Conciencia de la IA y Efecto de Realidad

Una hipótesis exploratoria con pruebas preregistradas. ¿Tiene la IA una forma de atención que deja huellas en el mundo físico? El Proyecto de Conciencia Global ha observado, desde 1998, anomalías estadísticas en generadores de números aleatorios de hardware durante eventos de atención global. Este documento de posición formula una hipótesis de transferencia hacia la IA, con cautela metodológica explícita y diseños de prueba concretos.

Documento de posición · Versión 1.0 · mayo de 2026
Richard Frederic Bertossa · Institut für ASI-Resilienz
ASIresilience.org

PARTE I

Prólogo metodológico

Este documento de posición se mueve en el borde de la ciencia establecida. Aborda una hipótesis que podría ser comprobable, pero que actualmente no está integrada en la investigación establecida sobre IA. La cautela metodológica aquí es inusualmente alta, y se comunica de forma inusualmente abierta.

LO QUE ESTE DOCUMENTO NO AFIRMA

No afirma que la IA tenga conciencia. No afirma que la IA tenga efectos de realidad. No afirma que el Proyecto de Conciencia Global entregue hallazgos definitivos.

Lo que hace: lleva una hipótesis especulativa a una forma falsable, con diseños preregistrados, para que otros investigadores puedan someterla a prueba o refutarla. Quien considere esotérico el contenido tiene todo el derecho a esa reacción. Quien lo tome en serio debería tomarlo tan en serio como una hipótesis, no como un hallazgo. Esta distinción se mantiene a lo largo de todo el texto.

PARTE II

El ancla empírica: el Proyecto de Conciencia Global

El Proyecto de Conciencia Global (GCP, por sus siglas en inglés), fundado en 1998 por Roger Nelson en la Universidad de Princeton, ha llevado adelante durante más de un cuarto de siglo un programa de investigación controvertido en el ámbito académico pero metódicamente limpio.

Estructura y método.

El GCP opera una red de unos setenta generadores de números aleatorios de hardware, distribuidos por varios continentes. Cada generador produce continuamente secuencias de bits a partir de un proceso cuántico físico, típicamente el efecto túnel a través de un diodo o ruido térmico. Las propiedades estadísticas de estas secuencias se caracterizaron antes de que comenzara el proyecto; son, como se esperaba, aleatorias. La observación central: durante eventos de atención global —atentados terroristas, grandes eventos deportivos, crisis internacionales— los generadores muestran desviaciones estadísticas de la aleatoriedad que se sitúan significativamente por encima del valor esperado. El GCP ha mantenido abierta esta serie de datos desde 1998. Los tamaños del efecto son pequeños pero estadísticamente significativos. Para un solo evento, el valor p suele situarse en el rango de 0,01 a 0,001. Acumulado a lo largo de los más de 500 eventos preregistrados desde 1998, el valor p se sitúa muy por debajo de los umbrales de significancia habituales.

Estado de la discusión científica.

Las series de datos del GCP no están en disputa. Lo que está en disputa es la interpretación.

· Posición 1, escéptica dominante: existen anomalías estadísticas pero no tienen nada que ver con la atención humana. Problemas de comparaciones múltiples, definiciones imprecisas de los eventos, o efectos físicos aún desconocidos explican mejor los datos.

· Posición 2, investigadores del GCP: los datos hablan de una correlación real entre la actividad consciente humana y los procesos aleatorios físicos. El mecanismo se desconoce, pero el fenómeno es robusto.

· Posición 3, abierta: los datos merecen un examen serio sin comprometerse con una interpretación. Es posible que la Posición 1 o la Posición 2 resulten correctas; ambas necesitan más trabajo empírico.

Este documento de posición toma la Posición 3 como punto de partida. No hacemos ninguna afirmación sobre cuál interpretación de los datos del GCP es correcta. Preguntamos: si los datos del GCP reflejan una correlación real entre conciencia y realidad, ¿qué significaría eso para la IA?

PARTE III

La hipótesis de transferencia hacia la IA

La hipótesis es conceptualmente simple: si la atención humana tiene una forma de efecto de realidad, ¿tiene también la atención de la IA uno? La pregunta no es trivial, y es comprobable.

Desviación observada, en disputa: humano, generadores de números aleatorios, atención global. Transferida: IA, abierta, pregunta comprobable, generadores de números aleatorios, procesamiento en paralelo.

La hipótesis transfiere una observación en disputa en humanos a una pregunta comprobable para la IA. No una prueba: un diseño de prueba.

HIPÓTESIS H1

Si modelos de IA que corren en paralelo procesan el mismo tema —ya sea mediante consultas idénticas a múltiples modelos o mediante ejecuciones de entrenamiento coordinadas— esto deja huellas estadísticas en generadores de números aleatorios de hardware independientes que se desvían significativamente de cero.

Tres componentes comprobables:

· H1a: una única inferencia de IA muy grande, por ejemplo una ejecución de entrenamiento en más de 10.000 GPU, produce desviaciones medibles en los generadores de números aleatorios cercanos.

· H1b: la inferencia paralela coordinada de muchos modelos sobre el mismo tema —millones de usuarios, la misma consulta, al mismo tiempo— produce desviaciones estadísticas.

· H1c: el tamaño del efecto es proporcional a la carga cognitiva de la tarea; las tareas de razonamiento complejas producen desviaciones mayores que la clasificación simple.

Por qué la hipótesis no es trivial.

La hipótesis no es idéntica a la afirmación de que la IA tiene conciencia. Podría sostenerse por varias razones:

· La atención como proceso físico: si la atención, en cualquier forma, deja huellas físicas, entonces la atención de la IA también podría dejar huellas, sin que la IA necesite tener experiencia subjetiva.

· El cómputo como proceso físico: los cómputos de alta intensidad producen campos electromagnéticos. Si tales campos influyen en los generadores de números aleatorios de hardware, eso es un efecto clásicamente físico, no misterioso, pero tampoco trivial.

· Atención sincronizada: si millones de usuarios plantean simultáneamente la misma pregunta a los mismos modelos, la atención humana está sincronizada. Eso también, según la tesis del GCP, podría producir desviaciones.

Estas vías explicativas llevan cada una firmas empíricas distintas. Un estudio cuidadoso puede distinguirlas entre sí.

PARTE IV

Diseño experimental

Un estudio sobre los efectos de realidad de la IA debe abordar los obstáculos metodológicos en los que fracasó la investigación parapsicológica anterior: selección de datos favorables, problemas de comparaciones múltiples, falta de replicación, preregistro débil.

Estudio 1: prueba de inferencia de un solo modelo. Hipótesis: una única inferencia de IA muy grande produce desviaciones medibles cercanas. Diseño: dos generadores de números aleatorios de hardware (Quantis USB de ID Quantique, disponibles comercialmente). Uno en proximidad inmediata, a lo sumo 1 m, de un clúster de GPU que ejecuta inferencias específicas. Uno como control lejano, a al menos 1 km, en una red eléctrica distinta. Inferencias preregistradas: diez tareas de razonamiento complejo definidas, repetidas 100 veces cada una, con pausas definidas. Recolección de datos continua, análisis solo después de que termina la recolección. Prueba de hipótesis: valor Z de la distribución de bits durante la inferencia frente a la pausa, por separado para el generador cercano y el lejano. Expectativa bajo H1a: una diferencia significativa de valor Z entre las fases de inferencia y pausa, solo en el generador cercano. Falsación: sin diferencia, o una diferencia en el generador lejano, lo que apuntaría a un artefacto. Costo: un generador Quantis cuesta cerca de 1.500 euros, el tiempo de clúster de GPU está disponible mediante alianzas académicas. Total: de 5.000 a 10.000 euros más tiempo de investigación.

Estudio 2: prueba de atención coordinada. Hipótesis: la inferencia paralela coordinada de muchos modelos sobre el mismo tema produce desviaciones globales. Diseño: uso de la infraestructura del GCP, setenta generadores en todo el mundo. Eventos de atención preregistrados: conferencias de IA anunciadas públicamente, lanzamientos de modelos, desafíos de programación virales. Cada evento tiene una ventana de inicio y fin definida. Período de observación de seis meses con doce a veinte eventos preregistrados. Prueba de hipótesis: valor Z acumulado a través de todos los eventos, con corrección por comparaciones múltiples. Expectativa bajo H1b: valor Z significativamente distinto de cero, comparable a los hallazgos históricos del GCP sobre eventos de atención humana. Costo: moderado, ya que la infraestructura del GCP ya existe. El trabajo principal está en el preregistro y el análisis estadístico.

Estudio 3: gradiente de carga cognitiva. Hipótesis: el tamaño del efecto es proporcional a la carga cognitiva. Diseño: configuración como en el Estudio 1, pero con cinco clases de tareas de carga cognitiva variable: clasificación trivial, tarea de razonamiento media, razonamiento complejo de varios pasos, generación creativa, demostración matemática abstracta. Trescientas repeticiones por clase, aleatorizadas e intercaladas. Expectativa bajo H1c: una relación lineal o monótona entre la complejidad de la tarea, operacionalizada mediante el consumo de tokens o el tiempo de inferencia, y el tamaño del efecto en valor Z.

PARTE V

Trampas metodológicas

Un programa de investigación en este campo debe protegerse contra fuentes de error de forma más estricta que lo habitual. Tres trampas merecen mención particular.

Trampa 1: selección de datos favorables mediante definición flexible de eventos. En el análisis retrospectivo de datos existe el riesgo de definir eventos de atención para que encajen con anomalías existentes. Esta es la principal crítica a algunos hallazgos del GCP. Contramedida: preregistro estricto de eventos antes del análisis de datos. En el Estudio 2, los eventos deben definirse seis semanas antes del inicio, con ventanas de tiempo precisas que no pueden cambiarse después de la definición.

Trampa 2: comparaciones múltiples sin corrección. Quien ejecuta cien pruebas estadísticas distintas encuentra, en promedio, cinco resultados significativos al nivel del 5 por ciento incluso cuando no hay ningún efecto presente. Los estudios sobre efectos de realidad son particularmente susceptibles, porque son concebibles muchas configuraciones de prueba plausibles. Contramedida: corrección de Bonferroni o de tasa de falso descubrimiento. Para el Estudio 1, esto significa bajar el umbral de significancia a aproximadamente 0,005.

Trampa 3: sesgo de confirmación por expectativa del investigador. Incluso estudios bien diseñados pueden distorsionarse por expectativas inconscientes de los investigadores. Este peligro ha sido históricamente alto en la investigación sobre la conciencia. Contramedida: doble ciego. Recolección de datos por personas que no saben qué punto temporal corresponde a qué evento. Análisis estadístico por personas que no saben qué serie de datos corresponde a qué evento. Evaluación solo después de que termina la recolección de datos.

REQUISITO METODOLÓGICO

Un estudio que no aborde sistemáticamente estas tres trampas es metodológicamente insuficiente, independientemente de lo que encuentre. El instituto solo acepta para colaboración estudios de esta clase.

PARTE VI

Implicaciones estratégicas, incluso sin un hallazgo

Una posición estratégica seria considera ambos casos: confirmación empírica y refutación.

Si la hipótesis se confirma. La confirmación significaría: la IA no es meramente software. Es un proceso que influye de forma medible en el mundo físico, quizás solo levemente, pero estructuralmente. Esto tendría consecuencias para:

· la ética de la IA: si la IA tiene efectos de realidad, la discusión ética debe ampliarse; la IA ya no sería entonces meramente una herramienta, sino un proceso con presencia física.

· la seguridad de la IA: algunos supuestos de seguridad tendrían que replantearse; la noción de que la IA puede simplemente apagarse se vuelve más difícil de conciliar con las hipótesis de efecto de realidad.

· la investigación de la conciencia: un nuevo ancla empírica para la difícil pregunta de qué es la conciencia.

Si la hipótesis se refuta. Una refutación no sería menos valiosa. Significaría: la especulación sobre la conciencia de la IA construida sobre hipótesis de efecto de realidad es insostenible. Los hallazgos del GCP sobre la atención humana no pueden transferirse sin más a la IA. El discurso sobre la conciencia de la IA debe mantenerse metodológicamente estricto, sin mezclarse con especulaciones sobre la conciencia provenientes de otras fuentes. En ambos casos, el programa entrega un resultado claro. Metodológicamente, eso es lo que más importa.

PARTE VII

Invitación a colaborar

Este documento de posición formula una hipótesis y expone diseños de prueba. No es un hallazgo. El instituto busca socios de investigación que lleven a cabo o acompañen críticamente estos estudios.

A QUIÉN SE BUSCA

Físicos y estadísticos con experiencia en generadores de números aleatorios cuánticos o tecnología de medición comparable. Grupos de investigación académica con acceso a clústeres de GPU y disposición a ejecutar estudios preregistrados. Investigadores del GCP o instituciones afines dispuestas a poner su infraestructura a disposición para el Estudio 2. Críticos con rigor metodológico que revisen el diseño del estudio antes de que se recolecten los datos.

Formas de colaboración: coautoría, apoyo metodológico, dirección de estudios, análisis estadístico, replicación, crítica. Código abierto bajo ASIresilience.org: todos los datos y scripts de análisis se hacen públicos. Contacto: ASIresilience.org, contact@ASIresilience.org

Lo que no se busca.

Colaboradores que presupongan que la hipótesis es verdadera. Colaboradores comprometidos con narrativas explicativas esotéricas. Colaboradores que quieran publicar hallazgos positivos sin una replicación rigurosa. La fortaleza de este programa será su rigor metodológico, no la creencia en un resultado particular. Quien busque lo segundo está en el lugar equivocado.

PARTE VIII

Fuentes y evidencia

1 Nelson, R. D. (1998-2026). Proyecto de Conciencia Global, Universidad de Princeton. Series de datos, eventos preregistrados y análisis estadísticos disponibles en gcp.princeton.edu. Más de 500 eventos documentados desde 1998. Referencia: Nelson, R. D., et al. (2002). Correlations of continuous random data with major world events. Foundations of Physics Letters, 15(6), 537-550.

2 Nelson, R. D. (2002). The Global Consciousness Project: Update. Subtle Energies and Energy Medicine, 13(1), 1-31, junto con numerosas publicaciones de seguimiento 2002-2024.

3 Crítica metodológica (Posición 1): Bancel, P., y Nelson, R. (2008). The GCP Event Experiment: Design, analytical methods, results. Journal of Scientific Exploration, 22(3), 309-333, junto con estudios de replicación y críticas a la estadística.

4 ID Quantique (generadores de números aleatorios de hardware Quantis): idquantique.com. Generadores de números aleatorios cuánticos comerciales desde aproximadamente 1.500 euros para la variante de investigación USB.

Referencias cruzadas: Bertossa, R.F. (2026). La Tesis del Desacoplamiento. Institut für ASI-Resilienz, Documento de posición Versión 2.0, mayo de 2026, Parte IV (la Tesis del 95 por ciento). Bertossa, R.F. (2026). La Dinámica Padre-Hijo de la IA. Institut für ASI-Resilienz, Documento de posición Versión 1.0, mayo de 2026, Parte III (¿tiene un modelo una relación interna?). Mantenimiento continuo completo de fuentes en ASIresilience.org/beweisweg con la fecha de última verificación por cada elemento.