

KNO-021 · Ciencia abierta y reproducibilidad

KNO · Formas de conocer

Juan Seguro

v0.3 · ES

Todos los derechos reservados

Resumen del artículo

Las afirmaciones científicas resultan más fáciles de inspeccionar cuando son accesibles los materiales necesarios para comprenderlas, repetirlas o discutirlos. La ciencia abierta reúne prácticas que intentan hacer posible esa inspección: compartir datos y código cuando corresponde, describir los métodos con precisión, distinguir análisis exploratorios de confirmatorios, conservar versiones y hacer visibles las correcciones. Reproducibilidad y replicabilidad están relacionadas, pero no son idénticas. Un resultado puede ser reproducible computacionalmente y seguir siendo erróneo, mientras que un estudio independiente puede no replicar por razones que no se reducen a mala conducta o incompetencia. Este artículo pregunta qué partes de un resultado científico deben seguir siendo inspeccionables para que investigadores posteriores reconstruyan el análisis, prueben la afirmación con nuevos datos y localicen dónde entra el desacuerdo. Trata la apertura como infraestructura para la crítica, no como certificado de verdad.

En 60 segundos

La ciencia abierta no es un interruptor llamado «compartirlo todo». Es un conjunto de controles que vuelve más inspeccionable el trabajo científico. La reproducibilidad suele preguntar si los mismos datos y procedimientos computacionales producen el mismo resultado; la replicabilidad pregunta si nuevos datos obtenidos para la misma cuestión generan resultados compatibles dentro de la incertidumbre pertinente.[1,2] El prerregistro puede separar predicciones de análisis escogidos tras ver los datos, pero no convierte un diseño débil en uno fuerte.[3] Los informes registrados trasladan parte de la evaluación a antes de conocer los resultados, mientras que los datos y el código abiertos permiten inspeccionar transformaciones, exclusiones e implementación. [4] Ninguna de estas prácticas garantiza la verdad. Su valor es procedimental: exponen más del camino entre observación y afirmación, conservan alternativas que de otro modo desaparecerían y abaratan la corrección. La apertura también tiene límites: privacidad, seguridad, consentimiento, propiedad intelectual y carga de curación pueden justificar acceso controlado. La meta es ciencia auditable, no divulgación sin fricción.

Por qué importa esta pregunta

La ciencia moderna depende a menudo de conjuntos de datos, scripts, instrumentos y decisiones que no aparecen en el artículo final. Cuando solo sobrevive la narración, investigadores posteriores pueden no saber si un desacuerdo nace de los datos, el preprocesamiento, las decisiones estadísticas, el software, la interpretación o una diferencia real en la naturaleza. Las prácticas de ciencia abierta intentan conservar esos puntos de decisión. Importan porque la crítica es uno de los mecanismos de corrección de errores de la ciencia, pero solo puede operar sobre lo inspeccionable. Las mismas prácticas revelan además una tensión: la apertura puede aumentar la auditabilidad y, al mismo tiempo, crear nuevas obligaciones de privacidad, gobernanza y conservación a largo plazo.

1. Reproducibilidad no es replicación

La pregunta Si dos investigadores obtienen el mismo resultado numérico, ¿qué se ha reproducido exactamente: el cálculo, la observación o la afirmación? La respuesta depende de qué entradas, procedimientos y condiciones permanecen fijos y cuáles pueden cambiar durante la nueva prueba.

Las National Academies distinguen reproducibilidad de replicabilidad para ganar claridad práctica. La reproducibilidad consiste en obtener resultados consistentes con los mismos datos de entrada, pasos computacionales, métodos y condiciones de análisis; la replicabilidad se refiere a nuevos estudios sobre la misma pregunta con datos obtenidos de nuevo.[1] Goodman, Fanelli e Ioannidis muestran igualmente que el vocabulario se ha usado de forma inconsistente entre campos y advierten contra tratar «reproducible» como sinónimo de «verdadero».[2]

La distinción separa dos modos de fallo. Si los mismos archivos y código no regeneran una tabla publicada, el problema puede ser una dependencia ausente, preprocesamiento no documentado, deriva de versiones o error de implementación. Si un estudio nuevo discrepa, la explicación puede estar en muestreo, medición, contexto, heterogeneidad del efecto o una afirmación original falsa. El diagnóstico empieza preguntando qué tipo de igualdad se esperaba.

Afirmación C01. La reproducibilidad y la replicabilidad ponen a prueba formas distintas de consistencia. [1,2]

Afirmación C10. La reproducibilidad no debe usarse como sinónimo de verdad. [2]

En qué punto estamos

Reproducibilidad y replicación prueban eslabones distintos de la cadena evidencial. Repetir un cálculo antiguo comprueba si el análisis comunicado puede reconstruirse; obtener nuevos datos prueba si la afirmación sobrevive a un nuevo encuentro con el mundo. Ningún resultado es certificado de verdad. La pregunta siguiente es qué debe conservarse para que la reconstrucción sea posible.

2. Un artículo no es todo el análisis

Un artículo comprime un proceso mucho mayor. Las observaciones brutas se seleccionan, calibran, limpian, transforman y resumen antes de convertirse en figura o coeficiente. Las bibliotecas de software aportan valores por defecto; las reglas de exclusión eliminan casos; los archivos intermedios conservan estado. Cuando esas capas desaparecen, la publicación puede seguir siendo legible mientras el análisis deja de ser reconstruible.

Un trabajo reproducible necesita por tanto más que prosa. Según el campo, puede requerir datos o un sustituto de acceso controlado, código, información del entorno, parámetros, registros de calibración y un mapa que conecte salidas con entradas. No se trata de conservar cada pulsación, sino de preservar suficiente grafo de dependencias para que otro investigador competente identifique cómo se produjo el resultado.[1,4]

Afirmación C02. La reproducibilidad puede fallar cuando los datos, el código, las dependencias o las condiciones del análisis no pueden reconstruirse. [1]

En qué punto estamos

El objeto científico sometido a revisión suele ser mayor que el artículo que lo describe. Datos, código, parámetros y procedencia pueden ser responsables del resultado aunque no ocupen ninguna página impresa. Conservarlos vuelve inspeccionables decisiones ocultas. Eso conduce al siguiente problema: algunas decisiones deben registrarse antes de conocer el resultado, no reconstruirse después.

3. El prerregistro separa predicción y retrospección

Los datos pueden sugerir hipótesis después de ser observados. Eso es científicamente útil, pero no equivale a probar una predicción que existía antes de la observación. Nosek, Ebersole, DeHaven y Mellor describen el prerregistro como una forma de hacer visible esa diferencia temporal registrando hipótesis y planes de análisis antes de inspeccionar los resultados.[3]

El prerregistro no elimina la exploración ni vuelve correcto un análisis. Los planes pueden ser imprecisos, las mediciones pueden fallar y pueden ser necesarias desviaciones justificadas. Su valor es que un lector posterior puede distinguir qué decisiones fueron prospectivas y cuáles respondieron a los datos. Eso cambia cómo se interpreta la fuerza evidencial sin prohibir el descubrimiento.

Afirmación C03. El prerregistro puede hacer visible el momento en que se formularon las hipótesis y se eligieron los análisis. [3]

Afirmación C04 - Síntesis declarada. El prerregistro no establece por sí solo la calidad del diseño ni la verdad. **Síntesis declarada:** [3] respalda el prerregistro como compromiso procedimental fijado antes de conocer los resultados; la cláusula adicional de que no establece por sí mismo la calidad del diseño ni la verdad es la síntesis.

En qué punto estamos

El prerregistro es una distinción con marca temporal, no un sello de calidad. Conserva cuándo entró una hipótesis o decisión analítica en el flujo de trabajo, permitiendo separar predicción de adaptación dependiente de los datos. Puede reducir la ambigüedad del estatus evidencial sin deslegitimar la exploración. El siguiente paso lleva esa lógica a la publicación: ¿puede empezar la evaluación antes de conocer el resultado?

4. Los informes registrados cambian el orden de la evaluación

Un artículo convencional suele revisarse cuando ya existen los resultados principales. Los informes registrados adelantan una parte importante de la revisión: la pregunta, el diseño y el análisis previsto pueden evaluarse antes de conocer el desenlace. Munafò y colegas incluyen este mecanismo entre reformas destinadas a reducir oportunidades de comunicación selectiva y narración retrospectiva.[4]

Cambiar el orden no garantiza una pregunta útil ni una ejecución impecable. Cambia incentivos y el registro evidencial. Un resultado nulo procedente de un diseño aceptado en principio tiene menos riesgo de desaparecer por ser poco llamativo, mientras que las desviaciones del plan pueden documentarse en lugar de ocultarse. La publicación queda más ligada a la calidad de pregunta y método que a la dirección del resultado.

Afirmación C05. Las reformas del tipo informe registrado desplazan parte de la evaluación de preguntas y métodos a una fase anterior del proceso de investigación. [4]

En qué punto estamos

Los informes registrados muestran que la apertura también trata de cronología: conservar qué se decidió antes de que los resultados pudieran influir en esas decisiones. Reducen una vía por la que el registro publicado puede quedar filtrado selectivamente, pero no eliminan todos los filtros. La pregunta siguiente es qué ocurre después de publicar, cuando otro investigador necesita acceder al material que hizo posible el resultado.

5. Datos y código abiertos vuelven inspeccionables las transformaciones

Compartir datos permite comprobar codificación, exclusiones, transformaciones y análisis alternativos. Compartir código puede mostrar cómo se calculó realmente un número publicado, no solo cómo se describió el método en prosa. Juntos reducen la distancia entre la afirmación y las operaciones que la produjeron.[1,4]

Pero «abierto» no basta. Un archivo sin definiciones de variables, procedencia o versión puede ser descargable y científicamente inútil. El código que depende de versiones de software no documentadas puede fallar años después. La apertura necesita curación: identificadores, metadatos, licencias, información de dependencias y estrategias de conservación. KNO-019 explicó por qué los registros estructurados necesitan contexto; la ciencia abierta añade que ese contexto debe seguir accesible a la crítica.

Afirmación C06. Los datos y el código abiertos pueden exponer a escrutinio transformaciones y decisiones de implementación. [1,4]

Afirmación C07 - Síntesis declarada. El acceso abierto sin metadatos, procedencia y contexto de versión puede ser científicamente insuficiente. **Síntesis declarada:** [1] respalda la importancia de metadatos, procedencia y contexto de versión para la reutilización; el juicio de que el acceso por sí solo puede ser científicamente insuficiente es la síntesis.

En qué punto estamos

Los materiales abiertos extienden el objeto inspeccionable desde el artículo hasta el flujo de trabajo, pero el acceso sin contexto es una infraestructura débil. Datos y código necesitan metadatos, versiones y procedencia para seguir siendo interpretables. La apertura se convierte así en un problema de custodia además de divulgación. También aparece un límite: algunos materiales científicamente importantes no pueden hacerse públicos responsablemente sin restricciones.

6. La apertura tiene límites legítimos

Los datos de participantes humanos pueden contener información que no fue consentida para publicación. Localizaciones ecológicas pueden exponer especies amenazadas. Métodos sensibles para la seguridad y obligaciones contractuales pueden restringir la difusión. Por eso ciencia abierta no puede significar publicación pública universal de todo objeto subyacente.

El principio más fuerte es inspeccionabilidad bajo gobernanza apropiada. Desidentificación, entornos seguros, acuerdos de uso, datos sintéticos, repositorios controlados y comités independientes pueden conservar parte de la auditabilidad sin fingir que el riesgo es cero. El problema de diseño consiste en hacer las restricciones explícitas y proporcionales para que «no puede ser público» no se convierta silenciosamente en «no puede comprobarse».

Afirmación C08. Las obligaciones legítimas de privacidad y seguridad pueden justificar un acceso controlado en vez de irrestricto. [1]

En qué punto estamos

La ciencia abierta es compatible con acceso controlado cuando la difusión irrestricta vulneraría obligaciones legítimas. El requisito epistémico no es publicidad máxima, sino una vía documentada para que el escrutinio cualificado siga siendo posible cuando sea viable. Eso convierte la gobernanza en parte de la infraestructura evidencial. La cuestión siguiente es la corrección: cuando se detecta un error, ¿queda registrado que el objeto científico cambió?

7. Versiones y correcciones forman parte de la evidencia

Un conjunto de datos corregido, un script revisado o una conclusión retractada no son el mismo objeto científico que la versión anterior al cambio. Identificadores persistentes, repositorios versionados y avisos de corrección permiten recuperar esa distinción. Sin ellos, usuarios posteriores pueden mezclar sin saberlo afirmaciones de una versión con datos o código de otra.

Una buena corrección hace más que sustituir un archivo. Conserva suficiente historia para mostrar qué cambió, por qué cambió y qué conclusiones quedan afectadas. Es una forma de memoria institucional. Permite que la ciencia cambie de opinión sin borrar el camino del cambio. KNO-024 volverá al problema de la conservación desde la dirección opuesta: qué desaparece cuando esos registros no sobreviven.

Afirmación C09 - Síntesis declarada. El versionado y los registros de correcciones distinguen objetos científicos modificados de sus versiones anteriores. **Síntesis declarada:** [1] respalda hacer explícitos el contexto de versión y las correcciones; tratar esos registros como lo que distingue un objeto científico modificado de su manifestación anterior es la síntesis.

En qué punto estamos

El versionado hace visible la corrección como cambio del objeto evidencial y no como sustitución silenciosa. Importa porque investigadores posteriores deben saber a qué datos, código e interpretación se refiere una afirmación. Un historial transparente de correcciones es evidencia sobre la historia de la evidencia. La pregunta final es cómo juzgar estas prácticas sin convertir la apertura en una insignia de corrección.

8. La ciencia abierta es infraestructura de corrección de errores

Las prácticas abiertas no garantizan que un conjunto de datos carezca de sesgo, que un modelo esté bien especificado o que una conclusión sea verdadera. Cambian qué errores pueden encontrarse y quién puede encontrarlos. Munafò y colegas presentan la práctica reproducible como un sistema de reformas que interactúan, no como una cura única.[4] Las National Academies tratan igualmente reproducibilidad y replicabilidad como propiedades que deben investigarse, no como veredictos simples sobre el valor científico.[1]

El argumento más fuerte a favor de la apertura es por tanto procedimental. Una afirmación acompañada de materiales inspeccionables expone más puntos posibles de crítica. Un resultado que no replica puede diagnosticarse con mayor precisión. Una corrección puede propagarse a los objetos que dependen de ella. La ganancia no es certeza, sino un camino más corto desde el desacuerdo al diagnóstico y del diagnóstico a la reparación.

Afirmación C11 - Síntesis declarada. Las prácticas de ciencia abierta pueden acortar el camino desde el desacuerdo hasta el diagnóstico al exponer una parte mayor del flujo evidencial. **Síntesis declarada:** [1,4] respaldan abrir registros y flujos de trabajo a inspección y reutilización; la afirmación de que eso puede acortar el camino del desacuerdo al diagnóstico es la síntesis.

Afirmación C12 - Síntesis declarada. La ciencia abierta se entiende mejor como infraestructura de corrección de errores que como certificado de verdad. **Síntesis declarada:** [1,2,4] respaldan transparencia, prácticas de reproducibilidad y mecanismos de corrección; describir la ciencia abierta en conjunto como infraestructura de corrección de errores y no como certificado de verdad es la síntesis.

En qué punto estamos

La ciencia abierta se entiende mejor como infraestructura para la crítica organizada. Conserva más del recorrido entre observación y afirmación, distingue algunas decisiones por cuándo se tomaron y facilita rastrear correcciones. Estas propiedades pueden mejorar la fiabilidad sin demostrar verdad. La frontera queda clara: KNO-020 posee revisión y consenso; KNO-021 posee inspeccionabilidad, reproducibilidad y prácticas de corrección.

Lo que sabemos y cómo lo sabemos

- Reproducibilidad y replicabilidad prueban tipos distintos de consistencia y no deben tratarse como sinónimos.[1,2]
- El prerregistro puede hacer visible el momento en que se formularon hipótesis y decisiones analíticas.[3]
- Datos, código y procedencia abiertos pueden volver inspeccionable una parte mayor del análisis, pero el acceso sin contexto es insuficiente.[1,4]
- Versionado y correcciones visibles conservan a qué objeto científico se refiere una afirmación.

Lo que todavía no sabemos

- Cuánto mejora cada intervención de ciencia abierta la fiabilidad en disciplinas y diseños distintos.
- Qué combinaciones de apertura y acceso controlado preservan mejor auditabilidad y privacidad legítima.
- Cuánto tiempo siguen siendo ejecutables los objetos de investigación actuales cuando cambian software y hardware.
- Cómo alteran los incentivos el uso real, y no solo la adopción nominal, de prácticas abiertas.

Qué podría hacernos cambiar de opinión

- Evaluaciones amplias entre disciplinas que muestren que una práctica promovida no añade auditabilidad ni fiabilidad en uso real.
- Evidencia de que un control destinado a reducir comunicación selectiva crea distorsiones mayores de las que evita.
- Nuevos métodos de preservación de privacidad que vuelvan rutinarias y seguras formas hoy restringidas de escrutinio.

- Estudios de conservación a largo plazo que muestren que repositorios y versionado actuales no mantienen interpretables los objetos de investigación.

Si solo recuerdas una cosa

La ciencia abierta no certifica la verdad; conserva más del camino evidencial para que otros puedan inspeccionarlo, reproducirlo, discutirlo, corregirlo y reutilizarlo.

Artículos relacionados

- KNO-019 · Datos, tablas y registros
- KNO-020 · Revisión, consenso y desacuerdo
- KNO-024 · Conocimiento perdido y sesgo de conservación

Referencias

[1] National Academies of Sciences, Engineering, and Medicine. Reproducibility and Replicability in Science. National Academies Press, Washington, DC (2019). DOI 10.17226/25303. <https://doi.org/10.17226/25303> [2] Steven N. Goodman; Daniele Fanelli; John P. A. Ioannidis. What does research reproducibility mean?. Science Translational Medicine 8(341), 341ps12 (2016). DOI 10.1126/scitranslmed.aaf5027. <https://doi.org/10.1126/scitranslmed.aaf5027> [3] Brian A. Nosek; Charles R. Ebersole; Alexander C. DeHaven; David T. Mellor. The preregistration revolution. Proceedings of the National Academy of Sciences 115(11), 2600–2606 (2018). DOI 10.1073/pnas.1708274114. <https://doi.org/10.1073/pnas.1708274114> [4] Marcus R. Munafò; Brian A. Nosek; Dorothy V. M. Bishop; Katherine S. Button; Christopher D. Chambers; Nathalie Percie du Sert; Uri Simonsohn; Eric-Jan Wagenmakers; Jennifer J. Ware; John P. A. Ioannidis. A manifesto for reproducible science. Nature Human Behaviour 1, 0021 (2017). DOI 10.1038/s41562-016-0021. <https://doi.org/10.1038/s41562-016-0021>