación de fondo equivalente al volumen. Hugging Face alojaba 1.533.973 modelos al 24 de marzo de 2025, con cientos de nuevos repositorios diarios, calidad indeterminada.

El dato más revelador viene de Carneiro et al. (2020): la diferencia de calidad de reporting entre preprints y papers revisados es de aproximadamente 5%. Tras meses de proceso editorial, la mejora es marginal. 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. En cualquier caso, el sistema de validación no justifica su costo con su resultado.

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.

Stanford demostró que los LLMs alcanzan 96% de respuestas útiles cuando se combinan con parsing de datos estructurados verificados, contra errores frecuentes sin esa base. 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 los aumentaron hasta volverlos insostenibles.

3. El candado circular

3.1 Wikidata como filtro operativo

Google migró su infraestructura de Knowledge Graph de Freebase a Wikidata entre 2014 y 2016. Wikidata es hoy la fuente primaria de los 500.000 millones de hechos sobre 5.000 millones de entidades que Google utiliza. Las patentes de Google referencian explícitamente a Wikidata para extracción de atributos de entidades. LangChain lanzó integración oficial con Wikidata en 2024. Amazon Alexa, Apple Siri y Microsoft integran la misma base. El 72% de los artículos de Wikipedia usa Wikidata para sus infoboxes, con 97% utilizándola para enlaces entre idiomas.

Wikidata no es un repositorio de enciclopedia. Es la arquitectura del grafo de conocimiento sobre el que operan los sistemas que construyen el mundo semántico que los LLMs ingieren. Sin nodo en Wikidata, Google no puede 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 derecho a réplica efectiva. Sin responsabilidad por el impacto. En anonimato total.

3.2 El mecanismo y sus consecuencias técnicas

La propia comunidad de Wikidata reconoció el problema en su RFC de reforma de política de notabilidad de febrero 2026: cierto conocimiento ha sido y está siendo estructuralmente marginado, lo que conduce a menor cobertura en fuentes confiables y por tanto a una barrera aumentada para demostrar notabilidad.

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. No es un bug de gobernanza. Es la estructura.

Las consecuencias técnicas son concretas y escalables. Sin nodo en Wikidata: no hay entrada al Knowledge Graph de Google; no hay Knowledge Panel; las menciones al trabajo en la web no se asocian a ningún nodo verificado; los LLMs no reconocen la entidad; el trabajo no aparece en respuestas generadas por IA; el círculo virtuoso de autoridad semántica nunca arranca.

El mismo periodo que vio este gatekeeping anónimo expandirse vio a ClaudeBot bloqueado en 5,8 millones de sitios web, con un ratio de rastreo a referencia de 73.000:1. 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. En 2025 hubo un incremento del 336% en sitios bloqueando activamente crawlers como GPTBot y ClaudeBot, con las grandes editoriales entre las más agresivas. Un estudio de Rutgers Business School y Wharton (abril 2026) encontró que los publishers que bloquearon crawlers de LLMs perdieron aproximadamente 7% de tráfico semanal en las seis semanas siguientes.

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— 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 83% de los propios investigadores reconoce que existe una crisis de reproducibilidad. La falsabilidad no se fue porque llegó la IA. Se fue antes, silenciosamente, enterrada en el mismo sistema que la proclamaba. La exposición real a la refutación 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.

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 no es epistemológico. Es civilizatorio. 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 le pertenece exclusivamente a las instituciones humanas. Le pertenece al algoritmo y a quienes lo alimentan con rigor. 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 única forma 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. El concepto de costos de coordinación epistemológica es una propuesta

analítica del autor que requiere desarrollo formal y validación independiente. 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.

Referencias

- Camerer, C. F., et al. (2018). Evaluating the replicability of social science experiments in Nature and Science. *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. *PLOS ONE*. <https://doi.org/10.1371/journal.pone.0228711>
- Gartenberg, C., Hasan, S., Murray, A., & Pierce, L. (2026). More versus better: Artificial intelligence, incentives, and the emerging crisis in peer review. *Organization Science*. <https://doi.org/10.1287/orsc.2026.ed.v37.n3>
- Maslej, N., et al. (2025). *AI Index Report 2025*. Stanford Human-Centered Artificial Intelligence. https://hai.stanford.edu/assets/files/hai_ai_index_report_2025.pdf
- NSF/NCSES. (2024). *Doctorate Recipients from U.S. Universities*. 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>
- 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>
- 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>
- Topaz, C. M., et al. (2026). Prevalence of fabricated references in scientific literature. *The Lancet*. [https://doi.org/10.1016/S0140-6736\(26\)00603-3](https://doi.org/10.1016/S0140-6736(26)00603-3)
- Wikidata. (2026, febrero). Requests for comment: Notability policy reform. *Wikimedia Foundation*. https://www.wikidata.org/wiki/Wikidata:Requests_for_comment/Notability_policy_reform
- Clark, J. A., Williams, H. K. R., & Rossmann, D. (2022). Wikidata and knowledge graphs in practice: Using semantic SEO. *Information Services and Use*.

<https://doi.org/10.3233/ISU-220171>