

¿Falta de concienciación o desconocimiento? La transparencia y la reproducibilidad en la Arqueología Computacional

Lack of Awareness or Lack of Knowledge? Transparency and Reproducibility in Computational Archaeology

Salvador Pardo-Gordó*
Universidad de la Laguna
Departamento de Geografía e Historia
<https://orcid.org/0000-0002-1060-1526>
spardogo@ull.edu.es

Alfredo Cortell-Nicolau
University of Cambridge
McDonald Institute for Archaeological Research
<https://orcid.org/0000-0003-1982-6403>
ac2320@cam.ac.uk

Recibido: 26/07/2022; Revisado: 05/12/2022; Aceptado: 22/12/2022

Resumen

La aplicación de la computación en Humanidades en la última década se ha consolidado como una disciplina clave para el estudio de los procesos históricos desde numerosas vertientes siendo necesaria la utilización de diferentes softwares y lenguajes para la elaboración de dichas aproximaciones y el posterior tratamiento de los resultados. No obstante, son muchos los trabajos que carecen de transparencia (acceso a los datos y código) por lo que se hace imposible su reproducción por otros equipos de investigación. Por ello, en este documento se realiza una revisión de los trabajos publicados en revistas peninsulares para evaluar el grado de transparencia de la comunidad computacional hispano-lusa.

Palabras clave: Ética, Transparencia, Reproducibilidad, Código Abierto, Simulación.

Abstract

The last decade has seen the consolidation of Computational Archaeology and the Digital Humanities

*Autor de correspondencia / *Corresponding author.*

as key disciplines in the multifaceted study of historical processes. In this regard, the use of different software packages or programming languages for data analysis has notably increased among archaeologists. However, in many cases open research protocols are not fully observed (when access to data and code are not always granted), which makes reproducibility impossible for other research groups. The present article reviews the transparency of works published in the most important journals of the Iberian Peninsula, to evaluate the degree of openness within the Iberian research community.

Keywords: Ethics, Transparency, Reproducibility, Open Source, Simulation.

1. INTRODUCCIÓN

La Arqueología Cuantitativa y/o Arqueología Computacional hace referencia a la aplicación de técnicas y métodos cuantitativos, incluyendo modelizaciones *in silico* para responder a cuestiones histórico-arqueológicas. A pesar de ser una disciplina tradicionalmente ligada a las escuelas británica y americana dónde los primeros trabajos de simulación se remontan a la década de 1970 (DORAN, 1970; FLANNERY, 1973; WOBST, 1974; AMMERMAN y CAVALLI-SFORZA, 1971, GARDIN, 1988), en la última década ésta se ha consolidado en la investigación hispano-lusa no solo por la consolidación de los escasos grupos focalizados en la computación-cuantificación sino también por la emergencia de nuevas investigaciones que bajo el «paraguas» de la Arqueología Computacional están llevando a cabo trabajos novedosos (CORTELL-NICOLAU *et al.*, 2021; ORENGO *et al.*, 2021; PARDO-GORDÓ *et al.*, 2022; RUBIO-CAMPILLO *et al.*, 2022). Todo ello nos sugiere que las nuevas tecnologías, y especialmente la Arqueología Computacional y Cuantitativa permiten nuevas perspectivas en la forma de visualizar y analizar el registro arqueológico. Es un hecho que las Humanidades, y en especial la Arqueología, se han beneficiado de todo tipo de aproximaciones desde distintos campos como las Ciencias Sociales, Ciencias Naturales y Ciencias de la Computación entre muchas otras.

La adopción de métodos computacionales por parte de la comunidad arqueológica implica la existencia de tres grandes retos intrínsecos asociados a las aplicaciones de la simulación para explorar procesos histórico-sociales, y que acertadamente se han explorado en el *Manifesto of Computational Social Science* (CONTE *et al.*, 2012):

a) La complejidad de la disciplina donde las relaciones entre individuo-sociedad-entorno implican la retroalimentación entre las diferentes escalas de análisis, y, por extensión, la aparición de fenómenos/propiedades emergentes las cuales no siempre se comprenden.

b) La utilización de métodos en los que se realice una integración del binomio computación-humanidades donde el modelo se convierta en el laboratorio virtual para generar explicaciones alternativas.

c) Las variables utilizadas a escala de modelado. Este reto implica la integración dentro del modelo de diferentes características, procesos, escalas de análisis etc. Siendo este el desafío más inmediato de la simulación social en Humanidades. En este sentido, la simulación computacional en Humanidades debe situarse entre el principio de la *Navaja de Ockham* y la *Ley contra la tacañería*

de K. Menger. Dicho de otra manera, las Humanidades deben virar hacia la construcción de modelos computacionales donde las variables seleccionadas sean las estrictamente necesarias para, sin caer en el reduccionismo, permitir explorar los resultados *in silico* en relación con el fenómeno histórico analizado. En definitiva, «un modelo debe ser tan sencillo como sea posible, pero no más» (atribuida a Estein en CALAPRICE, 2011: 385).

Paralelamente a los retos intrínsecos de la simulación existen tres desafíos de la ciencia contemporánea. En el caso específico de la Arqueología, el viraje hacia una *Arqueología en abierto* (MARWICK, 2020; MARWICK *et al.*, 2017) permitirá la consolidación y, sobre todo, la adopción de las técnicas computacionales y cuantitativas.

Los principales retos se focalizan en la generación de conciencia, transparencia y reproducción:

a) **Educación**. Desde la disciplina deben realizarse numerosos trabajos pedagógicos o de concienciación en relación con la necesidad de una investigación transdisciplinar donde se presente la simulación y el análisis cuantitativo de los datos como elementos clave para realizar aproximaciones alternativas a los procesos histórico-arqueológicos. Este aspecto implica convencer a la comunidad científica que considera la simulación social como un escollo más que una solución. En definitiva, tanto desde los centros docentes como los de investigación debe realizarse un trabajo de *alfabetización digital* (RAMÍREZ-SÁNCHEZ, 2018) que permita una adopción natural de las Humanidades Digitales.

b) **La transparencia y reproducibilidad**: Estas cuestiones se están discutiendo en aquellas disciplinas que hacen uso de las técnicas computacionales como método para explorar diferentes problemáticas. Entre ellas, destacamos la Psicología (DERKSEN, 2020; GERNSBACHER, 2018), la Inteligencia Artificial (HAIBE-KAINS *et al.*, 2020) o, recientemente, la Arqueología (MARWICK, 2017). La transparencia y la reproducibilidad son complementarias. La primera hace referencia a la necesidad de hacer públicos todos los datos utilizados en los diferentes análisis y el material gráfico asociado al trabajo. En este sentido, la comunidad científica trabaja para adoptar una serie de principios comunes a través de las buenas prácticas para la gestión y administración de datos científicos, conocidos como los principios FAIR¹ (WILKINSON *et al.*, 2016), que promueven un marco uniforme y común para la publicación y reutilización de datos; la segunda, reproducibilidad, remarca la obligación de hacer público el código de programación (entendiendo como tal el lenguaje de programación utilizado, cualquiera que este sea) y modelo computacional para que cualquier revisor o revisora pueda reproducir el proceso metodológico sin necesidad de realizar ninguna tarea más allá del cambio de la ruta de acceso a los datos y ejecutar el proceso.

Dichos retos, educación y transparencia-reproducibilidad permitirán desterrar la idea instalada en el imaginario colectivo de que los modelos computacionales están diseñados para que los resultados se ajusten a la hipótesis

¹ Siglas en inglés de los conceptos *Findable* (disponibles), *Accesible* (accesibles), *Interoperable* (interoperables) y *Reusable* (reutilizables). En este trabajo se utilizará el acrónimo en inglés ya que es el utilizado toda la comunidad científica. Véase el apartado 4.1 para la descripción de cada uno de los conceptos.

de investigación. En definitiva, desasociar la concepción de que una investigación basada en modelizaciones no puede llegar a conclusiones válidas sobre cualquier proceso.

Con todo ello, el objetivo de este trabajo es examinar hasta qué punto se está llevando a cabo una ciencia reproducible, en los términos mencionados más arriba, en la Arqueología Computacional ibérica. Para ello, se han seleccionado aquellos trabajos publicados en revistas de temática arqueológica gestionadas desde instituciones hispano-lusas para su posterior categorización siguiendo los criterios de transparencia y reproducibilidad. En el siguiente apartado se explican los criterios de selección de dichos trabajos y los métodos de análisis utilizados. En la sección tres se analizan los resultados y se proponen algunos de los repositorios y métodos más utilizados para compartir datos y código. Finalmente se hará una reflexión sobre dónde se sitúan las aportaciones científicas en el contexto de los principales retos del siglo XXI.

2. CRITERIOS DE SELECCIÓN Y ENFOQUE ANALÍTICO

Se ha decidido limitar el alcance de este trabajo al ámbito editorial hispano-luso, lo cual excluye las publicaciones realizadas por investigadoras peninsulares en revistas con sede fuera de la Península Ibérica, a la vez que incluye los trabajos de investigadores extranjeros en las editoriales analizadas.

Uno de los primeros elementos para tener en cuenta es qué consideramos como Arqueología Computacional. En este sentido, es importante establecer una primera distinción entre los términos computacional y digital. Mientras el primero hace referencia a la modelización computacional, entendiendo como tal la aplicación de simulación teórica, apoyada en el poder de cálculo de los ordenadores actuales (BARTON, 2013), el segundo se focaliza en procesos de digitalización generalmente a partir de modelado visual, como software de diseño computacional asistido (CAD), tridimensional y, en general, con gran aplicación en el ámbito patrimonial (WATRALL, 2016). Aunque el uso de ambos puede combinarse, y así lo atestiguan algunas técnicas de investigación actuales, como la morfometría geométrica bidimensional y tridimensional (MATZIG, *et al.*, 2021; OKUMURA y ARAUJO, 2019; VENEZIANO *et al.*, 2021), se suele recurrir al primero como herramienta de validación de hipótesis, y en él nos centraremos en el presente trabajo.

Atendiendo a que, como hemos visto, el uso de ordenadores no es suficiente para clasificar una técnica de análisis como computacional, no es sencillo seleccionar qué trabajos pueden considerarse como Arqueología Computacional y cuáles no. En este sentido, la utilización de Sistemas de Información Geográfica (SIG) supone uno de los grandes problemas. Si bien por un lado puede ser utilizados solo con fines ilustrativos o contextuales, por otro lado, el uso de modelización espacial en Arqueología para la construcción de modelos explicativos tiene hoy en día una tradición sólida, tanto en el ámbito peninsular como fuera de él (entre otros, BEVAN, 2015; ISERN *et al.*, 2014; SCHAUER *et al.*, 2020; SNITKER *et al.*, 2018).

Algo similar ocurre con la modelización matemática. Aunque parece fuera de duda que podemos considerar como computacional la modelización compleja, donde frecuentemente varios algoritmos y/o elementos de simulación pueden ser necesarios, no está tan claro qué elementos de estadística básica puedan considerarse como computacionales. En este sentido, creemos que deberían estimarse como computacionales aquellas metodologías donde sea necesario el cálculo intensivo. Más en particular, para la selección de trabajos hemos utilizado los siguientes criterios:

1. SIG: El trabajo realizado trasciende la realización de mapas y requiere, de una forma u otra, modelado estadístico espacial.

2. Modelización matemática: Se realizan análisis complejos, más allá de contrastes de hipótesis básicos, tanto paramétricos como no paramétricos.

3. Modelización *in silico*: Cualquier tipo de modelización realizada por ordenador puede considerarse como computacional.

Atendiendo a estos criterios, hemos revisado todos los trabajos publicados por editoriales hispano-lusas, indexadas en la categoría Arqueología (*arts and humanities*) que se encuentren dentro del primer cuartil (Scopus), desde el año 2015 hasta el año 2021, lo cual nos ha dejado con un total de 60 publicaciones. Ya que nuestro objetivo es establecer si dichos artículos son reproducibles o no, para todos ellos se ha registrado si el artículo incluye/comparte código y si incluye/comparte datos (Tabla 1). Somos conscientes de que, en muchos casos (sobre todo en la aplicación de técnicas SIG o en casos de sumatorios de probabilidad e inferencia bayesiana sobre dataciones radiocarbónicas), el código no es utilizado porque los mismos softwares ofrecen GUIs o interfaces que realizan la modelización solicitada sin necesidad de recurrir a la escritura de programas informáticos. A nuestra consideración, la utilización de interfaces, o GUIs, es una gran ayuda para la simplificación de procesos y tiene un gran sentido para la optimización del rendimiento y la producción en contextos industriales y/o empresariales. Sin embargo, permite la automatización de un proceso, como es la elección de las técnicas y parámetros más adecuados para el proceso de inferencia, que en ningún caso debería ser automático. Además, en el ámbito académico el objetivo no es la obtención de un resultado en el menor tiempo posible (o al menos no debería serlo), sino la generación de un modelo explicativo que pueda ser puesto a prueba y testado. En este sentido, la utilización de interfaces limita enormemente la reproducibilidad del proceso (mucho más si hablamos de software propietario), y es por ello por lo que consideramos los artículos sin inclusión de código como no completamente reproducibles, independientemente de la herramienta con la que se hayan llevado a cabo los análisis.

TABLA 1

Tabla resumen con la información analizada en este trabajo. En esta se indican las revistas consultadas como las variables analizadas

Revista	Nº Artículos	Acceso datos Total	Acceso datos Parcial	Acceso código
<i>Archeofauna</i>	5	0	1	0
<i>Arqueología de la arquitectura</i>	2	1	0	0
<i>Arqueología iberoamericana</i>	3	1	0	0
<i>SPAL</i>	1	1	0	0
<i>Trabajos de Prehistoria</i>	27	12	10	2
<i>Virtual Archaeology Review</i>	12	3	1	3
<i>Zephyrus</i>	10	2	3	0

Dada la naturaleza del presente trabajo, no ha sido necesario realizar análisis estadísticos complejos. En este sentido, el método de análisis se centra en una descripción gráfico-estadística del estado de la reproducibilidad en la Península Ibérica durante los últimos años (2015-2021), así como de las diferencias editoriales. Todo ello de acuerdo con los criterios ofrecidos más arriba. Dada la selección de artículos, se ha tenido en cuenta cuántos de ellos comparten datos, de forma parcial o total y cuántos de ellos comparten el código utilizado. En cuanto al hecho de compartir datos, hemos considerado como parciales aquellos artículos cuya información pueda extraerse o deducirse de tablas presentes en el artículo, pero no estén especificados claramente. De esta forma, debe tenerse en cuenta que, para considerar un artículo como completamente reproducible, código y datos deberían estar preparados para funcionar sin necesidad de alterarlos o modificarlos (obviamente, más allá de cambiar las rutas de los archivos, si esto es necesario). Tal y como se ha mencionado en el párrafo anterior, en el caso de la utilización de interfaces se considera código no compartido.

Puede consultarse el listado total de los artículos analizados en el material electrónico asociado a este trabajo. En ese mismo material suplementario se adjuntan también todas las referencias bibliográficas en formato *.Bib*. Por último, todo el código y los datos asociados a este trabajo se encuentran disponibles en el repositorio Zenodo (<https://doi.org/10.5281/zenodo.6861216>).

3. RESULTADOS

En primer lugar, si nos centramos en una observación general, los datos en cuanto a la reproducibilidad de la ciencia arqueológica en el contexto analizado no son demasiado halagüeños. De un total de 60 artículos para el periodo estudiado, un 66,67% de los mismos contiene algún elemento de reproducibilidad (Fig. 1); esto es, que contienen, al menos, código o una parte de los datos utilizados. Sin embargo, si nos centramos en los artículos que podrían considerarse completamente reproducibles, es decir que contienen código y todos los datos utilizados, la cifra

desciende hasta un pobre 5%, incluso sin entrar a valorar si la forma en que estos datos/código han sido presentados garantiza su reproducibilidad.

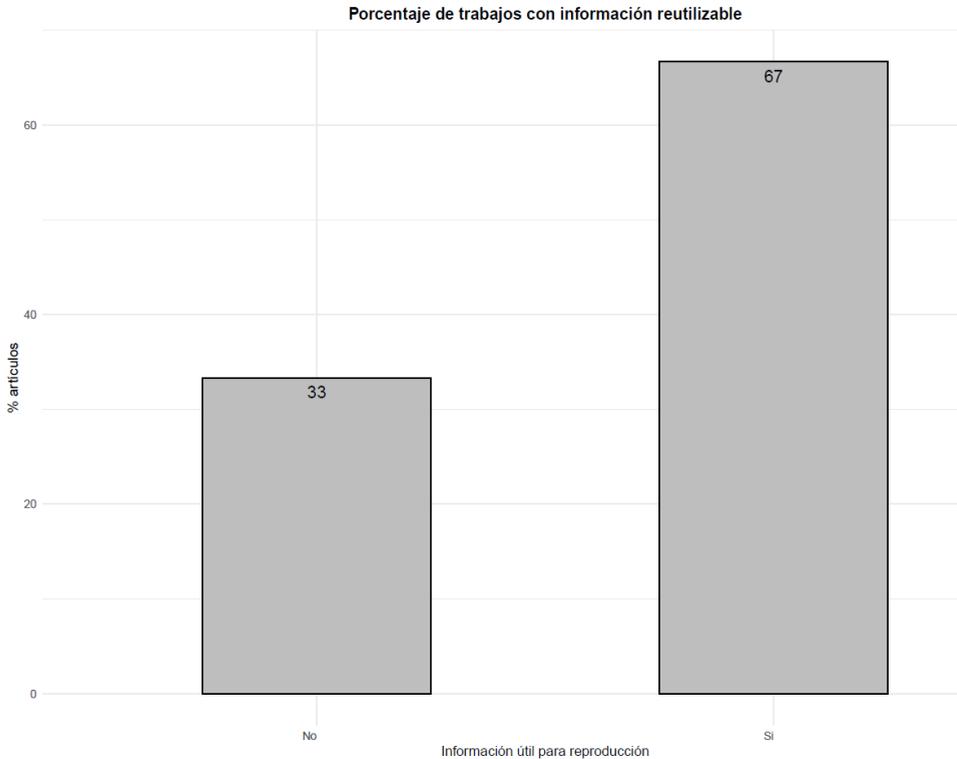


Figura 1. Número de artículos, expresado en porcentaje, que publican información útil para replicar los resultados.

En este sentido, considerando el ámbito internacional, existen diversas publicaciones que incluyen en su política editorial la sugerencia, e incluso el requerimiento de la presentación de los datos y código utilizados en el trabajo publicado. Mientras, obviamente, esto facilita la reproducibilidad científica, para el caso de las revistas seleccionadas (Fig. 2), tan solo dos de ellas, *Arqueología de la Arquitectura* y *Trabajos de Prehistoria* incluyen en su política editorial y/o en sus directrices la recomendación de la inclusión de los datos utilizados, así como ofrecen plataformas o soluciones para compartir estos mismos datos. No hemos podido encontrar recomendaciones similares para ninguna de las otras revistas consideradas. Es difícil valorar si, siendo únicamente una recomendación, esta política editorial tiene impacto real sobre la decisión de la autoría de compartir, o no, su información. En todo caso, sí que puede apreciarse que, para las revistas que incluyen estas recomendaciones, del total de artículos enviados, un 79,31%

incluye datos total o parcialmente, mientras que la cifra desciende a un 38,71% para las revistas que no incluyen en sus normas referencias o directrices en cuanto a su política de datos (Fig. 3). Por supuesto, estos resultados deben tomarse únicamente como orientativos, ya que el motivo real por el cual se comparte más información en unas u otras editoriales seguramente esconde un conjunto de razones mucho más complejo y variado que, en todo caso, excede los objetivos de este artículo.

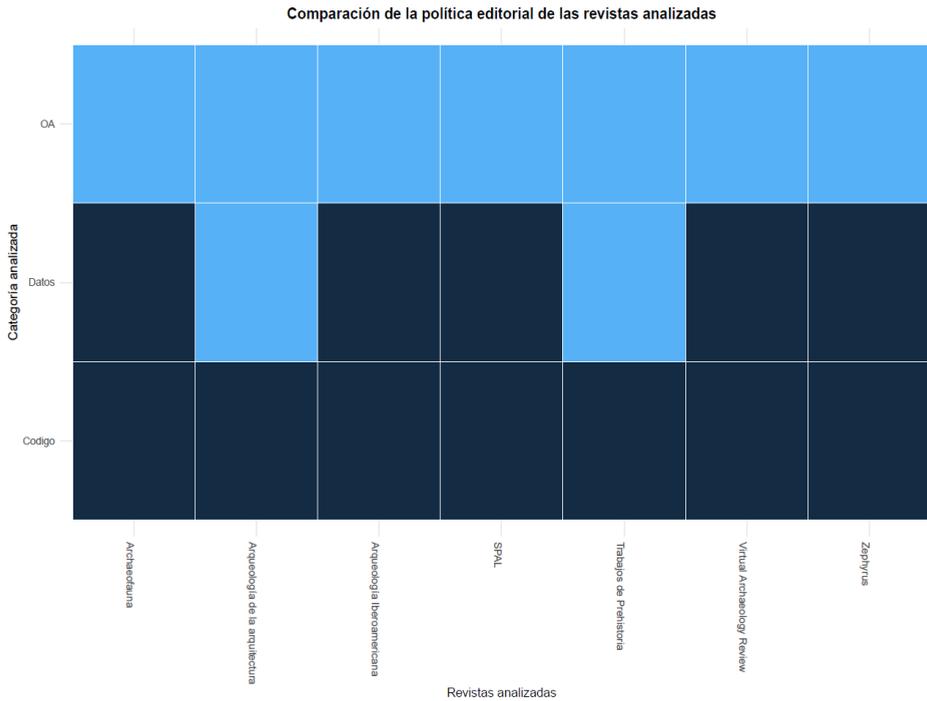


Figura 2. Análisis de la política editorial de las revistas exploradas. OA (publicación en abierto), datos (sugiere la publicación de los datos), código (sugerencia de publicación de código). El color azul claro implica que es un requisito mientras que el azul marino significa que no hay referencia en la política editorial a dicha cuestión.



Figura 3. Número de trabajos que comparten los datos/códigos según la política editorial de la revista.

Si observamos la tendencia global de publicaciones relacionadas con la Arqueología Computacional en la Península Ibérica, de acuerdo con el criterio establecido en el punto dos del presente trabajo, observamos cómo en general se observa una tendencia ascendente de dichas publicaciones (Fig. 4), si bien ciertamente irregular. En este sentido, hay varios elementos de interpretación en cuanto a esta tendencia. Por un lado, debe tenerse en cuenta que, al considerar únicamente revistas en el primer cuartil, esto puede estar reflejando en parte diferencias en el posicionamiento de las revistas peninsulares en cuanto a índices de calidad. Por otro lado, el ámbito computacional suele relacionarse con el ámbito teórico anglosajón, y es posible que la mayoría de estos trabajos realizados por la comunidad científica hispano-lusa estén enviándose a revistas internacionales, dejando para el contexto peninsular otros elementos arqueológicos más centrados en cuestiones de registro y debates internos. Finalmente, una tercera posibilidad reside en que la Arqueología Computacional es una disciplina incipiente, y si por un lado no siempre la aproximación metodológica planteada se ajusta a la línea editorial, por el otro pueden existir reticencias sobre los resultados y su utilidad en relación con la comprensión de fenómenos histórico-arqueológicos.

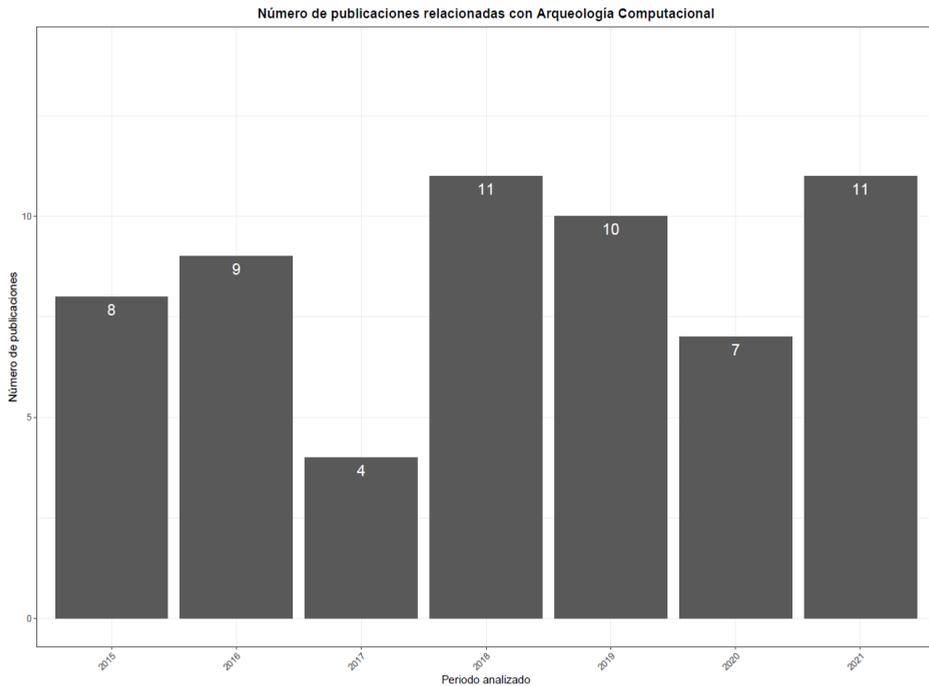


Figura 4. Evolución temporal del número de trabajos publicados asociados a la Arqueología Computacional. Los valores representados son en números absolutos y hacen referencia al sexenio analizado.

En todo caso, y este sí es un elemento positivo para tener en cuenta, si observamos la tendencia anual de datos y código publicados en cada revista (Fig. 5), observamos que la proporción de datos y/o código publicado con respecto al total de trabajos publicados anualmente crece consistentemente. Esto es especialmente acusado para el año 2021, cuya publicación de código/datos prácticamente dobla a cualquier año anterior, a pesar de que el número total de publicaciones para este año es similar al de los años 2016, 2018 o 2019. A este respecto, debería mencionarse que la concienciación de la comunidad científica en cuanto a la importancia de compartir datos o, al menos, su menor reticencia a compartirlos es un fenómeno relativamente reciente, donde muchas formas de colaboración, así como políticas que garanticen dicha cooperación, o plataformas seguras que preserven la atribución de autoría se están diseñando e implementando en este mismo momento.

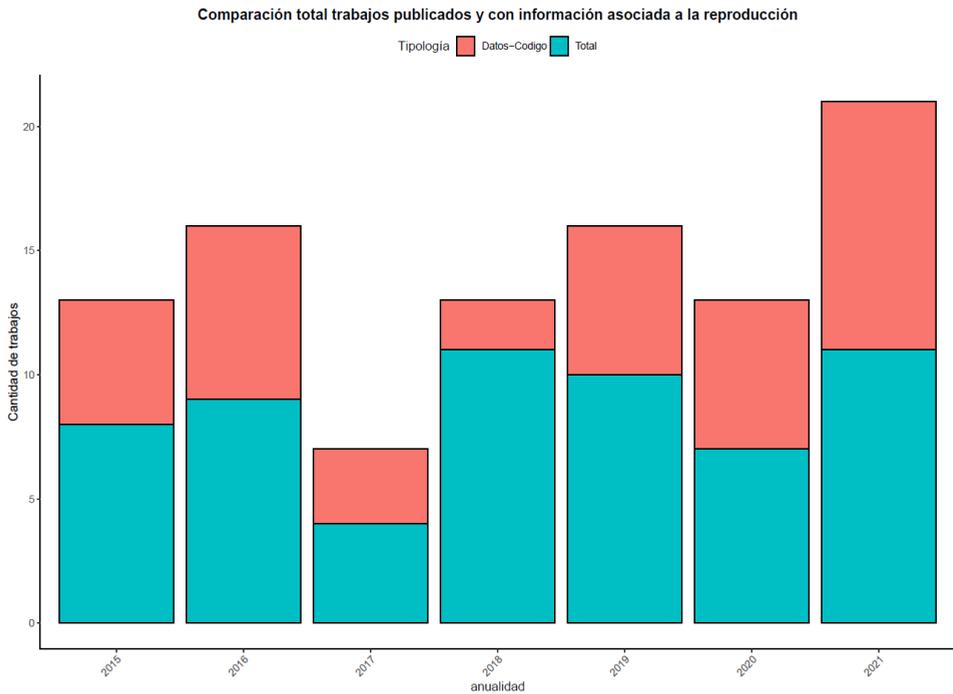


Figura 5. Número total de trabajos publicados por año en relación con el total de trabajos que publican información asociada a la transparencia/reproducibilidad.

Por último, debería mencionarse el aumento de código compartido producido también durante el último año 2021. De un total de 11 artículos, tres de ellos comparten código, representando casi un 30% del total, mientras que, para el resto de los años, tan solo en 2017 (25%) y en 2019 (10%) se compartió el código utilizado con un artículo por año respectivamente. Para el resto de los años estudiados, ningún artículo compartió el código utilizado.

4. DISCUSIÓN

4.1 Compartiendo datos: pros y contras

Según los principios FAIR, establecidos por WILKINSON *et al.* (2016), dentro de la comunidad científica los datos deben ser localizables, accesibles, interoperables y reutilizables. El debate sobre compartir datos y código no es exclusivo a la Arqueología, y, de hecho, muchas otras disciplinas científicas se encuentran, como la Psicología, la Ciencia de Datos o las Ciencias de la Salud, en procesos similares. Si bien cada una ha ido avanzando a su propio ritmo y el nivel de adopción de políticas y métodos de ciencia abierta es dispar (LINDSAY, 2017, WALLIS *et al.*, 2013,

ZHANG *et al.*, 2018). En términos prácticos, seguir estos principios FAIR significaría que, para que unos datos siguieran dichos principios, deberían tenerse en cuenta las siguientes premisas:

1. Disponibilidad: Los datos deben incluir un identificador único y global (comúnmente, DOI), deben describirse con suficientes metadatos, los cuales deben hacer alusión a dicho identificador y deben estar registrados e indexados en motores de búsqueda.

2. Accesibilidad: Los datos pueden conseguirse por medio de su identificador con protocolos estándar, abiertos, gratis y universales, donde pueda requerirse una autenticación de la persona usuaria y, en todo caso, los metadatos deben ser accesibles a perpetuidad.

3. Interoperabilidad: Debe tenerse en cuenta un formato en metadatos sin problemas derivados de alfabeto o lenguaje y con referencias a cualquier otro metadato necesario.

4. Reutilización: El lenguaje utilizado debe ser accesible, deben cumplirse las convenciones de la disciplina científica en que se enmarque el trabajo y las licencias de reutilización deben estar claramente especificadas.

Aunque es conveniente recordar y difundir estos principios para aumentar su uso, estos principios no han comenzado a tener una difusión y utilización real hasta momentos muy recientes. De hecho, incluso hoy en día siguen sin ser conocidos y/o utilizados por toda la comunidad científica. Es por ello de que en el presente trabajo nos hemos centrado en las publicaciones que comparten los datos y/o código utilizado, sin considerar hasta qué punto se siguen las normas FAIR.

En todo caso, desde los últimos años existe cada vez una mayor conciencia de la necesidad de reproducibilidad científica, tanto desde el punto de vista de la autoría, como de los organismos gubernamentales que financian la mayoría de los proyectos de investigación en Arqueología e Historia. En este sentido, y centrándonos en los marcos que puedan afectar a la investigación hispano-lusa, cualquier proyecto financiado por la Comisión Europea exige, aparte de distintos elementos de consideración ética, la elaboración de un plan de gestión de datos que sea plausible y acorde con las dimensiones del proyecto. Por otro lado, en las convocatorias nacionales no siempre existe directriz específica alguna en relación con los protocolos de publicación de los datos. En este aspecto, tanto la Agencia Estatal de Investigación (AEI) como la Agencia Nacional de Evaluación de la Calidad Académica (ANECA) han firmado la declaración de San Francisco (DORA) y el Manifiesto de Leiden donde la ciencia colaborativa, abierta y de calidad es un pilar fundamental. No obstante, en las convocatorias competitivas (proyectos de generación de conocimiento y programas postdoctorales) no suelen presentarse requisitos específicos en este sentido, más allá de la difusión de los resultados. Cierto es que la convocatoria *incentivación de la consolidación investigadora* del año 2022 ha incorporado en la memoria la necesidad de indicar un plan de gestión de datos (si procede); así que la pregunta que nos podemos plantear es ¿procede disponer de un plan de gestión de datos en proyectos

financiados por la administración pública? La respuesta es obvia, sí.

Cierto es que las instituciones, Universidades y Centros de Investigación, disponen de repositorios específicos (los cuales han sido recientemente objeto de evaluación (AGUILLO, 2020) y que, generalmente, la mayoría de ellas exigen a la autoría compartir sus publicaciones en dichos repositorios, de modo que estas se encuentren disponibles a perpetuidad y contribuyan a posicionarlos en lugares preferentes en los rankings de investigación. Sin embargo, y referente al tema que nos ocupa en este trabajo, esta política se ve lastrada por tres factores fundamentales:

1. La exigencia de publicación es del artículo en sí, pero no siempre de los datos y código que lo acompañan, con lo cual la reproducibilidad del artículo no está garantizada.

2. En muchas ocasiones los derechos de los artículos publicados no pertenecen a las personas que han generado el conocimiento científico, sino a las editoriales, lo que provoca que las versiones publicadas en los repositorios sean versiones previas a la impresión final.

3. La utilización por parte de diferentes grupos de la cláusula «*data availability upon request*» permite a la autoría del trabajo indicar que los datos son accesibles si se requieren específicamente a la persona que lidera el artículo o es la investigadora principal del proyecto que ha generado el *output*. Sin embargo, recientemente se ha demostrado que el personal científico que utiliza esta cláusula es reticente a compartir los datos. En este sentido, según el análisis realizado por TEDERSOO *et al.* (2021) solo un 39.4% ha compartido los datos cuando se le han requerido, cuestión criticada en diversos ambientes académicos, tanto formales como informales (Fig. 6). Por consiguiente, desde las editoriales y, sobre todo, desde las entidades financiadoras no debería permitirse esta práctica.



Figura 6. Ejemplo de diferentes críticas sobre la cláusula «*data availability upon request*» publicados en la red social Twitter.

Si bien hasta ahora nos hemos centrado en los problemas derivados del mundo editorial, la autoría no es completamente inocente en cuanto a una reproducibilidad y transparencia completas. Hay muchos motivos por los cuales la autoría puede decidir no compartir sus datos, que incluyen, pero no se limitan a los siguientes: (1) reservas ante la falta de reconocimiento y/o atribución de autoría, (2) trabajo adicional necesario para que datos y código cumplan unas condiciones mínimas de reproducibilidad, (3) inseguridad con respecto a la eficiencia/elegancia del propio código y/o preocupación ante la crítica del mismo o (4) los datos pertenecen a una investigación en curso.

Consideramos que ninguno de los motivos anteriores es suficiente para la negativa a compartir datos y todos son solucionables.

1. Atribución de autoría: Si los datos cumplen los principios FAIR, se les asigna un identificador único y pueden ser citados como cualquier otra propiedad intelectual generada por el personal investigador. En este caso, la probabilidad de que exista una falta de atribución de autoría es exactamente la misma que existe en publicaciones científicas completas. Por lo tanto, este debate afectaría a la ética investigadora individual, pero no al hecho de que esta ética se aplique sobre datos/código o artículos completos.

2. Adaptación de datos: Es cierto que proveer los datos en un formato inteligible exige un trabajo adicional por parte del personal científico, pero ese trabajo adicional frecuentemente permite detectar errores no observados previamente, a la vez que fuerza a una mayor auditoría de los datos y, por lo tanto, a una mayor calidad científica.

3. Ineficiencia de código: Siguiendo la línea del punto anterior, es frecuente que el personal investigador acumule soluciones *ad hoc* que acaban convirtiendo en ininteligibles tanto código como datos. Sin embargo, una revisión profunda de los mismos, bajo la consciencia de que pueden ser examinados, ayuda a la autoría a una reflexión adicional que, en última instancia, mejora la calidad científica.

4. Los datos pertenecen a una investigación en curso: A nuestro parecer, este caso es más complejo. Si bien es cierto que, en el momento de publicación de un artículo debería considerarse que los datos que han generado la investigación han sido estudiados exhaustivamente y, por lo tanto, no tendría sentido que pertenecieran a una investigación 'en curso', también es cierto que, dada la naturaleza y, en muchas ocasiones, el volumen de las colecciones arqueológicas, es frecuente que se publiquen investigaciones parciales, en ocasiones ligadas a proyectos de Tesis Doctorales o incluidas en proyectos de investigación más amplios. En esta situación, es potestad de la autoría considerar si es momento, o no, de publicar los datos, pero, en cualquier caso, la provisionalidad de la investigación no debería ser utilizada como excusa *ad aeternum* para evitar la publicación y reutilización de datos y código.

La publicación de datos y código debe seguir estándares que permitan su reproducibilidad, siendo insuficiente el recurrente mantra de «datos disponibles bajo requerimiento», cuyas consecuencias no es necesario nombrar aquí y, en última instancia, debe tenerse en cuenta que dicha publicación es beneficiosa de

varias maneras, tanto para quien publica como para quien consulta los datos. Publicar los datos y código puede ayudar a (1) tener una mejor comprensión de sus propios datos y formato, (2) ganar visibilidad, de la misma forma que la puede ganar con cualquier publicación científica y (3) promover una situación en la que, de la misma manera que el/la autor/a ha publicado sus propios datos, en un futuro podrá utilizar datos y código utilizados por otras personas.

Por último, y más allá de las reservas personales, los beneficios para la comunidad científica son incuestionables, ya que permiten una mayor y mejor auditoría de la investigación, a la vez que incentivan la colaboración y permiten la fijación de procesos y herramientas metodológicas, sobre las que se puede seguir una optimización que, en última instancia, consolide herramientas y construcciones teóricas específicas de cada disciplina.

4.2 ¿Qué podemos hacer?

Existen varias líneas de acción que se pueden tomar, dentro de la comunidad científica, para promover la publicación de datos y código asociados a un artículo. La primera de todas ellas sería, obviamente, dando ejemplo y publicando datos y código, en formato reproducible, en nuestros propios trabajos. Se pueden tomar líneas activas de comunicación tanto en las diferentes reuniones científicas, nacionales e internacionales, como en otros tipos de plataformas, ya sea en redes sociales o dentro de marcos institucionales.

Un arma especialmente poderosa para promover e incentivar la publicación de datos reside en nuestro papel como revisoras y revisores de artículos científicos. En primer lugar, y dentro del papel de revisión, debe partirse de la base de que es imposible asegurarse de que un artículo cumple con los mínimos estándares científicos, y de que sus conclusiones son válidas, sin una revisión del código utilizado. Por lo tanto, una revisión de código debería ser exigencia ya que, sin acceso al mismo, no se puede asegurar que los cálculos sean correctos y que las conclusiones estén debidamente fundamentadas. Esta exigencia de revisión ya obliga a la autoría a presentar código y datos en un formato inteligible y, por lo tanto, el siguiente paso, su publicación, es una consecuencia lógica, que también debería ser recomendada, si no exigida, por el personal investigador que revisa el trabajo.

Existen varias formas en las que la autoría puede compartir el código y datos utilizados. Una de las más obvias y sencillas es compartirlos como material suplementario asociado a la propia publicación. Dicha forma de publicación tiene la ventaja de que otorga una identificación inequívoca a datos y código que, además, están ligados al propio texto. Sin embargo, no todas las revistas ofrecen esta posibilidad y, además, en caso de que la revista no sea de acceso libre, los datos y código tampoco lo serían y, por lo tanto, se violarían los principios FAIR de reutilización de datos.

Existen plataformas digitales especialmente enfocadas a la investigación colaborativa y que, por lo tanto, ofrecen espacio para la reutilización de datos y

código. Algunas de las más utilizadas serían GitHub² o Zenodo,³ y alguno de sus elementos más positivos es que exigen una estandarización a la hora de almacenar material, lo cual garantiza unas mínimas condiciones de reproducibilidad. Sin embargo, no todas funcionan exactamente de la misma manera. Por ejemplo, GitHub se enfoca al trabajo colaborativo y, si bien, puede contener datos y código, no genera identificadores individuales por sí misma. Zenodo, por su parte, se centra en justamente lo contrario, en ofrecer identificadores individuales para la publicación de datos, código, software u otros tipos de materiales. En cualquier caso, ambas pueden ser interconectadas. Finalmente, mención especial merece Digital.CSIC⁴ puesto que es un repositorio de datos en abierto gestionado por la primera institución científica del estado español, el Consejo Superior de Investigaciones Científicas (CSIC). En este no solo se puede consultar la producción científica de su personal sino también las publicaciones gestionadas por el propio Consejo.

Se han mostrado estas dos únicamente como ejemplo, pero existen toda una serie de plataformas donde se puede llevar a cabo una investigación colaborativa a diferentes niveles, como servicios de revisión por pares previos a la publicación como PCI Archaeology⁵ o publicaciones de 'preprints' (artículos en formato previo a la edición final) en distintos repositorios, como bioRxiv⁶ o SocArxiv,⁷ y donde algunas de estas plataformas permiten la publicación de datos y código. Nuestro objetivo no es, sin embargo, entrar a valorar cada una de ellas, pero sí mostrar que existen diferentes formas en las que cualquier investigador puede compartir las herramientas utilizadas en su trabajo y que, en última instancia, este tipo de prácticas favorecen el crecimiento científico tanto general como individual.

4.3. La situación en el contexto hispano-luso

Como hemos visto, aunque existe un crecimiento tanto de publicaciones dedicadas a la arqueología computacional, como de la publicación de datos y código asociados a las mismas en la Península Ibérica, la situación dista mucho de ser ideal. Sin embargo, esta no es excesivamente diferente de lo que ocurre a escala internacional, si nos guiamos por el compendio de publicaciones realizado por Marwick. De acuerdo con una recopilación automática hecha por este autor,⁸ la tendencia creciente de trabajos incluyendo código y datos en la escena internacional no dista demasiado de la escena analizada en este trabajo, donde 2021, en ambos casos, dobla prácticamente a cualquier año anterior.

A esto debe añadirse la particular estructura en la investigación arqueológica en cada país. En el caso de la Arqueología hispano-lusa, existen pocos centros

2 <https://github.com>

3 <https://zenodo.org>

4 <https://digital.csic.es>

5 <https://archaeo.peercommunityin.org>

6 <https://www.biorxiv.org>

7 <https://socopen.org>

8 <https://github.com/benmarwick/archaeopaperswithcode> (consultado 06/07/2022).

dedicados a la Arqueología Computacional, más allá de la aplicación de los SIG. Además, la formación de profesionales dedicados en la actualidad en estos centros frecuentemente se ha completado en el extranjero, lo cual conlleva que la mayoría de las publicaciones de este cariz acaben destinándose a editoriales anglosajonas. Por último, y refiriéndonos a la formación, debe mencionarse que todavía existe, no únicamente en la investigación arqueológica hispano-lusa, sino también entre la comunidad internacional, cierta reticencia a la aplicación de técnicas cuantitativas en la disciplina. El debate entre los pros y contras del empleo de métodos cuantitativos en Arqueología, por más que interesante, trasciende en todo caso los objetivos del presente trabajo.

5. CONCLUSIONES: PROTOCOLO HACIA UNA ARQUEOLOGÍA EN ABIERTO Y TRANSPARENTE

Este trabajo se ha focalizado en la revisión, bajo el prisma de la Arqueología en abierto, de los diferentes artículos de temática computacional publicados en revistas hispano-lusas indexadas en el primer cuartil durante el sexenio 2015-2021. A partir de la identificación de los diferentes problemas, no parece inoportuno remarcar la necesidad de elaborar un protocolo focalizado en la transparencia y la reproducibilidad de las investigaciones arqueológicas. En definitiva, un protocolo de concienciación sobre la Arqueología en abierto. Por ello, la propuesta debe focalizarse en los dos sectores implicados como son el sistema editorial y la autoría.

La necesidad de una conciencia editorial que debe abarcar tanto la propia elaboración del manuscrito como los criterios de revisión. Por este motivo cualquier revista debería tener un manifiesto indicando de forma clara y concisa la adopción de la ciencia en abierto y la obligación de publicar los datos y/o código en aras de la transparencia y la reproducibilidad de los resultados. Paralelamente, el propio equipo editorial y de redacción deben plantear un cambio en los criterios de revisión de manuscritos. En este sentido, junto a las exigencias asociadas a la propia calidad *lato sensu* del mismo debería estipularse realizar un análisis de la reproducibilidad y la disponibilidad de los datos. Y, al igual que un artículo es rechazado si no supera la revisión desde una perspectiva de contenido científico, también debería ser descartado, hasta su subsanación, si no hay transparencia en el trabajo: esto significa que no puede accederse a los datos o bien los resultados no se pueden reproducir. Aunque existen numerosas opciones para alojar los datos, tal y como se ha indicado en el apartado anterior, consideramos oportuno que las propias revistas habiliten en sus plataformas una sección para incorporar los datos y códigos asociados a cada trabajo. Por otro lado, la autoría debe abandonar las reticencias para compartir los datos, ya indicadas anteriormente, y realizar revisiones focalizadas también en la parte de transparencia y reproducción de la investigación. No hay duda de que, si ambos sectores realizan este viraje hacia una ciencia histórico-arqueológica en abierto, las carencias observadas en este

trabajo serán superadas.

En definitiva, a pesar de que en los últimos años se está realizando un esfuerzo en aras de la transparencia, desgraciadamente aún queda mucho camino por recorrer. Por ello, este trabajo se ha focalizado en señalar la situación actual en torno a la accesibilidad de los datos como punto de partida para concienciar a la comunidad científica en la necesidad de adoptar una política del acceso abierto que abarque no solo a la propia publicación, siguiendo las directrices de la ANECA, así como los principios FAIR, sino también el acceso abierto de los datos y otra información necesaria para la reproducción del trabajo.

6. AGRADECIMIENTOS

Este trabajo se ha realizado en el contexto de los proyectos *ArchBiMod* (MSCA-IF: 101020631) y el proyecto PID2021-127141NA-I00 financiado por MCIN/AEI/10.13039/501100011033/ y por FEDER *una manera de hacer Europa*. SPG es beneficiario del programa Ramón y Cajal (ayuda RYC2021-033700-I) financiada por MCIN/AEI/10.13013/501100011033 y por la Unión Europea NextGenerationEU/PRTR. ACN es beneficiario del programa Marie Skłodowska-Curie Actions (H2020-MSCA-IF-2020) financiado por el programa Horizonte 2020 de la Comisión Europea.

Gracias a las dos personas revisoras cuyos comentarios nos han permitido incrementar la calidad del trabajo.

7. REFERENCIAS

- AGUILLO, I.F. (2020): «Altmetrics of the Open Access Institutional Repositories: a webometrics approach», *Scientometrics*, 123: 1181-92. <https://doi.org/10.1007/s11192-020-03424-6>.
- AMMERMAN, A.J. y CAVALLI-SFORZA, L. L. (1971): «Measuring the Rate of Spread of Early Farming in Europe», *Man*, 6 (4): 674-88.
- BARTON, M. (2013): «Stories of the Past or Science of the Future? Archaeology and Computational Social Science», en A. BEVAN y M. LAKE (Ed.), *Computational Approaches to Archaeological Spaces*, Routledge, Nueva York: 151-78.
- BEVAN, A. (2015): «The Data Deluge», *Antiquity*, 89 (348): 1473-84. <https://doi.org/10.15184/aqy.2015.102>.
- CALAPRICE, A. (2011): *The ultimate quotable Einstein*, Princeton University Press, Princeton-New Jersey.
- CONTE, R.; GILBERT, N.; BONELLI, G.; CIOFFI-REVILLA, C.; DEFFUANT, G.; KERTESZ, J.; LORETO, V.; MOAT, S.; NADAL, J-P.; SÁNCHEZ, A.; NOWAK, A.; FLACHE, A.; SAN MIGUEL, M.; HELBING, D. (2012): «Manifiesto of computational social science», *The European Physical Journal Special Topics*, 214: 325-46.
- CORTELL-NICOLAU, A.; GARCÍA-PUCHOL, O.; BARRERA-CRUZ, M.; GARCÍA-RIVERO, D. (2021): «The Spread of Agriculture in Iberia through Approximate Bayesian

- Computation and Neolithic Projectile Tools», *PLOS ONE*, 16 (12): e0261813. <https://doi.org/10.1371/journal.pone.0261813>.
- DERKSEN, M. (2020): «Open Psychology: transparency and reproducibility», *Psicología, Conocimiento y Sociedad* 9, (2): 318-30.
- DORAN, J. E. (1970): «Systems theory, computer simulations, and archaeology», *World Archaeology*, 1: 289-98.
- FLANNERY, K.V. (1973): «Archaeology with a capital S», en C. REDMAN (Ed.), *Research and theory in current archaeology*, John Wiley & Sons, New York: 47-52.
- GARDIN, J.C. (1988): *Artificial Intelligence and Expert System: a Case Study in the Knowledge Domain in Archaeology*, John & Wiley Sons, Hoboken.
- GERNSBACHER, M.A. (2018): «Writing Empirical Articles: Transparency, Reproducibility, Clarity and Memorability», *Advances in Methods and Practices in Psychological Science*, 1 (3): 403-414.
- HAIBE-KAINS, B.; ADAM, G.A.; HOSNY, A.; KHODAKARAMI, F.; WALDRON, L.; WANG, B.; MCINTOSH, C.; GOLDENBERG, A.; KUNDAJE, A.; GREENE, C.S.; BRODERICK, T.; HOFFMAN, M.M.; LEEK, J.T.; KORTHAUER, K.; HUBER, W.; BRAZMA, A.; PINEAU, J.; TIBSHIRANI, R.; HASTIE, T.; IOANNIDIS, J.P.A.; QUACKENBUSH, J.; AERTS, H.J. W. L. (2020): «Transparency and reproducibility in artificial intelligence», *Nature*, 586 (7829): E14-16. <https://doi.org/10.1038/s41586-020-2766-y>.
- ISERN, N.; FORT, J.; CARVALHO, A. F.; GIBAJA, J. F.; IBAÑEZ, J. J. (2014): «The Neolithic Transition in the Iberian Peninsula: Data Analysis and Modeling», *Journal of Archaeological Method and Theory*, 21 (2): 447-60. <https://doi.org/10.1007/s10816-013-9193-4>.
- LINDSAY, D.S. (2017): «Sharing data and materials in Psychological Science», *Psychological Science*, 28 (6): 669-702.
- MARWICK, B. (2017): «Computational Reproducibility in Archaeological Research: Basic Principles and a Case Study of Their Implementation», *Journal of Archaeological Method and Theory*, 24: 424-50. <https://doi.org/10.1007/s10816-015-9272-9>.
- MARWICK, B. (2020): «Open Access to Publications to Expand Participation in Archaeology», *Norwegian Archaeological Review*, 53 (2): 163-69. <https://doi.org/10.1080/00293652.2020.1837233>.
- MARWICK, B.; D'ALPOIM GUEDES, J.; BARTON, C.M.; BATES, L.A.; BAXTER, M.; BEVAN, A.; BOLLWERK, E.A.; BOCINSKY, R.K.; BRUGHMANS, T.; CARTER, A.K.; CONRAD, C.; CONTRERAS, D.; COSTA, S.; CREMA, E.R.; DAGGET, A.; DAVIES, B.; DRAKEM L.B.; DYE, T.S.; FRANCE, P.; FULLAGAR, R.; GIUSTI, D.; GRAHAM, S.; HARRIS, M.D.; HAWKS, J.; HEARTH, S.; HUFFER, D.; KANSA, E.C.; KANSA, S.W.; MADEN, M.E.; MELCHER, J.; NEGRE, J.; NEIMAN, F.; OPITZ, R.; ORTON, D.C.; PRZYSTUPA, P.; RAVIELE, M.; RIEL-SALVATORE, J.; RIRIS, P.; ROMANOWSKA, I.; SMITH, J.; STRUPLER, N.; ULLAH, I.I.; VAN VLACK, H.G.; VAN VALKENBURGH, N.; WATRALL, E.C.; WEBSTER, C.; WELLS, J.; WINTERS, J.; WREN C.D. (2017): «Open science in archaeology», *SAA Archaeological Record*, 17 (4): 8-14.
- MATZIG, D.N.; SHUMON, T.H.; RIEDE, F. (2021): «Design Space Constraints and the Cultural Taxonomy of European Final Palaeolithic Large Tanged Points: A Comparison of Typological, Landmark-Based and Whole-Outline Geometric

- Morphometric Approaches», *Journal of Paleolithic Archaeology*, 4 (4): 27. <https://doi.org/10.1007/s41982-021-00097-2>.
- OKUMURA, M. y ARAUJO, A.G.M. (2019): «Archaeology, Biology, and Borrowing: A Critical Examination of Geometric Morphometrics in Archaeology», *Journal of Archaeological Science*, 101 (enero): 149-58. <https://doi.org/10.1016/j.jas.2017.09.015>.
- ORENGO, H.A.; GARCIA-MOLSOSA, A.; BERGANZO-BESGA, I.; LANDAUER, J.; ALIENDE, P.; TRES-MARTÍNEZ, S. (2021): «New Developments in Drone-Based Automated Surface Survey: Towards a Functional and Effective Survey System», *Archaeological Prospection*, 28 (4): 519-26. <https://doi.org/10.1002/arp.1822>.
- PARDO-GORDÓ, S.; BERNABEU, J.; JIMÉNEZ-PUERTO, J.; ARMERO, C.; GARCÍA-DONATO, G. (2022): «The Chronology of Archaeological Assemblages Based on an Automatic Bayesian Procedure: Eastern Iberia as Study Case», *Journal of Archaeological Science*, 139: 105555. <https://doi.org/10.1016/j.jas.2022.105555>.
- RAMÍREZ-SÁNCHEZ, M. (2018): «Las Ciencias de la Antigüedad en España en la encrucijada: retos y oportunidades en la sociedad digital», en E. ORTIZ-DE-URBINA y J.M. VALLEJO (Ed.), *Métodos y técnicas en Ciencias de la Antigüedad: estudios sobre investigación y docencia*, Universidad del País Vasco, Vitoria: Anejos de Veleia 16: 13-36.
- RUBIO-CAMPILLO, X.; BLE, E.; PUJOL, À.; SALA, R.; TAMBA, R. (2022): «A Spatial Connectivity Approach to Landscapes of Conflict: Julius Caesar and the Assault to Puig Ciutat (NE Iberian Peninsula)», *Journal of Archaeological Method and Theory*, 29: 1059-1089. <https://doi.org/10.1007/s10816-022-09549-7>.
- SCHAUER, P.; BEVAN, A.; SHENNAN, S.; EDINBOROUGH, K.; KERIG, T.; PARKER PEARSON, M. (2020): «British Neolithic Axehead Distributions and Their Implications», *Journal of Archaeological Method and Theory*, 27 (4): 836-59. <https://doi.org/10.1007/s10816-019-09438-6>.
- SNITKER, G.; DIEZ CASTILLO, A.; BARTON, C. M.; BERNABEU AUBÁN, J.; GARCÍA PUCHOL, O.; PARDO-GORDÓ, S. (2018): «Patch-Based Survey Methods for Studying Prehistoric Human Land-Use in Agriculturally Modified Landscapes: A Case Study from the Canal de Navarrés, Eastern Spain», *Quaternary International*, 483 (julio): 5-22. <https://doi.org/10.1016/j.quaint.2018.01.034>.
- TEDERSOO, L.; KÜNGAS, R.; ORAS, E.; KÖSTER, K.; EENMAA, H.; LEIJEN, Ä.; PEDASTE, M.; RAJU, M.; ASTAPOVA, A.; LUKNER, H.; KOGERMANN, K.; SEPP, T. (2021): «Data sharing practices and data availability upon request differ across scientific disciplines», *Scientific Data*, 8 (1): 192. <https://doi.org/10.1038/s41597-021-00981-0>.
- VENEZIANO, A.; CAZENAVE, M.; ALFIERI, F.; PANETTA, D.; MARCHI, D. (2021): «Novel Strategies for the Characterization of Cancellous Bone Morphology: Virtual Isolation and Analysis», *American Journal of Physical Anthropology*, 175 (4): 920-30. <https://doi.org/10.1002/ajpa.24272>.
- WALLIS, J.C.; RONALDO, E.; BORGMAN, C.L. (2013): «If we share data, will anyone use them? Data Sharing and Reuse in the Long Tail of Science and Technology», *PLOS ONE*, 8 (7): e67332.
- WATRALL, E.C. (2016): «Archaeology, the Digital Humanities, and the “Big Tent”»,

en M. K. GOLD y L.F. KLEIN (Ed.), *Debates in the Digital Humanities 2016*, University of Minnesota Press, Minneapolis: 345-58.

- WILKINSON, M.D.; DUMONTIER, M.; AALBERSBERG, I. J.; APPLETON, G.; AXTON, M.; BAAK, A.; BLOMBERG, N.; BOITEN, J.W.; DA SILVA SANTOS, L.B.; BOURNE, P.E.; BOUWMAN, J.; BROOKES, A.J.; CLARK, T.; CROSAS, M.; DILLO, I.; DUMON, O.; EDMUNDS, S.; EVELO, C.T.; FINKERS, R.; GONZALEZ-BELTRAN, A.; GRAY, A.J.G.; GROTH, P.; GOBLE, C.; GRETHE, J.S.; HERINGA, J.; HOEN, P.A.C.; HOOFT, R.; KUHN, T.; KOK, R.; KOK, J.; LUSHER, S.J.; MARTONE, M.E.; MONS, A.; PACKER, A.L.; PERSOSON, B.; ROCCA-SERRA, P.; ROOS, M.; SCHAIK, R.; SANSONE, R.A.; SCHULTES, E.; SENGSTAG, T.; SLATER, T.; STRAWN, G.; SWERTZ, M.A.; THOMPSON, M.; VAN DER LEI, J.; VAN MULLIGEN, E.; VELTEROP, J.; WAAGMEESTER, A.; WITTENBURG, P.; WOLSTENCROFT, K.; ZHAO, J.; MONS, B. (2016): «The FAIR Guiding Principles for Scientific Data Management and Stewardship», *Scientific Data*, 3 (1): 160018. <https://doi.org/10.1038/sdata.2016.18>.
- WOBST, H.M. (1974): «Bounday conditions for Paleolithic social system: a simulation approach», *American Antiquity*, 39: 147-78.
- ZHANG, P.; WHILTE, J.; SCHMIDT, D.C.; LENZ, G.; TRENT ROSENBLUM, S. (2018): «FHIRChain: Applying blockchain to securely and Scalably Share Clinical Data», *Computational and Structural Biotechnology Journal*, 16: 267-278.

