

KNO-024 · Conocimiento perdido y sesgo de preservación

KNO · Formas de conocer

Juan Seguro

v0.3 · ES

Todos los derechos reservados

Resumen del artículo

Lo que sobrevive del pasado no es una muestra aleatoria de lo que existió. Los huesos se degradan, los manuscritos arden, las instituciones descartan registros, los formatos dejan de ser legibles y los conjuntos de datos desaparecen cuando fallan dispositivos o direcciones de contacto. El archivo superviviente queda así moldeado por procesos de conservación, selección y destrucción antes de que historiador o científico alguno lo interprete. La arqueología formalizó este problema mediante procesos de formación que separan la actividad pasada del registro material recuperado después. [1] La paleontología estudia igualmente cómo la conservación filtra organismos y ambientes.[2,3] Los archivos añaden decisiones humanas sobre qué se crea, guarda, describe y hace accesible.[4] La información digital elimina algunas restricciones físicas, pero introduce dependencia de formatos, software, migración y custodia activa.[5,6] Este artículo pregunta cómo deben cambiar las inferencias cuando la ausencia puede reflejar no conservación y no inexistencia, y cómo los sistemas de preservación pueden hacer visibles sus propios efectos selectivos.

En 60 segundos

Los registros sobreviven a través de filtros. En arqueología, los objetos pasan del uso activo al contexto arqueológico mediante procesos culturales y naturales de formación, de modo que lo excavado ya está transformado.[1] En paleontología, meteorización, transporte, enterramiento y mineralización hacen que unos organismos y ambientes tengan mucha más probabilidad que otros de entrar en el registro fósil.[2,3] Los archivos añaden otra capa: personas e instituciones deciden qué documentar, conservar, clasificar y permitir consultar.[4] Los medios digitales no eliminan la pérdida. Vines y colegas hallaron que la disponibilidad de conjuntos de datos disminuía fuertemente con la antigüedad del artículo en su muestra, ilustrando la fragilidad de datos dependientes de custodios individuales.[5] La preservación digital exige migración, comprobación de integridad, metadatos y continuidad institucional, no mera copia.[6] La lección metodológica es simple pero exigente: la ausencia en un registro superviviente solo es evidencia después de considerar los procesos que pudieron eliminar, ocultar o nunca registrar el objeto.

Por qué importa esta pregunta

Razonamos constantemente a partir de supervivientes: fósiles, cuadernos de laboratorio, expedientes administrativos, yacimientos, fotografías, conjuntos de datos y código fuente. Como esos supervivientes son visibles, pueden parecer el pasado mismo. No lo son. Todo registro tiene una historia de supervivencia, y esa historia puede estar correlacionada con la propiedad que queremos inferir. Los huesos grandes se conservan de forma distinta a tejidos frágiles; los archivos oficiales, de forma distinta al habla informal; los repositorios mantenidos, de forma distinta a discos olvidados. El sesgo de conservación importa cuando el mecanismo de supervivencia puede distorsionar frecuencia, diversidad o cronología aparentes.

1. El registro es producido por la supervivencia

La pregunta Si la mayoría de las huellas desaparece antes de poder inspeccionarse, ¿qué muestrea realmente el registro superviviente? La respuesta depende no solo de lo que ocurrió en el pasado, sino también de qué huellas se crearon, conservaron, recopilaron, describieron y quedaron disponibles para observadores posteriores.

Un acontecimiento pasado y una huella superviviente no son el mismo objeto. Entre ambos actúan procesos que crean, transforman, mueven, conservan y destruyen registros. La distinción de Schiffer entre contexto sistémico y arqueológico convirtió los procesos de formación en parte central de la inferencia: los objetos recuperados tienen historias posteriores al uso que afectan ubicación y asociación.[1] Antes de tratar una huella como muestra del pasado, hay que preguntar qué hizo probable su supervivencia.

Afirmación C01. La evidencia arqueológica está moldeada por procesos de formación entre el uso activo y su recuperación posterior. [1]

En qué punto estamos

El registro superviviente se genera dos veces: primero por los acontecimientos que crean huellas y después por los procesos que conservan unas y eliminan otras. La inferencia debe modelar ambos. La conservación no es una molestia añadida tras observar; forma parte del propio proceso de observación. La sección siguiente muestra cuánto importa esto en el registro físico de organismos y ambientes.

2. La tafonomía filtra el pasado biológico

Behrensmeyer calibró etapas de meteorización ósea con cadáveres de edad conocida y mostró que la meteorización depende del tiempo y de condiciones locales.[2] Kidwell y Flessa revisan variaciones en completitud, resolución temporal y fidelidad ecológica del registro fósil.[3] La abundancia fósil no equivale por tanto a abundancia pasada: partes duras, enterramiento rápido y ambiente deposicional alteran la probabilidad de conservación. La tafonomía trata ese filtro como evidencia y no como ruido.

Afirmación C02. La meteorización ósea varía con el tiempo y las condiciones ambientales y puede informar la interpretación tafonómica. [2]

Afirmación C03. El registro fósil difiere en completitud y en resolución temporal y ecológica según el contexto. [3]

En qué punto estamos

Los fósiles son observaciones filtradas. Meteorización, transporte, enterramiento y mineralización cambian qué organismos, partes corporales e intervalos temporales se vuelven visibles. La tafonomía no vuelve inútil el registro fósil; convierte el proceso de conservación en parte de la evidencia. El mismo principio se aplica a registros humanos, con una complicación adicional: personas e instituciones deciden intencionadamente qué documentar y conservar.

3. Los archivos se construyen, no solo se heredan

Schwartz y Cook sostienen que archivos y registros participan en la construcción de la memoria social en vez de funcionar como contenedores neutrales de hechos.[4] Decisiones sobre creación, valoración, descripción, acceso y custodia modelan lo que usuarios posteriores pueden encontrar. Un registro ausente puede significar que algo no ocurrió, pero también que no se creó, se destruyó, quedó fuera de adquisición, fue mal descrito o tuvo acceso restringido.

Afirmación C04. Los archivos y registros están moldeados por decisiones humanas sobre creación, valoración, descripción y acceso. [4]

En qué punto estamos

Los archivos humanos añaden selección intencional a la supervivencia física. Los registros pueden faltar porque nadie los creó, porque una institución no los conservó, porque la descripción los volvió difíciles de localizar o porque el acceso quedó restringido. La procedencia archivística incluye por tanto custodia y política. La siguiente cuestión es la asimetría: los sistemas oficiales suelen conservar unas voces y actividades más sistemáticamente que otras.

4. La ausencia suele estar estructurada

El sesgo de conservación es más peligroso cuando la ausencia se correlaciona con la pregunta estudiada. Piedra y metal duraderos sobreviven de forma distinta a textiles, madera o habla; las instituciones formales documentan de forma distinta a comunidades informales; las crisis pueden dejar más papeles que el éxito rutinario. Las preguntas clave son qué categorías podían crear huellas, cuáles tenían probabilidad de sobrevivir y cuáles de ser recogidas y descubiertas.

Afirmación C05 - Síntesis declarada. La ausencia puede estar estructurada cuando la probabilidad de supervivencia depende de la propiedad que se estudia. **Síntesis declarada:** [1-4] respaldan procesos selectivos de formación, supervivencia y archivo; describir la ausencia como estructurada cuando la supervivencia depende de la propiedad estudiada es la síntesis.

Afirmación C06 - Síntesis declarada. La ausencia en un archivo no es automáticamente evidencia de que un acontecimiento no ocurrió. **Síntesis declarada:** [1,4] respaldan varias vías por las que un registro puede no formarse o no sobrevivir; inferir que la ausencia archivística no equivale automáticamente a ausencia del acontecimiento es la síntesis.

En qué punto estamos

El sesgo de conservación es más fuerte cuando el mecanismo de pérdida está relacionado con el fenómeno estudiado. Entonces el registro superviviente puede exagerar materiales duraderos, oficiales, excepcionales o valorados institucionalmente. Interpretar correctamente exige un modelo de quién podía crear registros y cuáles podían sobrevivir. Los medios digitales parecen eliminar muchos filtros físicos, pero introducen dependencias y modos de fallo propios.

5. Digital no significa permanente

Vines y colegas solicitaron conjuntos de datos de 516 artículos biológicos con edades entre dos y veintidós años y hallaron que, entre los casos donde los autores informaron del estado de los datos, las probabilidades en favor de que un conjunto siguiera existiendo disminuyeron un 17 % por año en su muestra.[5] Direcciones de contacto rotas y almacenamiento obsoleto fueron obstáculos importantes. La estimación es específica de la muestra, pero demuestra que los datos digitales pueden desaparecer rápidamente sin preservación institucional.

Afirmación C07. En la muestra de Vines y colaboradores, la disponibilidad de datos disminuyó con fuerza al aumentar la antigüedad del artículo. [5]

Afirmación C08. La preservación digital requiere una custodia activa que va más allá de la simple copia. [5,6]

En qué punto estamos

La información digital puede copiarse con bajo coste, pero la persistencia sigue requiriendo custodia activa. Los datos pueden desaparecer por dispositivos obsoletos, enlaces rotos, credenciales perdidas, deriva de formatos o pérdida de contexto. El estudio de Vines muestra una forma empírica de ese deterioro en datos de investigación. La preservación pasa así de proteger un objeto físico a mantener una cadena interpretable de archivos, software, metadatos e instituciones.

6. Digitalizar no equivale a preservar digitalmente

Conway distingue digitalización para conservación de preservación digital como actividad sociotécnica.[6] Escanear puede reducir manipulación del original y ampliar acceso, pero el escaneo se convierte en un objeto digital que requiere gestión. Resolución, color, compresión, metadatos y formato determinan qué captura el sustituto, mientras migración y validación determinan si sigue siendo utilizable.

Afirmación C09. La digitalización y la preservación digital son actividades distintas. [6]

Afirmación C10 - Síntesis declarada. Un sustituto digital conserva solo lo que codifican sus decisiones de captura y metadatos. **Síntesis declarada:** [6] respalda distinguir decisiones de digitalización de preservación digital; la cláusula de que el sustituto conserva solo lo que codifican captura y metadatos es la síntesis.

En qué punto estamos

La digitalización crea un nuevo objeto de conservación en vez de terminar el trabajo de preservar. Lo capturado depende de decisiones de conversión, y los archivos resultantes requieren gestión continua. Un sustituto digital puede mejorar acceso y omitir rasgos del original o quedar obsoleto. El problema siguiente es que los sistemas de conservación deben hacer visibles sus propias selecciones y transformaciones para que usuarios futuros juzguen qué sobrevivió.

7. La conservación necesita procedencia de pérdidas y cambios

Un archivo útil puede registrar no solo lo que contiene sino, cuando es posible, qué fue transformado, migrado, retirado o se sabe que falta. Las sumas de comprobación documentan integridad de bytes; los registros de migración, cambios de formato; las políticas de colección, límites de selección. Estos registros no reconstruyen todo lo perdido, pero vuelven auditables algunos procesos de conservación y ayudan a distinguir evidencia original de transformaciones posteriores.

Afirmación C11 - Síntesis declarada. Los sistemas de preservación pueden mejorar la auditabilidad registrando migraciones, integridad e historial de selección. **Síntesis declarada:** [6] respalda la necesidad de gestionar objetos digitales a través de acciones de preservación; afirmar que registrar migraciones, integridad e historial de selección mejora la auditabilidad es la síntesis.

En qué punto estamos

Los sistemas de conservación pueden reducir sesgo no solo guardando más, sino documentando los filtros que aplican. Registros de integridad, historiales de migración y políticas de colección exponen partes del proceso de supervivencia al escrutinio posterior. No recuperan lo que nunca se registró, pero evitan que una transformación silenciosa se presente como evidencia original. La sección final pregunta cómo formular afirmaciones cuando el filtro de conservación solo se conoce parcialmente.

8. La ausencia requiere un modelo de conservación

Una ausencia es evidencia más fuerte cuando el sistema habría tenido alta probabilidad de crear, conservar y exponer el registro pertinente si el acontecimiento hubiera ocurrido. Es más débil cuando alguna de esas probabilidades es baja o desconocida. Arqueólogos modelan procesos de formación, paleontólogos estudian tafonomía, archiveros examinan valoración y custodia, y la preservación digital estudia formatos y custodia. Dominios distintos expresan la misma disciplina: interpretar el registro mediante el proceso que lo hizo observable.

Afirmación C12 - Síntesis declarada. La ausencia es evidencia más fuerte cuando la creación, preservación y posibilidad de descubrimiento habrían sido todas probables. **Síntesis declarada:** [1-6] respaldan filtros distintos de creación, supervivencia, archivo y preservación digital; combinarlos en este criterio sobre cuándo la ausencia constituye evidencia más fuerte es la síntesis.

En qué punto estamos

El conocimiento perdido no es solo contenido ausente; es incertidumbre sobre el proceso de filtrado que produjo lo que queda. La ausencia se vuelve informativa cuando creación, conservación y descubribilidad habrían sido probables. Cuando esas condiciones son débiles, la confianza debe bajar. KNO-024 cierra la secuencia Formas de conocer convirtiendo la conservación en parte de la evidencia: lo que sobrevive limita el conocimiento solo mediante un modelo de cómo sobrevivió.

Lo que sabemos y cómo lo sabemos

- Los registros arqueológicos están moldeados por procesos de formación entre uso y recuperación.[1]
- La meteorización ósea y otros procesos tafonómicos afectan qué huellas biológicas sobreviven y cómo se interpretan.[2,3]
- Los archivos incorporan decisiones de creación, valoración, descripción y acceso en vez de ser contenedores neutrales.[4]
- Los conjuntos de datos pueden dejar de estar disponibles con el tiempo cuando la conservación depende de custodios individuales.[5]

Lo que todavía no sabemos

- Cuánto material no documentado se ha perdido ya de ecosistemas contemporáneos de investigación digital.
- Cómo varían las probabilidades de conservación entre comunidades infrarrepresentadas y sistemas informales de registro.

- Qué estrategias de preservación digital seguirán siendo interpretables tras cambios tecnológicos de siglos.
- Cómo cuantificar mejor el sesgo de conservación cuando creación, pérdida y descubrimiento son inciertos.

Qué podría hacernos cambiar de opinión

- Registros empíricos que muestren que clases importantes de evidencia supuestamente frágil se conservan al mismo ritmo que evidencia duradera en contextos relevantes.
- Estudios archivísticos que demuestren que valoración y acceso no afectan de forma medible inferencias históricas posteriores en colecciones comparables.
- Estudios longitudinales digitales que muestren pérdida insignificante de datos sin custodia institucional.
- Nuevos métodos de conservación que vuelvan irrelevante a largo plazo la obsolescencia de formatos, software y metadatos.

Si solo recuerdas una cosa

Un registro superviviente es evidencia sobre el pasado y sobre los procesos que seleccionaron qué podía sobrevivir; la ausencia solo se interpreta considerando ambos.

Artículos relacionados

- KNO-009 · Archivos, bibliotecas y memoria colectiva
- KNO-019 · Datos, tablas y registros
- KNO-021 · Ciencia abierta y reproducibilidad

Referencias

[1] Michael B. Schiffer. Archaeological Context and Systemic Context. *American Antiquity* 37(2), 156–165 (1972). DOI 10.2307/278203. <https://doi.org/10.2307/278203> [2] Anna K. Behrensmeyer. Taphonomic and ecologic information from bone weathering. *Paleobiology* 4(2), 150–162 (1978). DOI 10.1017/S0094837300005820. <https://doi.org/10.1017/S0094837300005820> [3] Susan M. Kidwell; Karl W. Flessa. The quality of the fossil record: populations, species, and communities. *Annual Review of Earth and Planetary Sciences* 24, 433–464 (1996). DOI 10.1146/annurev.earth.24.1.433. <https://doi.org/10.1146/annurev.earth.24.1.433> [4] Joan M. Schwartz; Terry Cook. Archives, Records, and Power: The Making of Modern Memory. *Archival Science* 2(1–2), 1–19 (2002). DOI 10.1007/BF02435628. <https://doi.org/10.1007/BF02435628> [5] Timothy H. Vines; Arianne Y.K. Albert; Rose L. Andrew; Florence Débarre; Dan G. Bock; Michelle T. Franklin; Kimberly J. Gilbert; Jean-Sébastien Moore; Sébastien Renaut; Diana J. Rennison. The availability of research data declines rapidly with article age. *Current Biology* 24(1), 94–97 (2014). DOI 10.1016/j.cub.2013.11.014. <https://doi.org/10.1016/j.cub.2013.11.014> [6] Paul Conway. Preservation in the Age of Google: Digitization, Digital Preservation, and Dilemmas. *The Library Quarterly* 80(1), 61–79 (2010). DOI 10.1086/648463. <https://doi.org/10.1086/648463>