

Cuidar de los datos y ser justos con ellos: ¿debemos complementar los principios FAIR con los principios CARE?

Caring for and being fair with data: Should we complement the FAIR principles with the CARE principles?

Paloma Marín-Arraiza

Marín-Arraiza, Paloma (2023). "Cuidar de los datos y ser justos con ellos: ¿debemos complementar los principios FAIR con los principios CARE?". *Anuario ThinkEPI*, v. 17, e17a07.

<https://doi.org/10.3145/thinkepi.2023.e17a08>

Publicado en *IweTel* el 11 de abril de 2023

Paloma Marín-Arraiza

<https://orcid.org/0000-0001-7460-7794>

ORCID

p.arraiza@orcid.org



Resumen: En los últimos años, los principios FAIR se han incluido completamente en el diálogo sobre la gestión de datos de investigación desde un punto de vista técnico. Paralelamente han surgido los principios CARE (o CREA) (Control –Autoridad para controlar–, Responsabilidad, Ética, Aprovechamiento para el beneficio colectivo), que se centran en aspectos éticos y de gobernanza, con un enfoque en la soberanía de los datos indígenas. A diferencia de los principios FAIR, aún no existen modelos de madurez o de "CREAficación" para evaluar la adopción de los principios CARE, y su implementación requiere un desarrollo comunitario de servicios y gestión de datos. A pesar de las diferencias, una infraestructura que se adapta a los principios FAIR puede articularse más fácilmente a los principios CARE. En esta nota, se presentan los principios CARE y su posible encuadre más allá de la gestión técnica de los datos de investigación.

Palabras clave: Principios FAIR; Principios CARE; Ética; Gobernanza de datos; Infraestructura de información.

Abstract: In recent years, the FAIR (Findable, Accessible, Interoperable, and Reusable) principles have been fully included in the technical dialogue on research data management, and parallel to them, the CARE (Collective Benefit, Authority to Control, Responsibility, Ethics) principles have also emerged. These focus on ethical and governance aspects, placing particular emphasis on indigenous data sovereignty. Unlike the FAIR principles, there are as yet no maturity or "CREAfication" models to assess the adoption of the CARE principles, and their implementation requires community development of services and data management. Despite the differences, an infrastructure that conforms to the FAIR principles can be more easily linked with the CARE principles. This work presents the CARE principles and their possible framing beyond technical management of research data.

Keywords: FAIR principles; CARE principles; Ethics; Data governance; Information infrastructure.

1. Introducción

En los últimos años, hemos visto un enfoque creciente por la inclusión de los principios FAIR en las actividades de gestión de datos de investigación. En 2016 se publica el artículo *The FAIR guiding principles for scientific data management and stewardship* (Wilkinson et al., 2016) que oficializa el trabajo en torno a la encontrabilidad (*findability*), accesibilidad o acceso (*accessibility*), interoperabilidad (*interoperability*) y reusabilidad (*reusability*) de datos, metadatos e infraestructura. Actualmente, estos principios aparecen en guías de gestión institucional, políticas públicas e institucionales, federación de infraestructuras y proyectos de diversa índole, entre otros. Empezamos a hablar de la FAIRificación de datos, metadatos e infraestructuras. En el contexto español, la *Ley Orgánica 2/2023, de 22 de marzo, del Sistema Universitario (España, 2023)*, en su artículo 12 (Fomento de la ciencia abierta y ciencia ciudadana) también incide en dichos principios, aunque solo para datos.

En paralelo, surge el diálogo en torno a la compartición y preservación de los datos más allá de los aspectos técnicos o centrados en el dato en sí. Un diálogo basado en las relaciones y las condiciones históricas asociadas a la recopilación de datos y su repercusión en el uso ético y socialmente responsable de los datos recopilados. Tal diálogo tiene especial relevancia en los contextos de datos indígenas, aunque no únicamente. Es aquí donde surgen los principios CARE (*Collective Benefit, Authority to Control, Responsibility, Ethics*), o en castellano CREA (Control –Autoridad para controlar–, Responsabilidad, Ética, Aprovechamiento para el beneficio colectivo).

2. Los principios CARE (o CREA)

Los principios CREA surgen en 2019 a partir de la conferencia de la *Global Indigenous Data Alliance (GIDA)* que tuvo lugar precisamente en Oñati (Guipúzcoa) con el objetivo de entender y preservar la soberanía de los datos indígenas y lograr entender los aspectos sociales de la colecta, preservación y reuso de datos.

Los principios CREA se estructuran en cuatro bloques con tres subapartados cada uno.

Tabla 1. Resumen de los principios CREA

Control (Autoridad para controlar)	C1. Reconocimiento de derechos e intereses
	C2. Datos para gobernanza
	C3. Gobernanza de datos
Responsabilidad	R1. Para las relaciones positivas
	R2. Para ampliar capacidades
	R3. Para los idiomas y cosmovisiones indígenas
Ética	E1. Para minimizar daños y maximizar beneficios
	E2. Por la justicia
	E3. Para su uso futuro
Aprovechamiento para el beneficio colectivo	A1. Para el desarrollo inclusivo y la innovación
	A2. Para mejorar la gobernanza y el compromiso ciudadano
	A3. Para obtener resultados equitativos

Fuente: *Grupo de Interés en Soberanía de Datos Indígenas (2019)*

A diferencia del caso de los principios FAIR, aún existen carencias a la hora de evaluar la adopción de los principios CREA. No se tienen modelos de madurez ni de “CREAficación” como sí se tienen de FAIR. Si bien, sí que existen algunas iniciativas que están trabajando en modelos para operacionalizar los principios CREA en conjunto con los FAIR en repositorios y otras infraestructuras de información científica (Carroll et al., 2021).

Otra de las diferencias principales es que los principios FAIR hacen referencia a aspectos técnicos, cuya implementación y medición resulta mucho más directa. Sin embargo, los principios CREA hacen alusión a aspectos de ética y gobernanza, deben tener un desarrollo comunitario de servicios y gestión de datos.

De la misma forma que la correcta documentación e identificación de los datos de investigación y sus procesos, así como el uso de formatos abiertos e interoperables, contribuye a la FAIRificación de un conjunto de datos abierto (Higman et al., 2019), una infraestructura que se haya adaptado a los principios FAIR podrá articularse más fácilmente a los principios CREA (Carroll et al., 2021).

Cabe también señalar que no estaríamos ante un ámbito de aplicación cerrado, sino ante un contexto que se inicia con la preservación de y respeto a datos indígenas que puede extenderse a otros tipos de datos, como datos genéticos (anonimizados) o medioambientales (Carroll et al., 2022) tanto de dichas comunidades como de forma más amplia.

Por otro lado, la ética, la gobernanza de datos y el desarrollo para el bien común deben ser puntos principales de una correcta gestión de datos de investigación.

3. Ideas en el contexto español

El concepto de “datos indígenas” puede sonar un poco ajeno en el contexto español dado que no existen movimientos como el maorí o el guaraní en la península. Sin embargo, este concepto es más próximo a nuestra realidad de lo que pueda inicialmente parecer.

La propia *GIDA* toma su acrónimo de la palabra *guía* en euskera, siendo que en el contexto español también conviven otras lenguas —oficiales como el catalán, el gallego o el aranés y no oficiales como el asturiano o el castúo, así como el lenguaje silbado silbo gomero. Todas ellas asociadas a una serie de contextos culturales y conocimientos que también requieren de representación y gestión en investigación.

Así, un punto de trabajo de los principios CREA es la responsabilidad para los idiomas y cosmovisiones (R3) en el que se debate ampliamente si las infraestructuras actuales están preparadas para soportar metadatos en varios idiomas y cómo se deben adaptar a conocimientos más allá del textual.

También es posible correlacionar los aspectos éticos indicados en los principios CREA con aspectos éticos considerados en la investigación y con la reutilización indicada por los principios FAIR. De hecho, el objetivo de la investigación debería ser —aunque infelizmente no siempre lo sea— el uso futuro y la innovación inclusiva.

Cabe también el desarrollo más allá de la parte técnica de los principios FAIR en torno a la gobernanza de datos y a cómo la investigación científica debe facilitar el compromiso ciudadano. Aquí podemos destacar diversas acciones de ciencia ciudadana —como *LabIN Granada* o la *Oficina de Ciencia Ciudadana de Barcelona*— que se han acabado convirtiendo en beneficios para esa propia ciudadanía y en establecimiento de comunidades de práctica.

Como reflexión final, considerar los aspectos técnicos en la gestión de datos de investigación, sus metadatos y las infraestructuras donde están alojados a partir de los principios FAIR en combinación con aspectos éticos, de protección de datos, de gobernanza y responsabilidad para con la sociedad puede llevarnos a valorar la gestión de datos de investigación más allá de la productividad y en pro del desarrollo y la innovación comunes.

4. Referencias

Carroll, Stephanie R.; Herczog, Edit; Hudson, Maui; Russell, Keith; Stall, Shelley (2021). “Operationalizing the CARE and FAIR Principles for Indigenous data futures”. *Scientific data*, v. 8, n. 1, 108.

<https://doi.org/10.1038/s41597-021-00892-0>

Carroll, Stephanie R.; Plevel, Rebecca; Jennings, Lydia L.; Garba, Ibrahim; Sterling, Rogena; Cordova-Marks, Felina M.; Hiratsuka, Vanessa; Hudson, Maui; Garrison, Nanibaa’ A. (2022). “Extending the CARE Principles from tribal research policies to benefit sharing in genomic research”. *Frontiers in genetics*, v. 13, 1052620.

<https://doi.org/10.3389/fgene.2022.1052620>

España (2023). “Ley Orgánica 2/2023, de 22 de marzo, del Sistema Universitario”. *BOE*, n. 70, pp. 43267-43339.

<https://boe.es/boe/dias/2023/03/23/pdfs/BOE-A-2023-7500.pdf>

Grupo de Interés en Soberanía de Datos Indígenas (2019). *Principios CREA para la Gobernanza de Datos Indígenas*. Alianza Global de Datos Indígenas.

<https://static1.squarespace.com/static/5d3799de845604000199cd24/t/5f0621e031811d08f9c4cd2711594237427481/Principios+CREA.pdf>

Higman, Rosie; Bangert, Daniel; Jones, Sarah (2019). “Three camps, one destination: The intersections of research data management, FAIR and Open”. *Insights the UKSG Journal*, 32, 18.

<https://doi.org/10.1629/uksg.468>



Figura 1. Principios FAIR y CARE (o CREA)

Fuente: Grupo de Interés en Soberanía de Datos Indígenas (2019, p. 1).

Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, Ijsbrand-Jan; Appleton, Gabrielle; Axton, Myles; Baak, Arie; Blomberg, Niklas; Boiten, Jan-Willem; Da-Silva-Santos, Luiz B.; Bourne, Philip E.; Bouwman, Jildau; Brookes, Anthony J.; Clark, Tim; Crosas, Mercè; Dillo, Ingrid; Dumon, Olivier; Edmunds, Scott; Evelo, Chris T.; Finkers, Richard; ... Mons, Barend (2016). "The FAIR Guiding Principles for scientific data management and stewardship". *Scientific data*, n. 3, 160018. <https://doi.org/10.1038/sdata.2016.18>

* * *

COAR Task Force on Supporting Multilingualism and non-English Content in Repositories

Remedios Melero



En relación con el multilingüismo, COAR creó a finales de 2022 el *COAR Task Force on Supporting Multilingualism and non-English Content in Repositories* para desarrollar y promover buenas prácticas para los repositorios en la gestión de contenido multilingüe y que no está en inglés. El grupo de trabajo se centra en identificar buenas prácticas para metadatos, palabras clave multilingües, interfaces de usuario, traducciones, formatos, licencias e indexación que mejorarán la visibilidad del contenido multilingüe y no inglés en todo el mundo. Y su primera recomendación (seguirán apareciendo), es la siguiente (es un texto adaptado de <https://www.coar-repositories.org/news-updates/what-we-do/multilingual-and-non-english-content>, donde podéis encontrar más información)

Recomendación 1. Todos los registros en el repositorio deben incluir una etiqueta en el campo de metadatos de idioma que identifique el idioma del recurso y una etiqueta que identifique el idioma de los metadatos (incluso si los recursos están en inglés).

¿Por qué?

Esta es una recomendación muy simple, pero extremadamente poderosa. Cuando el idioma de los metadatos y el idioma del recurso se atribuyen correctamente, permite que los servicios de descubrimiento e indexación procesen y analicen correctamente el texto. La indexación implica prácticas de análisis de texto como la derivación, la lematización y el tratamiento adecuado de las palabras vacías, las cuales son específicas del idioma. La inclusión de la etiqueta de idioma permite a los buscadores de información, agregadores y otros servicios de descubrimiento identificar correctamente el idioma de los metadatos y el texto completo y tratar los elementos en consecuencia.

Remedios Melero
IATA-CSIC
rmelero@iata.csic.es

* * *

European Science Vocabulary

Paloma Marín-Arraiza

Otro buen ejemplo es la taxonomía que se utiliza en *CORDIS* y que se ha implementado en CRIS nacionales como el finlandés (*research.fi*), *European Science Vocabulary*. Aun así, suelen estar basados en los códigos de la ISO 639-1 y aún habría lenguas (especialmente aquellas no escritas) que no estarían representadas.

Esto es sin duda un gran desafío al que no se le encontrará una solución inmediata. Como hay que empezar por algo, las recomendaciones de COAR son un muy buen inicio.

Paloma Marín-Arraiza
p.arraiza@orcid.org