

AUTENTICIDAD DE LOS DATOS DE INVESTIGACIÓN: ANÁLISIS DE MODELOS DE METADATOS EN LOS REPOSITORIOS ESPAÑOLES DE RE3DATA

AUTHENTICITY OF RESEARCH DATA: ANALYSIS OF METADATA STANDARDS IN SPANISH REPOSITORIES AT RE3DATA

Marieta MARKS LÖW¹ y Josefa GALLEGO LORENZO²

¹ *Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Santa Catarina. Brasil*

marietalow@gmail.com

 <https://orcid.org/0000-0002-4442-9365>

² *Departamento de Patrimonio Artístico y Documental. Área de Biblioteconomía y Documentación. Universidad de León. España*

josefa.gallego@unileon.es

 <https://orcid.org/0000-0002-5218-517X>

RESUMEN: Este estudio analiza los modelos de metadatos utilizados en repositorios españoles de la plataforma RE3DATA y su principal objetivo es evaluar los mecanismos de verificación que garantizan la autenticidad de los datos de investigación en estos repositorios, analizando la implementación de estándares de metadatos. La autenticidad de los datos es un aspecto esencial para la gestión de datos de investigación en el contexto de la ciencia abierta y para la confiabilidad y la reproducibilidad de los resultados científicos, ya que los modelos de metadatos desempeñan un papel fundamental en la preservación digital. Para ello se ha aplicado un enfoque metodológico de carácter bibliográfico y exploratorio orientado al análisis comparativo de los modelos más utilizados. Se identificaron 63 repositorios españoles en el agregador de repositorios RE3DATA y los cinco modelos más utilizados fueron: Dublin Core, DataCite, OAI-ORE, DDI e ISO 19115. Los resultados demuestran que Dublin Core es el modelo más utilizado, a pesar de ser el que presenta más debilidades estructurales para garantizar la autenticidad. En el análisis se aprecia que hay diferencias significativas entre los cinco modelos y se sugiere crear directrices tanto nacionales como internacionales claras sobre la autenticidad en datos para garantizar la confiabilidad y la reutilización de los datos científicos a largo plazo. Nuestra principal aportación consiste en proporcionar una evaluación sobre las fortalezas y las debilidades de los modelos de metadatos en cuanto a la autenticidad,

con información relevante para la toma de decisiones de gestores y responsables de políticas científicas para la preservación digital y la gestión de datos de investigación.

PALABRAS CLAVE: autenticidad; datos de investigación; modelos de metadatos; principios FAIR; repositorios de datos.

ABSTRACT: This study analyzes the metadata models used in Spanish repositories on the RE3DATA platform. Its main objective is to evaluate the verification mechanisms that guarantee the authenticity of research data in these repositories, analyzing the implementation of metadata standards. Data authenticity is an essential aspect of research data management in the context of open science and for the reliability and reproducibility of scientific results, as metadata models play a fundamental role in digital preservation. To this end, a bibliographic and exploratory methodological approach was applied, aimed at the comparative analysis of the most widely used models. Sixty-three Spanish repositories were identified in the RE3DATA repository aggregator, and the five most widely used models were: Dublin Core, DataCite, OAI-ORE, DDI, and ISO 19115. The results show that Dublin Core is the most widely used model, despite being the one with the most structural weaknesses in terms of ensuring authenticity. The analysis shows that there are significant differences between the five models and suggests that clear national and international guidelines on data authenticity should be created to ensure the long-term reliability and reuse of scientific data. Our main contribution is to provide an assessment of the strengths and weaknesses of metadata models in terms of authenticity, with relevant information for decision-making by managers and scientific policy makers for digital preservation and research data management.

KEYWORDS: authenticity; data repositories; FAIR principles; metadata models; research data.

1. INTRODUCCIÓN

Vivimos en una sociedad «datificada» donde los datos digitalizados forman parte de nuestras vidas y cualquier operación en internet deja rastro de datos e información que, una vez analizados, pueden convertirse en conocimiento (Martín González e Iglesias Rodríguez, 2023). La ciencia siempre ha usado datos como materia prima para la investigación y el avance del conocimiento y estos datos pueden ser de diferente naturaleza y procedencia, como datos de investigación (ciencia abierta), datos públicos (*open data*) o datos gubernamentales. Autores como Abadal y Anglada (2020) sostienen que la ciencia abierta es un cambio de modelo en la manera de hacer ciencia que supone realizar con una visión «abierta» (*open*) todas las etapas o fases de la investigación científica (diseño, recolección de datos, revisión, publicación, etc.).

Describir y documentar los datos constituye una de las acciones más relevantes en el proceso de gestión, ya que permitirá conocer el origen de los datos, contextualizarlos para posteriormente preservarlos, visualizarlos y recuperarlos adecuadamente (Cunha *et al.*, 2024). Por ello, la autenticidad de los datos es esencial para garantizar la validez e implica que los datos sean originales, obtenidos de manera ética y sin alteraciones que comprometan su integridad.

Su autenticidad es esencial para garantizar la reproducción y la reutilización confiable de los datos.

La correcta gestión, reutilización y preservación de datos de investigación representa desafíos del ecosistema científico actual. El exceso de datos puede generar duda acerca de la confiabilidad de los datos, especialmente en un momento en que las redes sociales permiten que los datos de investigación sean utilizados para generar desinformación y contenido sensacionalista (Jurado, 2023). Por lo tanto, contar con repositorios de datos de investigación que utilicen modelos de metadatos capaces de garantizar la autenticidad de los datos es una necesidad cada vez más reconocida por la comunidad científica.

Para que los datos sean seguros deben estar en repositorios confiables que cumplan con los requerimientos FAIR y los principios TRUST de transparencia, responsabilidad, foco en el usuario, sostenibilidad y tecnología, componentes esenciales para definir repositorios de datos confiables (De Giusti, 2021).

En este contexto, los principios FAIR (Findable, Accessible, Interoperable, Reusable) y TRUST (Transparency, Responsibility, User Focus, Sustainability, Technology) se han convertido en marcos normativos clave para la evaluación y el diseño de repositorios confiables. FAIR establece los criterios mínimos para asegurar la disponibilidad y la reutilización de los datos, mientras que TRUST se centra en los atributos organizativos e infraestructurales que garantizan la confianza en los repositorios digitales. La autenticidad, aunque no explícita en ambos marcos, es transversal a sus componentes, particularmente en lo que respecta a la integridad y la trazabilidad de los datos.

La garantía de autenticidad es crucial para asegurar la reproducibilidad de los resultados de investigación, un principio fundamental de la ciencia abierta (Silva y Silva, 2020). Los datos de investigación, como bienes estratégicos del patrimonio científico, exigen modelos de metadatos que aseguren su autenticidad durante su ciclo de vida, particularmente en repositorios abiertos (Gallego Lorenzo *et al.*, 2024).

Este estudio tiene por objetivo evaluar los mecanismos de verificación de autenticidad en repositorios españoles de la plataforma RE3DATA, analizando la implementación de estándares de metadatos para preservación digital. En la metodología se comparan cinco modelos predominantes (Dublin Core, DataCite, OAI-ORE, DDI e ISO 19115), utilizando como referencia las directrices de RECOLECTA sobre integridad y autenticidad. RECOLECTA, junto con iniciativas internacionales como las OpenAIRE Guidelines, de los datos FAIR (Wilkinson *et al.*, 2016) y las recomendaciones de COAR, promueven buenas prácticas para repositorios digitales en aspectos como descubrimiento, acceso, reutilización e integridad de los contenidos.

Los metadatos desempeñan un papel fundamental en la preservación digital y en la garantía de autenticidad, al proporcionar información estructurada que describe, contextualiza y localiza recursos digitales, facilitando su recuperación, uso y gestión a largo plazo. La autenticidad se define como la propiedad de un objeto digital de corresponder a su estado original y de no haber sufrido modificaciones no documentadas, siendo este un aspecto central para la confiabilidad de los datos de investigación. La autenticidad, en términos archivísticos y documentales, se refiere a la propiedad de un objeto digital de ser lo que pretende ser y de no

haber sido alterado de manera no autorizada. Para que esta autenticidad sea comprobable, es indispensable que los datos estén acompañados de metadatos capaces de documentar su procedencia, registrar transformaciones, establecer relaciones lógicas entre versiones y permitir la verificación técnica mediante mecanismos como *checksums* o firmas digitales. De tal modo que los sistemas de metadatos deben incluir elementos que registren la procedencia de los datos, documenten todas las transformaciones aplicadas y establezcan mecanismos de verificación para asegurar su integridad (Duranti y Rogers, 2022; Pacheco *et al.*, 2023).

En el contexto español, iniciativas como RECOLECTA (Recolector de Ciencia Abierta) han establecido directrices para la evaluación de repositorios institucionales (FECYT, 2021), incorporando criterios específicos sobre seguridad, autenticidad e integridad de los datos. Estas directrices están alineadas con estándares internacionales como las OpenAIRE Guidelines y las recomendaciones de COAR (Confederation of Open Access Repositories), que promueven buenas prácticas para repositorios digitales en aspectos como descubrimiento, acceso, reutilización e integridad de los contenidos (COAR, 2022).

El objetivo principal del trabajo es ofrecer una evaluación comparativa de estos modelos de metadatos desde la perspectiva de la autenticidad digital, considerando tanto sus fortalezas como sus limitaciones en el contexto específico de los repositorios españoles. Se busca así contribuir a la discusión sobre buenas prácticas en la gestión de datos de investigación y proporcionar insumos para la toma de decisiones por parte de gestores de repositorios, desarrolladores de políticas científicas y la comunidad académica en general.

2. LA AUTENTICIDAD DE LOS DATOS EN REPOSITARIOS DIGITALES

La preservación digital se ha consolidado como un campo estratégico dentro de la Ciencia de la Información, orientado a asegurar el acceso a largo plazo de los recursos digitales mediante procesos técnicos, organizativos y normativos que salvaguardan su integridad, legibilidad y autenticidad. En este contexto, los principios FAIR y TRUST han sido adoptados ampliamente como marcos de referencia para evaluar la calidad y la sostenibilidad de los repositorios de datos de investigación.

Los principios FAIR –Findable, Accessible, Interoperable y Reusable–, formulados por Wilkinson *et al.* (2016), establecen que los datos deben estar adecuadamente descritos mediante metadatos; contar con identificadores persistentes y sistemas de recuperación; ser accesibles bajo condiciones claras, interoperables mediante vocabularios estandarizados, y reutilizables con información sobre su procedencia, licencias y contexto. La autenticidad se vincula estrechamente con los principios de interoperabilidad y reutilización, ya que la validez de los datos reutilizados depende de su integridad y trazabilidad documentada.

Por su parte, los principios TRUST –Transparency, Responsibility, User focus, Sustainability y Technology–, propuestos por Lin *et al.* (2020), aportan una perspectiva institucional y organizativa. TRUST promueve la existencia de políticas claras sobre preservación, estructuras de gobernanza responsables, atención a las necesidades de las comunidades usuarias, sostenibilidad financiera y adopción de tecnologías robustas. Aunque no establece criterios

técnicos específicos, este marco refuerza la necesidad de repositorios que integren prácticas orientadas a mantener la autenticidad documental de los datos científicos en contextos complejos y cambiantes.

En conjunto, FAIR y TRUST articulan una visión holística de la preservación digital que supera las preocupaciones técnicas aisladas e incorpora dimensiones éticas, legales y de gobernanza. La autenticidad de los datos, como atributo transversal, requiere no solo mecanismos técnicos, sino también marcos institucionales que garanticen su implementación coherente y sistemática.

La noción de autenticidad en entornos digitales ha sido objeto de creciente atención por parte de la archivística contemporánea, particularmente, desde el proyecto InterPARES (International Research on Permanent Authentic Records in Electronic Systems), que define la autenticidad como la confianza en que un documento es lo que declara ser y no ha sido alterado sin autorización durante su ciclo de vida (Duranti, 2007).

En el contexto de los datos de investigación, la autenticidad se articula a través de tres dimensiones principales: integridad, procedencia y fijeza (Duranti *et al.*, 2022). La integridad se refiere a la totalidad y la consistencia del objeto digital; la procedencia, a su origen y contexto de producción; y la fijeza, a su estabilidad material desde el punto de vista informático. Estos elementos deben estar documentados y ser verificables por medio de evidencias registradas en metadatos estructurados.

Entre los mecanismos más frecuentes para garantizar la autenticidad se encuentran los *checksums* (funciones *hash* que permiten verificar la integridad del archivo); las firmas digitales (que certifican la autoría y el no repudio), y, en escenarios más recientes, la tecnología *blockchain*, que permite registrar transacciones y cambios en cadenas inmutables, aunque su aplicación a gran escala en repositorios aún está en fase experimental.

Es importante destacar que la autenticidad digital no puede garantizarse de manera retrospectiva; requiere ser diseñada desde la planificación del ciclo de vida de los datos y sustentada en políticas, procedimientos y estructuras tecnológicas integradas. Por tanto, la elección del modelo de metadatos y su implementación rigurosa son elementos cruciales para esta garantía. Como señalan Kanza y Knight (2022), la gestión efectiva de los datos de investigación debe planificarse desde las primeras fases del proyecto, ya que la ausencia de dicha planificación puede comprometer la integridad y la reutilización futura de los datos. Igualmente ocurre con los datos insertados en los repositorios.

2.1. MODELOS DE METADATOS Y LA AUTENTICIDAD

Los metadatos, entendidos como datos sobre los datos, cumplen funciones esenciales en la preservación digital: describen contenido, contexto y estructura de los objetos digitales; documentan su historia de gestión, y facilitan su recuperación, interoperabilidad y autenticación. En el ámbito específico de la preservación, los metadatos deben incluir elementos que permitan rastrear la procedencia, registrar transformaciones, aplicar mecanismos de verificación y establecer relaciones entre versiones.

Existen diversos modelos de metadatos orientados a diferentes dominios científicos y técnicos. Los más ampliamente utilizados en repositorios de datos de investigación incluyen Dublin Core, DataCite y OAI-ORE (Khan *et al.*, 2024). Cada uno de estos esquemas incorpora, en distintos grados, elementos relevantes para la documentación de la autenticidad.

La literatura especializada ha destacado la necesidad de emplear modelos de metadatos que vayan más allá de la descripción bibliográfica básica e incorporen estructuras capaces de reflejar el ciclo de vida de los datos, sus agentes responsables, eventos de transformación, licencias, versiones y validaciones (Silva y Silva, 2021). Elementos como *provenance*, *version*, *checksum*, *data curator* y *lineage* son fundamentales para este propósito.

En este sentido, diversas organizaciones han desarrollado recomendaciones que orientan la implementación de metadatos con foco en la autenticidad. Las directrices de RECOLECTA (FECYT, 2021), por ejemplo, establecen que los repositorios deben contar con políticas que aseguren la integridad, la trazabilidad y la seguridad de los datos, y recomiendan el uso de estándares interoperables que incluyan elementos de verificación. Asimismo, las OpenAIRE Guidelines for Data Archives promueven la adopción de prácticas de metadatos alineadas con FAIR y que faciliten la documentación de la procedencia, las condiciones de acceso y los derechos de uso (OpenAIRE, 2022).

En el ámbito internacional, COAR (Confederation of Open Access Repositories) ha desarrollado un Framework of Good Practices que articula principios de interoperabilidad, sostenibilidad y calidad en la gestión de repositorios, insistiendo en la necesidad de documentar eventos de preservación, auditorías, cambios y autorizaciones como evidencia de la autenticidad (COAR, 2022).

Por lo tanto, el análisis de los modelos de metadatos debe considerar no solo su adopción generalizada, sino también su capacidad estructural para registrar la procedencia y garantizar la trazabilidad de los datos a través del tiempo. Esta perspectiva resulta clave para evaluar en qué medida los repositorios españoles cumplen con los requisitos internacionales de autenticidad y qué mejoras podrían ser implementadas para fortalecer su confiabilidad a largo plazo.

3. METODOLOGÍA

Este estudio adopta un enfoque metodológico de carácter exploratorio y bibliográfico, orientado al análisis comparativo de la capacidad de los modelos de metadatos implementados en repositorios españoles para garantizar la autenticidad de los datos de investigación. La naturaleza exploratoria se justifica por la escasez de investigaciones previas que aborden de manera sistemática esta dimensión en el contexto específico de España. Asimismo, la dimensión comparativa permite establecer contrastes entre distintos esquemas de metadatos, mientras que el enfoque bibliográfico garantiza la investigación en términos conceptuales y normativos a través de la literatura actualizada.

El proceso metodológico se desarrolló en las siguientes fases:

1. Identificación de repositorios españoles registrados en RE3DATA (Registry of Research Data Repositories), un registro global de repositorios de datos de investigación.

2. Recopilación de información sobre los modelos de metadatos implementados en cada repositorio identificado y selección de los modelos de metadatos más representativos según su prevalencia en los repositorios analizados, excluyendo aquellos que utilizan solamente un repositorio.
3. Evaluación de la conformidad de estos modelos con directrices nacionales e internacionales.

El análisis se centró en repositorios registrados en RE3DATA, por tratarse de un popular registro global de repositorios de datos de investigación, reconocido por su cobertura, fiabilidad y nivel de actualización. RE3DATA (Registry of Research Data Repositories) es una iniciativa internacional lanzada en 2012 con el objetivo de proporcionar un punto de acceso global y estandarizado de información sobre estos repositorios de datos de investigación. Surgió como un proyecto conjunto entre DataCite, la German Research Foundation (DFG) y varias instituciones académicas europeas, consolidándose posteriormente como un servicio de referencia en el ecosistema de la ciencia abierta (Pampel *et al.*, 2013). RE3DATA no solo funciona como registro, sino también como mecanismo de monitoreo de la evolución de los repositorios en términos de calidad, cobertura temática y cumplimiento de estándares.

Para el estudio se realizó una selección de todos los repositorios con sede institucional en España, identificando sus características, políticas de metadatos y cumplimiento con estándares internacionales.

A partir del total de 63 repositorios españoles registrados en RE3DATA (consulta realizada en febrero de 2024), se procedió a identificar los modelos de metadatos declarados por cada uno. Se contabilizaron un total de 10 modelos diferentes; para el análisis comparativo se seleccionaron solo aquellos que fueron utilizados por dos o más repositorios con el fin de garantizar una muestra representativa y evitar casos atípicos.

Los cinco modelos seleccionados fueron:

- Dublin Core (33 repositorios)
- DataCite Metadata Schema (7 repositorios)
- OAI-ORE (Open Archives Initiative - Object Reuse and Exchange) (4 repositorios)
- DDI (Data Documentation Initiative) (3 repositorios)
- ISO 19115 (2 repositorios)

Además del análisis documental de los modelos de metadatos y de los repositorios, se consultaron normativas relevantes como las OpenAIRE Guidelines for Data Archives, el Community Framework for Good Practices de COAR (2022) y la *Guía para la evaluación de repositorios institucionales* de RECOLECTA (FECYT, 2023), con el fin de establecer un marco normativo para la evaluación.

Cada modelo fue analizado con base en su estructura formal, sus elementos obligatorios y opcionales y su implementación concreta en los repositorios españoles identificados. Además, se consultaron perfiles de aplicación locales o extensiones institucionales para evaluar la personalización del modelo base.

Se reconocen varias limitaciones metodológicas. En primer lugar, la información declarada por los repositorios en RE3DATA puede no reflejar de forma completa o actualizada las prácticas reales de implementación de los esquemas de metadatos. En segundo lugar, no todos los repositorios disponen de documentación pública detallada, lo que obligó a inferir prácticas a partir de descripciones generales. Y, en tercer lugar, la evaluación no incluye análisis empírico de conjuntos de datos ni inspección directa de registros XML, lo que limita el alcance a una evaluación de tipo estructural y normativo.

Asimismo, el estudio no contempla aspectos relacionados con la gobernanza institucional, los recursos técnicos disponibles ni el grado de capacitación de los agentes responsables, factores que pueden incidir en la calidad de la implementación del modelo. Finalmente, la rápida evolución del ecosistema de ciencia abierta puede generar cambios significativos a corto plazo, por lo que se recomienda considerar estos resultados como una radiografía situada en un momento específico del desarrollo de los repositorios españoles.

4. RESULTADOS Y DISCUSIÓN

El análisis inicial de los datos levantados en el estudio reveló la existencia de 63 repositorios españoles registrados en RE3DATA, los cuales implementan un total de 10 modelos diferentes de metadatos. Aplicando el criterio de exclusión para aquellos modelos utilizados por un único repositorio, se seleccionaron cinco esquemas para un análisis en profundidad: Dublin Core (33 repositorios), DataCite Metadata Schema (7 repositorios), OAI-ORE - Open Archives Initiative Object Reuse and Exchange (4 repositorios), DDI - Data Documentation Initiative (3 repositorios) e ISO 19115 (2 repositorios). Este análisis ha permitido identificar diferencias significativas en los modelos de metadatos utilizados y, en particular, en su capacidad para garantizar la autenticidad de los datos de investigación. A continuación, presentamos un análisis detallado de cada modelo y una comparación entre los modelos en términos de su capacidad para garantizar la autenticidad de los datos en los repositorios de investigación.

4.1. DUBLIN CORE: FLEXIBILIDAD VERSUS ESPECIFICIDAD

El Dublin Core Metadata Initiative (DCMI) surgió en 1995 como un esfuerzo colaborativo orientado para crear un conjunto básico de elementos descriptivos que facilitara la interoperabilidad entre distintos sistemas de información. Originalmente concebido para describir recursos web, Dublin Core ha evolucionado hasta convertirse en uno de los estándares de metadatos más utilizados a nivel internacional, gracias a su simplicidad, flexibilidad y amplia adopción (DCMI, 2025).

La estructura original está compuesta por 15 elementos básicos, como *title*, *creator*, *subject*, *description* y *identifier*, que permiten representar de forma generalizada la información sobre un recurso digital. Estos elementos pueden ser extendidos o refinados mediante cualificadores, lo que ha dado lugar a numerosas adaptaciones locales y perfiles de aplicación.

En el contexto de los repositorios españoles registrados en RE3DATA, Dublin Core es el modelo predominante, implementado en 33 de los 63 repositorios analizados. Esta preferencia se explica por su bajo umbral técnico de implementación, compatibilidad con los protocolos OAI-PMH y amplia disponibilidad en plataformas de gestión como DSpace o EPrints, comúnmente empleadas por instituciones académicas.

No obstante, esta generalización ha derivado en una adopción frecuentemente limitada a los elementos más básicos del estándar, sin cualificaciones ni extensiones orientadas a la preservación digital o autenticidad de los datos de investigación.

Una de las principales fortalezas de Dublin Core es su alta interoperabilidad semántica y sintáctica, derivada de su estandarización por la ISO (ISO 15836) y su compatibilidad con múltiples formatos de serialización como XML y RDF. Su uso generalizado favorece la agregación de metadatos en catálogos nacionales e internacionales y facilita la integración en ecosistemas de ciencia abierta.

Además, su estructura simple lo convierte en una opción accesible para repositorios con recursos técnicos limitados o en etapas iniciales de desarrollo. También permite ser utilizado como núcleo sobre el cual pueden integrarse otros esquemas más especializados mediante esquemas compuestos o mapeos.

Pese a sus ventajas en términos de adopción e interoperabilidad, Dublin Core presenta carencias significativas en relación con la autenticidad digital. En su forma estándar, no contempla elementos explícitos para:

- Documentar la procedencia detallada de los datos (ej. agentes de recolección, transformación o validación).
- Registrar versiones, modificaciones o transformaciones del objeto digital.
- Implementar mecanismos de verificación técnica, como *checksums* o firmas digitales.
- Representar relaciones complejas entre múltiples objetos digitales interdependientes.

Algunos repositorios han intentado subsanar estas limitaciones mediante la incorporación de elementos adicionales o el uso de perfiles de aplicación locales. No obstante, estas soluciones no siempre están normalizadas ni documentadas públicamente, lo que compromete la interoperabilidad y dificulta la evaluación sistemática de su eficacia.

Entre los perfiles más relevantes se encuentran el Dublin Core Qualified y el uso de metadatos compuestos, integrando esquemas como PREMIS o METS. Sin embargo, la adopción de estos perfiles en repositorios españoles es aún incipiente y heterogénea, y no existe un consenso sobre su implementación mínima recomendada para contextos de preservación.

En conclusión, Dublin Core representa una solución práctica y ampliamente adoptada para la descripción básica de recursos, pero su capacidad para garantizar la autenticidad de los datos de investigación es limitada sin complementos técnicos y normativos adicionales. Esta situación plantea un desafío estratégico para los repositorios españoles que aspiran a alinearse con los estándares de calidad y fiabilidad exigidos por la ciencia abierta contemporánea.

4.2. DATACITE METADATA SCHEMA: ROBUSTEZ EN LA VERIFICACIÓN

El DataCite Metadata Schema fue desarrollado por la organización DataCite como parte de su misión de facilitar la atribución, el descubrimiento y la citación de datos de investigación. Su objetivo principal es apoyar la asignación de identificadores persistentes (DOI) a conjuntos de datos y garantizar que estos estén acompañados por metadatos ricos, estructurados y alineados con principios FAIR (DataCite, 2021). A diferencia de Dublin Core, DataCite fue concebido específicamente para describir objetos digitales en el ámbito de la investigación científica, incorporando elementos diseñados para representar información clave sobre procedencia, integridad y versiones de los datos.

El esquema de DataCite incluye varios elementos particularmente relevantes para la verificación de la autenticidad de los datos, entre ellos:

- *version*: permite indicar la versión específica del conjunto de datos, facilitando la trazabilidad de modificaciones.
- *relatedIdentifiers*: establece relaciones entre objetos digitales, como versiones anteriores o archivos complementarios.
- *contributors* con roles específicos (ej. *DataCollector*, *DataCurator*, *DataManager*): documentan la cadena de custodia de los datos, permitiendo identificar responsables en distintas etapas del ciclo de vida.
- *formats* y *sizes*: ayudan a validar las características físicas del archivo.
- *description* con subtipos como *technicalInfo* o *methods*: ofrecen espacio para incluir documentación detallada de procesos técnicos o metodológicos.

Además, aunque no obligatorio, el esquema es compatible con la inclusión de *checksum* y *checksumType* como extensiones, lo que permite integrar mecanismos de verificación de integridad en la documentación del objeto digital.

Una de las contribuciones más significativas del esquema DataCite es su integración nativa con el sistema de identificadores persistentes DOI, lo cual asegura que cada conjunto de datos pueda ser referenciado de manera única, estable y verificable. Esta característica fortalece la trazabilidad y evita ambigüedades en la reutilización de datos. El DOI también actúa como ancla para metadatos actualizables, permitiendo reflejar cambios sin comprometer la integridad del identificador original.

A través de la estructura de *contributors* y del elemento *publisher*, el esquema permite registrar con claridad no solo quién creó los datos, sino también quién los recopiló, procesó, almacenó y curó. Esta capacidad para documentar la procedencia institucional y técnica de los datos es fundamental para verificar su autenticidad y evaluar su fiabilidad para usos secundarios.

En el contexto español, siete repositorios analizados en RE3DATA utilizan el esquema DataCite. Estos repositorios presentan una orientación más explícita hacia buenas prácticas de documentación y preservación. Algunos incluso integran DataCite con otros esquemas como PREMIS o METS, conformando estructuras de metadatos compuestas que fortalecen su capacidad para representar eventos, agentes y relaciones entre objetos digitales.

Aunque la implementación aún es limitada en número, los repositorios que adoptan DataCite muestran una tendencia a alinear sus políticas con los principios FAIR y con las directrices de iniciativas como RECOLECTA o OpenAIRE, incluyendo mecanismos de control de versiones, registro de transformaciones y documentación de integridad.

Comparado con Dublin Core, DataCite ofrece una estructura significativamente más rica y orientada a los desafíos actuales de la ciencia de datos. Su adopción favorece no solo la autenticidad, sino también la visibilidad y la interoperabilidad internacional. Se configura, así, como un modelo de referencia para repositorios que buscan elevar su nivel de calidad, trazabilidad y cumplimiento normativo.

4.3. OAI-ORE: AGREGACIONES Y RELACIONES COMPLEJAS

El modelo OAI-ORE (Open Archives Initiative - Object Reuse and Exchange) fue desarrollado por la Open Archives Initiative para representar y describir agregaciones de recursos digitales distribuidos en la web. Su objetivo es permitir la especificación formal de conjuntos de objetos interrelacionados que, aunque almacenados por separado, forman una unidad lógica –como colecciones de datos, publicaciones científicas vinculadas a conjuntos de datos, o versiones sucesivas de un mismo archivo (Lagoze *et al.*, 2008)–.

OAI-ORE utiliza como base el modelo de datos Resource Description Framework (RDF) y serializaciones como RDF/XML o Atom. La unidad fundamental es el Resource Map (ReM), que describe la agregación (*Aggregation*) y sus recursos componentes, así como las relaciones entre ellos mediante propiedades RDF estándar o personalizadas.

El esquema OAI-ORE permite describir explícitamente relaciones semánticas y estructurales complejas, como versiones, derivados, documentación asociada, conjuntos de archivos o elementos multimedia vinculados. Además, ofrece la posibilidad de declarar metadatos sobre la agregación como un todo y sobre cada uno de sus componentes individualmente.

Esto convierte a OAI-ORE en una solución adecuada para contextos donde los datos de investigación no son archivos únicos, sino objetos compuestos distribuidos que requieren gestión colectiva y coherente.

De los 63 repositorios españoles identificados en RE3DATA, 4 implementan el modelo OAI-ORE. Estos repositorios suelen tener una orientación avanzada en cuanto a gestión de objetos digitales complejos, especialmente en disciplinas que generan conjuntos de datos estructurados, multimedia o vinculados a otras formas de documentación científica.

Sin embargo, el análisis revela que la implementación efectiva de OAI-ORE es aún incipiente y desigual. En algunos casos, el modelo se emplea como complemento de otros esquemas, sin aprovechar plenamente sus capacidades descriptivas. Además, no todos los repositorios proporcionan documentación pública sobre su uso específico de agregaciones, lo que dificulta evaluar su integración con prácticas de verificación de autenticidad.

Desde la perspectiva de la autenticidad digital, OAI-ORE presenta ventajas notables en cuanto a trazabilidad relacional. Su estructura permite declarar explícitamente qué recursos conforman un conjunto de datos, qué versiones o transformaciones han ocurrido y cómo se relacionan con otros objetos digitales. Esta capacidad de reconstruir el contexto documental

es esencial para establecer la autenticidad en entornos donde la procedencia y la organización interna del contenido son fundamentales.

4.3.1. Desafíos de implementación

Pese a su potencial, OAI-ORE enfrenta varios desafíos para su adopción generalizada:

- Complejidad técnica: requiere conocimiento de RDF y tecnologías semánticas, lo que limita su implementación a instituciones con capacidades avanzadas.
- Falta de estandarización en perfiles de aplicación: la libertad estructural que ofrece el modelo puede generar inconsistencia semántica entre implementaciones.
- Escasa integración con mecanismos de verificación técnica como *checksums* o firmas digitales, los cuales deben ser implementados mediante esquemas adicionales.

OAI-ORE no fue diseñado originalmente como un modelo de preservación, pero su capacidad para describir relaciones y dependencias lo convierte en una herramienta valiosa para representar la procedencia y la estructura lógica de los datos. Integrado con otros modelos orientados a la verificación técnica, puede fortalecer significativamente la capacidad de los repositorios para garantizar la autenticidad de objetos digitales complejos.

4.4. DDI: ESPECIALIZACIÓN EN CIENCIAS SOCIALES

La Data Documentation Initiative (DDI) es un modelo de metadatos desarrollado específicamente para describir datos utilizados en las ciencias sociales, en particular aquellos derivados de encuestas, estudios longitudinales, censos y otras formas de investigación empírica. Desde su primera versión en 2000, DDI ha evolucionado hacia una estructura modular y jerárquica capaz de describir todo el ciclo de vida de los datos, desde su recolección hasta su preservación y diseminación (Vardigan y Whiteman, 2007).

DDI ha sido promovido por un consorcio internacional de instituciones académicas y centros de datos, y cuenta con múltiples versiones (incluidas DDI-Codebook y DDI-Lifecycle), cada una adaptada a distintos niveles de complejidad y granularidad documental.

DDI destaca por su capacidad para representar la estructura interna, la historia de transformación y el contexto metodológico de los datos. El modelo permite describir:

- Instrumentos de recolección (ej. cuestionarios, entrevistas).
- Procesos de codificación, limpieza y edición.
- Versiones y modificaciones sucesivas.
- Información sobre anonimización y acceso.
- Procedencia institucional y técnica de cada etapa del ciclo.

A través de elementos como *Methodology*, *DataCollection*, *Processing*, *Version*, *Lineage* y *OtherMaterial*, DDI documenta con alta precisión cada evento asociado a la creación y la

modificación del conjunto de datos. Esta capacidad resulta crítica para establecer la autenticidad documental, especialmente en contextos donde la replicabilidad y la trazabilidad metodológica son esenciales.

Una de las características distintivas de DDI es su enfoque en el modelo de ciclo de vida completo. A diferencia de otros esquemas orientados principalmente a la descripción bibliográfica o técnica, DDI permite representar las transiciones temporales y estructurales que experimentan los datos desde su concepción hasta su archivo. Esta representación es esencial para validar que los cambios sufridos fueron autorizados, documentados y atribuibles a agentes identificados.

DDI ofrece mecanismos explícitos para asegurar la autenticidad, incluyendo:

- Registro detallado de transformaciones metodológicas (*processing steps*)
- Identificación de actores responsables por cada fase del proceso
- Versión del estudio y justificación de cambios
- Relación entre versiones y objetos derivados
- Información sobre restricciones de acceso y derechos de uso

Además, su diseño permite integrarse con otros esquemas como PREMIS o METS, incorporando funcionalidades adicionales para verificación técnica y preservación a largo plazo.

En el ámbito español, tres repositorios especializados en ciencias sociales utilizan DDI como modelo principal de metadatos. Estos repositorios están vinculados a instituciones académicas con tradición en investigación empírica y manejo de encuestas longitudinales. Su implementación de DDI suele estar alineada con estándares internacionales y con buenas prácticas documentales, incluyendo metadatos enriquecidos, versiones detalladas y políticas claras de curación de datos.

Aunque minoritaria en número, la presencia de DDI refleja un nivel de madurez documental elevado en estos repositorios, y sugiere un compromiso con la transparencia metodológica y la reproducibilidad científica.

DDI es uno de los pocos esquemas diseñados explícitamente para capturar la dimensión temporal, metodológica y contextual de los datos. Su potencial para garantizar la autenticidad es particularmente alto en disciplinas donde los datos son construidos mediante procesos complejos y donde la validación metodológica es parte del proceso científico.

Por tanto, se configura como un modelo de referencia para aquellos repositorios que gestionan datos cuantitativos estructurados, especialmente en áreas como sociología, economía, ciencia política y psicología.

4.5. ISO 19115: ESTÁNDAR PARA INFORMACIÓN GEOGRÁFICA

La norma ISO 19115 forma parte de la familia de estándares ISO 19100, desarrollados por el comité técnico ISO/TC 211 para abordar las necesidades específicas de la información geoespacial. Publicada inicialmente en 2003 y actualizada en versiones posteriores (ISO 19115-1:2014; ISO 19115-2:2019), esta norma proporciona un marco integral para describir

recursos geográficos digitales, con especial énfasis en calidad, procedencia, referencia espacial y condiciones de uso (ISO, 2019).

Su adopción ha sido promovida por agencias nacionales de cartografía, instituciones académicas, infraestructuras de datos espaciales (SDI) y organismos internacionales como INSPIRE (en Europa) y GEOSS (a nivel global).

ISO 19115 se basa en un modelo de datos modular y extensible, articulado en torno a elementos centrales como identificación, calidad, distribución, referencia temporal y espacial y mantenimiento. Además, permite la descripción detallada de:

- Procedencia de los datos (*lineage*)
- Procesos de transformación y edición
- Restricciones legales y técnicas de uso
- Formatos y estructuras del archivo
- Contactos institucionales y técnicos responsables

El estándar se implementa comúnmente utilizando XML bajo el esquema Geographic Markup Language (GML) y se integra con otros estándares relacionados, como ISO 19139 (para serialización) e ISO 19157 (para evaluación de calidad).

4.5.1. *Linaje y calidad de datos*

Uno de los componentes más relevantes de ISO 19115 para la autenticidad es el elemento LI_Lineage, que documenta el historial completo de producción del recurso geoespacial. Incluye:

- Descripción de los procesos de obtención y transformación
- Identificación de fuentes utilizadas
- Agentes responsables de cada fase
- Métodos de verificación y control de calidad

Este enfoque estructurado proporciona evidencia detallada sobre la fiabilidad, la exactitud y la procedencia de los datos geográficos, facilitando su evaluación crítica y la reconstrucción de su historia documental.

En el análisis realizado, dos repositorios españoles especializados en información geográfica utilizan el modelo ISO 19115. Estos repositorios están asociados a centros de investigación o servicios cartográficos que operan en coordinación con redes nacionales o europeas de datos espaciales. Su implementación de ISO 19115 suele estar integrada en sistemas más amplios de gestión geoespacial, como Infraestructuras de Datos Espaciales (IDE), lo que asegura una alineación con estándares internacionales.

Aunque su número es reducido, estos repositorios presentan un alto nivel de sofisticación técnica y documental, y sus metadatos incluyen información detallada sobre calidad, procesamiento y validación, lo cual refuerza significativamente la verificación de autenticidad.

ISO 19115 incorpora múltiples mecanismos para garantizar la autenticidad de los datos:

- Documentación estructurada del linaje técnico e institucional
- Control de calidad a través de evaluaciones normalizadas

- Identificación de agentes y fuentes en cada etapa del ciclo de vida
- Integración con sistemas de preservación digital y auditoría

Si bien no incluye de forma nativa elementos como *checksums* o firmas digitales, estos pueden ser incorporados a través de extensiones o en combinación con otros estándares complementarios (ej. METS, PREMIS).

Una ventaja estratégica de ISO 19115 es su alta interoperabilidad con otros modelos, permitiendo su integración en infraestructuras científicas multidisciplinares. Su adopción facilita el cumplimiento de los principios FAIR, particularmente en lo relativo a la trazabilidad y la reutilización de los datos, siempre que se implementen mecanismos técnicos adicionales de verificación.

En síntesis, ISO 19115 representa una solución robusta y especializada para describir y verificar la autenticidad de datos geoespaciales, siendo particularmente adecuada para repositorios con funciones cartográficas, ambientales o territoriales.

4.6. ANÁLISIS COMPARATIVO DE MODELOS

Para evaluar de manera sistemática las capacidades de los cinco modelos de metadatos analizados (Dublin Core, DataCite, OAI-ORE, DDI e ISO 19115) en relación con la verificación de autenticidad, se construyó una matriz comparativa basada en tres dimensiones fundamentales: procedencia, integridad y trazabilidad. A cada modelo se le asignó una puntuación cualitativa (alta, media o baja) según su estructura formal y las prácticas observadas en su implementación por los repositorios españoles.

Tabla 1. Matriz comparativa

Modelo	Procedencia	Integridad técnica	Trazabilidad / versionado
Dublin Core	Baja	Baja	Baja
DataCite	Alta	Media-alta	Alta
OAI-ORE	Media	Baja	Media-alta
DDI	Alta	Media	Alta
ISO 19115	Alta	Alta	Alta

La Tabla 1 revela que los esquemas más robustos en autenticidad son aquellos especializados y diseñados para contextos específicos, como ISO 19115 en datos geográficos, DDI en ciencias sociales y DataCite en datos científicos en general. En contraste, Dublin Core, pese a su amplia adopción, presenta debilidades estructurales en los tres aspectos evaluados.

Un análisis DAFO (debilidades, amenazas, fortalezas y oportunidades) puede resumir cada uno de los modelos, desde la perspectiva de la autenticidad:

- Dublin Core
 - Debilidades: escasa capacidad para documentar versiones y agentes; falta de soporte para verificación técnica.
 - Fortalezas: amplia adopción, interoperabilidad básica.
 - Amenazas: estancamiento normativo; dependencia de perfiles ad hoc.
 - Oportunidades: desarrollo de extensiones interoperables; mapeo con modelos más ricos.
- DataCite
 - Debilidades: requiere integración adicional para preservación técnica.
 - Fortalezas: identificación persistente; buena documentación de procedencia.
 - Amenazas: subutilización en repositorios de tamaño pequeño.
 - Oportunidades: alineación con FAIR; interoperabilidad con ORCID, Crossref y OpenAIRE.
- OAI-ORE
 - Debilidades: falta de estandarización semántica; escaso soporte documental en español.
 - Fortalezas: representación de agregaciones complejas.
 - Amenazas: alta complejidad técnica.
 - Oportunidades: uso combinado con otros modelos; trazabilidad relacional avanzada.
- DDI
 - Debilidades: especializado; curva de aprendizaje elevada.
 - Fortalezas: ciclo de vida completo de los datos; contexto metodológico.
 - Amenazas: implementación limitada fuera de ciencias sociales.
 - Oportunidades: interoperabilidad con estándares estadísticos; uso creciente en estudios longitudinales.
- ISO 19115
 - Debilidades: no nativo en ambientes no geoespaciales.
 - Fortalezas: documentación exhaustiva del linaje y calidad.
 - Amenazas: bajo número de implementaciones en repositorios generales.
 - Oportunidades: integración con IDE; cumplimiento con políticas europeas (INSPIRE).

El análisis sugiere que no existe un modelo «universalmente superior», sino que la selección debe responder a las características del dominio disciplinar, la complejidad de los datos y la infraestructura tecnológica del repositorio. Por ejemplo:

- *DataCite* resulta adecuado para repositorios interdisciplinarios que priorizan la citación y trazabilidad de datos.
- *DDI* es idóneo para archivos de datos de encuestas y estudios sociales.

- *ISO 19115* se recomienda para información geoespacial que requiere validación precisa del linaje.
- *OAI-ORE* es útil en repositorios que gestionan objetos digitales compuestos o agregados.
- *Dublin Core* puede mantenerse como capa básica interoperable, pero debería complementarse con otros esquemas más específicos.

El patrón de adopción observado en los repositorios españoles refleja una tensión entre facilidad de implementación y sofisticación funcional. Mientras que Dublin Core domina en número, su limitada capacidad técnica plantea desafíos para la calidad y la autenticidad de los datos preservados. En contraste, los esquemas más avanzados (como DataCite, DDI o ISO 19115) tienden a ser utilizados por repositorios con mayor madurez técnica o enfoque temático específico.

Asimismo, se observan tendencias emergentes de hibridación, en las que un modelo básico (como Dublin Core) se utiliza junto a modelos más ricos, mediante extensiones, mapeos o perfiles de aplicación locales. Esta estrategia permite mejorar la capacidad de verificación sin abandonar la interoperabilidad mínima.

Por último, la evolución normativa internacional –particularmente, en torno a los principios FAIR y TRUST– está generando incentivos para que los repositorios revisen sus modelos de metadatos y adopten soluciones que permitan una verificación estructurada de la autenticidad, como requisito para la certificación y el financiamiento público.

Por lo tanto, se puede concluir que Dublin Core, pese a ser el modelo más utilizado (33 repositorios), ofrece una estructura limitada para documentar procedencia, verificar integridad o registrar cambios en los datos. Su implementación en la mayoría de los casos se reduce a los elementos básicos, sin cualificadores ni extensiones orientadas a la preservación digital. La generalización de este modelo parece obedecer más a su facilidad técnica e integración con plataformas comunes (como DSpace) que a una decisión basada en criterios de autenticidad documental.

En contraste, DataCite (7 repositorios), DDI (3) e ISO 19115 (2) demuestran un mayor grado de especialización y robustez estructural para documentar aspectos críticos de la autenticidad. DataCite destaca por su integración con identificadores persistentes y su capacidad para representar agentes y versiones; DDI sobresale por su exhaustividad en describir procesos metodológicos y transformaciones de datos; e ISO 19115 aporta mecanismos precisos para rastrear el linaje y la calidad de datos geográficos.

OAI-ORE, por su parte, ofrece un enfoque distintivo orientado a representar agregaciones complejas, lo que resulta útil en contextos donde los objetos digitales están distribuidos y organizados de forma modular. No obstante, su implementación en repositorios españoles es limitada y técnicamente heterogénea.

Los resultados sugieren varias oportunidades concretas para fortalecer la capacidad del sistema español en materia de autenticidad de datos de investigación:

- Revisión y ampliación de los perfiles de metadatos existentes, especialmente en repositorios que utilizan Dublin Core, incorporando calificadores, elementos de preservación y mecanismos de verificación técnica.
- Fomento de la adopción de modelos especializados, como DataCite, DDI o ISO 19115, mediante capacitación técnica, incentivos institucionales y marcos de apoyo normativo.
- Desarrollo de perfiles de aplicación interoperables que integren múltiples esquemas en una arquitectura coherente, permitiendo mantener la simplicidad en el acceso y la interoperabilidad básica, sin renunciar a una descripción documental robusta.

Asimismo, podría promoverse el uso combinado de OAI-ORE y esquemas de preservación (como PREMIS), particularmente en repositorios que gestionan objetos digitales compuestos, versiones múltiples o conjuntos distribuidos.

5. CONCLUSIONES

Este estudio ha ofrecido una evaluación de cinco modelos de metadatos implementados en repositorios españoles de datos de investigación, con especial atención a su capacidad para garantizar la autenticidad digital. A partir del análisis de 63 repositorios registrados en RE3DATA, se identificaron fortalezas y limitaciones estructurales en los modelos Dublin Core, DataCite, OAI-ORE, DDI e ISO 19115.

Los resultados demuestran una brecha entre la adopción de modelos de metadatos y la efectiva implementación de mecanismos que garantizan la autenticidad de los datos en repositorios. Mientras Dublin Core continúa siendo el esquema predominante por su simplicidad, los modelos especializados como DataCite, DDI e ISO 19115 ofrecen soluciones más robustas para preservar la autenticidad, condición indispensable para la reproducibilidad científica y la confianza en los datos.

La investigación también destaca el papel estratégico de RE3DATA como registro global y herramienta de monitoreo del ecosistema de repositorios, así como la necesidad de alinear las prácticas nacionales con las directrices internacionales (FAIR, TRUST, COAR, OpenAIRE).

Estas observaciones plantean importantes consideraciones para el desarrollo futuro de repositorios de datos de investigación en España. La adopción de modelos de metadatos que incorporen elementos robustos para la verificación de autenticidad resulta fundamental para garantizar la confiabilidad y la reutilización de los datos científicos a largo plazo. Se recomienda la implementación de extensiones específicas o bien la adopción de esquemas complementarios que subsanen las limitaciones identificadas, particularmente en repositorios que utilizan Dublin Core como esquema principal.

Adicionalmente, se sugiere la creación de directrices específicas a nivel nacional, que establezcan requisitos mínimos para la documentación de autenticidad en los metadatos de datos de investigación, alineados con estándares internacionales como los Principios FAIR y las recomendaciones de COAR sobre buenas prácticas para repositorios.

Para concluir, este estudio contribuye al campo emergente de la curación y la preservación de datos de investigación, proporcionando una evaluación sistemática de las capacidades de los modelos de metadatos actuales para garantizar un aspecto crítico de la preservación digital: la autenticidad. Los resultados obtenidos pueden informar tanto a gestores de repositorios como a responsables de políticas científicas sobre las fortalezas y las debilidades de los diferentes esquemas de metadatos en relación con la verificación de autenticidad, facilitando decisiones informadas en la selección y la implementación de estos esquemas en futuros desarrollos.

La autenticidad de los datos de investigación constituye un pilar necesario para una ciencia abierta rigurosa, responsable y socialmente confiable. Garantizar esta autenticidad requiere no solo tecnología, sino también gobernanza, formación y compromiso institucional. En este sentido, los modelos de metadatos no son simples herramientas descriptivas, sino instrumentos de garantía epistemológica y transparencia científica.

REFERENCIAS BIBLIOGRÁFICAS

- Abadal Falgueras, E. y Anglada Ferrer, L. (2020). Ciencia abierta: cómo han evolucionado la denominación y el concepto. *Anales de Documentación*, 23(1), 1. <https://doi.org/10.6018/analesdoc.378171>
- COAR. (2022). *Community Framework for Good Practices in Repositories*, version 2. Zenodo. <https://zenodo.org/records/7108101>.
- Cunha, P., Marks Löw, M., Santos, A. A. dos y Cándido, A. C. (2024). Competência informacional para uso de Dados abertos: tendências e perspectivas. *Em Questao*, 30. <https://doi.org/10.1590/1808-5245.30.136545>
- DataCite. *Annual Report*. (2021). https://datacite.org/wp-content/uploads/2023/06/DataCite_2021_Annual_Report.pdf
- DataCite Metadata Working Group. (2024). *DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs*. Version 4.6. DataCite e.V. <https://doi.org/10.14454/mzv1-5b55>
- De Giusti, M. R. (2021). Calidad en los repositorios digitales. Los principios TRUST para repositorio de datos. *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*, 29, 55-59. <https://doi.org/10.24215/18509959.29.e6>
- DublinCore. (2025). *DublinCore Specifications*. <https://www.dublincore.org/specifications/dublin-core/>
- Duranti, L. (2007). The InterPARES 2 Project (2002-2007): An Overview. *Archivaria*, 64, 113-121. <https://www.archivaria.ca/index.php/archivaria/article/view/13155>
- Duranti, L., Rogers, C. y Thibodeau, K. (2022). Authenticity. *Archives and Records*, 43(2), 188-203. <https://doi.org/10.1080/23257962.2022.2054406>
- Fundación Española para la Ciencia y la Tecnología (FECYT). (2021). *Guía para la evaluación de repositorios institucionales de investigación v4.0*. <https://www.fecyt.es/publicaciones/guia-para-la-evaluacion-de-repositorios-institucionales-de-investigacion>
- Gallego Lorenzo, J., Martín González, Y. e Iglesias Rodríguez, A. (2024). Calidad y uso de datos confiables en la práctica docente universitaria. *Ibersid: Revista de Sistemas de Información y Documentación*, 18(2), 87-96. <https://doi.org/10.54886/ibersid.v18i2.5019>
- Jurado Martín, M. (2023). Competencia crítica y desinformación en teoría de la comunicación: Análisis en los programas de los grados en periodismo en España. *Documentación de las Ciencias de la Información*, 46(2), 169-176. <https://doi.org/10.5209/dcin.87513>

- Kanza, S. y Knight, N. J. (2022). Behind every great research project is great data management. *BMC Research Notes*, 15(20). <https://doi.org/10.1186/s13104-022-05908-5>
- Khan, A. M., Loan, F. A., Parray, U. Y. y Rashid, S. (2024). Global overview of research data repositories: an analysis of re3data registry. *Information Discovery and Delivery*, 52(1), 53-61. <https://doi.org/10.1108/IDD-07-2022-0069>
- Lagoze, C., Van De Sompel, H., Nelson, M. L., Warner, S., Sanderson, R. y Johnston, P. (2008). *Object re-use and exchange: A resource-centric approach*. Indianapolis. <http://arxiv.org/abs/0804.2273>
- Lin, D., Crabtree, J., Dillo, I., Downs, R., Edmunds, R., Giaretta, D., De Giusti, M., L'Hours, H., Hugo, W., Jenkyns, R., Khodiyar, V., Martone, M. E., Mokrane, M., Navale, V., Petters, J., Sierman, B., Sokolova, D. V., Stochhause, M. y Westbrook, J. (2020). The TRUST Principles for digital repositories. *Scientific Data*, 7(144). <https://www.nature.com/articles/s41597-020-0486-7>
- Martín González, Y. e Iglesias Rodríguez, A. (2023). Alfabetización en datos: Prácticas y escenarios formativos. *Revista EDICIC*, 3(3), 1-15. <https://ojs.edicic.org/revistaedicic/article/view/202>
- OpenAIRE. (2022). *OpenAIRE Guidelines*. <https://guidelines.openaire.eu/en/latest/index.html>
- Pacheco, A., Da Silva, C. G. y De Freitas, M. C. V. (2023). A metadata model for authenticity in digital archival descriptions. *Archival Science*, 23, 629-673. <https://doi.org/10.1007/s10502-023-09422-w>
- Pampel, H., Vierkant, P., Scholze, F., Bertelmann, R., Kindling, M., Klump, J., Goebelbecker, H.-J., Gundlach, J., Schirmbacher, P. y Dierolf, U. (2013) Making Research Data Repositories Visible: The re3data.org Registry. *PLoS ONE*, 8(11), e78080. <https://doi.org/10.1371/journal.pone.0078080>
- Silva, V. de A. y Silva, M. B. da. (2021). Metadatos para la preservación digital de datos abiertos: Un estudio de identificación. *Biblios Journal of Librarianship and Information Science*, 78, 44-60. <https://doi.org/10.5195/biblios.2020.793>
- Vardigan, M. y Whiteman, C. (2007). ICPSR meets OAIS: applying the OAIS reference model to the social science archive context. *Archival Science*, 7, 73-87. <https://doi.org/10.1007/s10502-006-9037-z>
- 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, Chris T., Finkers, R., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>