

SEMINÁRIO INTERNACIONAL DE PRESERVAÇÃO DIGITAL e
ENCONTRO DA REDE BRASILEIRA DE SERVIÇOS DE PRESERVAÇÃO DIGITAL
PROGRAMAS E PROJETOS DE PRESERVAÇÃO DIGITAL
junho de 2023, dias 14, 15 e 16



Mesa 1

Preservação digital de dados científicos com Marisa De Giusti e
Kadidja de Oliveira

Preservación, memoria, datos, principios FAIR y CARE

Vincular la preservación digital con la memoria de los pueblos, con los principios FAIR y los CARE

Dra. Marisa R. De Giusti

14 de junio de 2023



Esta obra está bajo una [Licencia Creative Commons](#)
Atribución-NoComercial-CompartirIgual 4.0 Internacional



Noción de preservación de UNESCO



“La preservación digital puede definirse como el conjunto de los procesos destinados a garantizar la continuidad de los elementos del patrimonio digital durante todo el tiempo que se consideren necesarios”.

“La mayor amenaza para la continuidad digital es la desaparición de los medios de acceso. No puede decirse que se han conservado los objetos digitales si, al haber dejado de existir los medios de acceso a ellos, resulta imposible utilizarlos. El objetivo de la preservación de los objetos digitales es mantener su accesibilidad, es decir, la capacidad de tener acceso a su mensaje o propósito esencial y auténtico”. (UNESCO, 2003: p. 37).



- Procesamiento**
- Sistematización
 - Depuración
 - Integración

- Razonamiento**
- Inductivo
 - Deductivo
 - Analítico
 - Sintético

- Reflexión**
- Concientización
 - Aprehensión
 - Interiorización

Datos

Representación simbólica de un atributo o variable cuantitativa o cualitativa

Información

Conjunto de datos convenientemente agrupados, estructurados e interpretados

Conocimiento

Información adquirida (teórica o empíricamente) sobre la que se razona, argumenta y concluye

Sabiduría

Carácter que se desarrolla a partir de generar mayor entendimiento y reflexión

Principios FAIR - Origen

SCIENTIFIC DATA

Open Access | Published: 15 March 2016

The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson, Michel Dumontier, [...] Barend Mons

Scientific Data 3, Article number: 160018 (2016) | Cite this article

135k Accesses | 1736 Citations | 1575 Altmetric | Metrics

El 15 de marzo de 2016 fue publicado en la revista Scientific Data de Nature el artículo: “*Principios FAIR para el manejo y administración de datos científicos*”. Los **Principios FAIR** ofrecen un conjunto de cualidades precisas y medibles que una publicación de datos debería seguir para que los datos sean Encontrables, Accesibles, Interoperables y Reutilizables (del inglés *FAIR – Findable, Accessible, Interoperable, and Reusable*).

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3(1), 1-9. <https://www.nature.com/articles/sdata201618>

Principios FAIR



- No son un estándar.
- Se pueden aplicar de manera gradual.
- Son importantes para avanzar cuando se trata de datos abiertos.
- Su ámbito (en principio) son los datos científicos.
- Se pueden aplicar usando distintas tecnologías o herramientas.
- **El acrónimo tiene que ver con equidad y justicia.**

Principios FAIR

1) Encontrable

- 1.1 A los (meta)datos se les asigna un identificador persistente y único global.
- 1.2 Los datos se describen con metadatos enriquecidos.
- 1.3 Los metadatos incluyen de forma clara y explícita el identificador de los datos que describen.
- 1.4 Los (meta)datos están registrados o indexados en un recurso de búsqueda.

2) Accesible

- 2.1 Los (meta)datos son recuperables por su identificador utilizando un protocolo de comunicaciones estandarizado.
 - 2.1.1 El protocolo es abierto, gratuito y universal.
 - 2.1.2 El protocolo permite un procedimiento de autenticación y autorización, cuando sea necesario.
- 2.2 Los metadatos son accesibles, incluso cuando los datos ya no están disponibles.











3) Interoperable

- 3.1 Los (meta)datos utilizan un lenguaje formal, accesible, compartido y de amplia aplicación para la representación del conocimiento.
- 3.2 Los (meta)datos utilizan vocabularios que siguen los principios FAIR.
- 3.3 Los (meta)datos incluyen referencias calificadas a otros (meta)datos.

4) Reutilizable

- 4.1 Los (meta)datos se describen detalladamente con una multitud de atributos precisos y relevantes.
 - 4.1.1 Los (meta)datos se publican con una licencia de uso de datos clara y accesible.
 - 4.1.2 Los (meta)datos están asociados con la procedencia detallada.
 - 4.1.3 Los (meta)datos cumplen con los estándares de la comunidad del dominio concreto.

En qué pensar cuando se trabaja con datos para hacerlos FAIR

-   Documentación y Metadatos
-   Formato de los Datos
-   Acceso a los Datos
-   Identificadores Persistentes
-   Licencias de Datos Abiertos



Sin importar la tipología de los datos, esto es lo que se precisa en la preservación digital: si se cumple con FAIR muchos aspectos requeridos para asegurar la preservación digital están cubiertos.



Documentación y metadatos

Documentación



Además de la búsqueda de recursos, los metadatos pueden ayudar a organizar los documentos electrónicos, Si se usan estándares, facilitar la interoperabilidad y la integración de estos recursos. **Son importantes pra la reusabilidad y reproducibilidad** y garantizar la identificación digital, el archivo y su conservación.



Documentación y metadatos

- 1) La descripción del proceso de recolección: qué instrumentos se usaron, si son datos de campo, cuál es la localización de la muestra, muchos etcétera.
- 2) Una descripción de la procedencia y el historial de los datos por ejemplo: quién los recolectó, quién los estandarizó, analizó y modificó los datos crudos.
- 3) Una descripción de los datos mismos (metadatos) por ejemplo para describir un archivo: con qué se lee el archivo (software), cuál es el tamaño del archivo, el formato del archivo, con qué software se puede abrir el archivo, quién generó el archivo, etcétera.
- 4) Un paso adicional (no es FAIR) es la estructuración de las carpetas del proyecto, cuando se definen las convenciones que se van a utilizar en el proyecto.
- 5) Utilizar archivos README cuyo nombre describa los datos que contiene, estructurado en secciones y con términos o vocabularios estándares disciplinarios cuando sea posible.
- 6) Llamar a los archivos/versiones siguiendo convenciones p.e.: 2023-04-18_Cultivo1_humedad_día_1.csv



- Mantener separados los archivos de datos, los métodos y las salidas.
- Los archivos de datos separados de archivos de códigos.
- Especificar la relación entre qué código opera sobre qué datos y en qué orden para producir qué salidas.

Metadatos y estándares

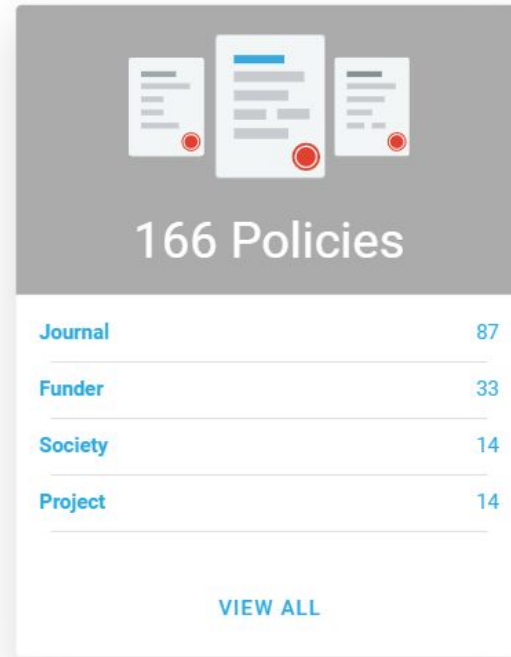
Los metadatos:

- son relevantes para acceder, entender y procesar datos por humanos y máquinas;
- saber el quién, qué, porqué, dónde y cuándo de los datos;
- pueden estar incluidos o ser externos a un archivo de datos.

Los estándares de metadatos:


- proveen una estructura que describe los datos con:
 - términos comunes (p.ej., en una disciplina) que permiten consistencia entre archivos o sets de datos;
 - definiciones comunes para una interpretación más fácil;
 - una estructura común para encontrar la información rápidamente.
- En la búsqueda y adquisición de datos proveen:
 - una estructura de documentación en un formato reconocible y confiable para la interpretación por una computadora/procesador;
 - una descripción uniforme y resumida de los datos.

Estándares de Metadatos ¿Dónde encontrarlos?



Fuente <https://fairsharing.org/>


Estándares de Metadatos ¿Dónde encontrarlos?

 standards, databases, policies

Search all of FAIRsharing


Standards Databases Policies Collections Add/Claim Content Stats PeteMcQ

Find

 **Recommendations**


Standards and/or databases recommended by journal or funder data policies.

Discover

 **Collections**

Standards and/or databases grouped by domain, species or organization.

Learn

 **Educational**

About standards, their use in databases and policies, and how we can help you.


<https://www.youtube.com/watch?v=xBIWXI6NA4>

Search all of FAIRsharing


Search

Standards Databases Policies Collections/Recommendations

Advanced Search

 Fine grained control over your search.

Search Wizard

 ask FAIRsharing

Let us guide you to your results.

Herramientas para crear metadatos



Data package Creator: <https://create.frictionlessdata.io/>



MRPHO (Ecological Metadata Language): <https://knb.ecoinformatics.org/tools/morpho>



Nesstar Publisher (Data Documentation Initiative - DDI) <http://www.nesstar.com/software/publisher.html>



Formato de los datos

- El formato en que se guardan y comparten los datos pueden ayudar a maximizar:
 - Interoperabilidad
 - Re-utilización
- Cuando sea posible elegir un formato de datos que:
 - sea abierto
 - contenga metadatos embebidos (mínimo o extendido);
 - sea sostenible (que permita el archivo de datos por un largo tiempo)
- Guías disponibles:
 - <http://www.ukdataservice.ac.uk/manage-data/format/recommended-formats>
 - <http://www.loc.gov/preservation/resources/rfs/>



Acceso a los datos

- Identificar tempranamente en el proyecto:
 - ¿Qué datos se pueden compartir?
 - ¿Con quién se pueden compartir? (abiertamente o con colaboradores)
 - ¿Cómo se pueden compartir?
- Si los datos no se pueden compartir abiertamente:
 - Proveer acceso a los metadatos
 - Indicar en los metadatos cómo se puede acceder a los datos

Identificadores persistentes

- Referencia duradera a un objeto digital
- Hace al objeto localizable y accesible
- Hace al objeto citable de una manera precisa



Ejemplos:



Libros - International Standard Book Number (ISBN)



Para Investigadores - ORCID



Publicaciones y Datos - Digital Object Identifiers (DOIs)



Licencias de datos abiertos



- Deben ser distribuidos de la manera **menos restrictiva posible**.
- Para eso se recomienda usar **Creative Commons CC0** or **CC-BY 4.0**.
- La **citación** apropiada es siempre una **buena práctica**.



La persona que asoció una obra con este resumen ha **dedicado** la obra al dominio público, mediante la renuncia a todos sus derechos a la obra bajo las leyes de derechos autorales en todo el mundo, incluyendo todos los derechos conexos y afines, en la medida permitida por la ley.

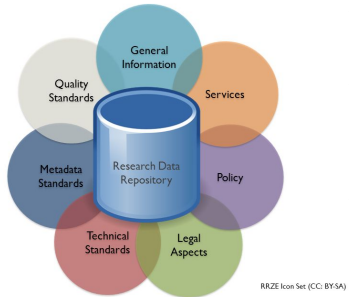
Puede **copiar, modificar, distribuir e interpretar la obra**, incluso para propósitos comerciales, sin pedir permiso.



Atribución — Usted debe dar crédito de manera adecuada, brindar un enlace a la licencia, e indicar si se han realizado cambios. Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que usted o su uso tienen el apoyo de la licenciante.

No hay restricciones adicionales — No puede aplicar términos legales ni medidas tecnológicas que restrinjan legalmente a otras a hacer cualquier uso permitido por la licencia.

Repositorios de datos



- Depositar datos en un repositorio ayuda a hacer los datos encontrables, accesibles y reutilizables.
- Es una herramienta que permite agregar metadatos.
- Es una herramienta que permite asignar un Identificador Persistente y una Licencia de Datos.
- Los datos publicados en un repositorio pueden ir acompañados de documentación relevante incluso de un PGD.
- Permite relacionar los datos con otros datos, publicaciones y referencias relevantes mejorando la interoperabilidad.

Si los datos no se pueden compartir abiertamente:

- Proveer **acceso a los metadatos**
- Indicar en los metadatos cómo se puede acceder a los datos

Acceso a los datos y a los datos indígenas

- Datos **FAIR** no siempre es equivalente a **Datos Abiertos**
- No se debe publicar datos confidenciales como:
 - datos personales;
 - datos que puedan comprometer la seguridad pública;
 - datos sujetos a un acuerdo de no-divulgación.



Ilustración realizada por Benny Ríos para la microserie “Yo me quedo en mi comunidad” <https://youtu.be/gT2ySRtJC78>

El concepto de **soberanía de los datos indígenas** es relativamente reciente pues la primera publicación importante sobre el tema apareció tan solo en 2016.

La soberanía de los datos indígenas se define como el derecho de los pueblos indígenas a controlar, consultar, poseer y ser propietarios de datos que proceden de ellos y que se refieren a sus miembros, sistemas de conocimiento, costumbres o territorios

RDA - Research Data Alliance

La **Research Data Alliance** es una organización global que dio inicio en marzo 2013 por iniciativa de la Comisión Europea, Estados Unidos y la Fundación de Ciencias de Estados Unidos y el gobierno de innovación de Australia, con el propósito de **construir los puentes técnicos que permiten el compartir el uso de datos.**

Principios

- Apertura
- Consenso
- Representación comunitaria
- Sin fines de lucro

Cuenta con más de **12.000 miembros** de 148 países.





- Grupos de trabajo (44)
- **Grupos de interés (57)**
- Grupos históricos
- Grupos de coordinación (16)
- Grupos nacionales (29)
- Comunidades de práctica

Grupo de interés internacional en **Soberanía de Datos Indígenas** (ID-Sov):

- Desarrollo y adopción de infraestructura que promueva el compartir datos y la investigación basada en datos indígenas para su uso particularmente por Pueblos Indígenas (PI).
- Desarrollo de capacidades relacionadas con el uso y manejo de datos más allá del ámbito académico, particularmente en beneficio de los PI.
- Proveer una plataforma internacional visible con relación a la SDI que integre e impulse el tema.
- Atraer nuevos actores incluyendo academia, usuarios de datos y por supuesto comunidades indígenas.

Comisión Económica para América Latina y el Caribe. (2021, julio 1). Gobernanza de datos indígenas: Principios FAIR y CARE [Text].
<https://www.cepal.org/es/notas/gobernanza-datos-indigenas-principios-fair-care>

FAIR/JUSTOS y CARE/CUIDADO

Los **Principios CARE/CREA** pretenden que las relaciones con los datos pasen de ser una consulta regulada a un diálogo basado en el valor que ponga en primer plano las culturas y los sistemas de conocimiento indígenas dentro de los ecosistemas de datos. Los Principios **CARE** complementan los Principios **FAIR**.

El hacer que los datos se ajusten tanto a los principios **CARE** como **FAIR** genera un espacio para difundir la procedencia, el protocolo y los permisos a lo largo del ciclo de vida de los datos con el fin de promover resultados y beneficios equitativos del acceso, uso, reutilización y atribución de los datos. La puesta en práctica de los principios **CARE** y **FAIR** para los datos de los pueblos indígenas requiere herramientas que guíen la inclusión del conocimiento indígena en los sistemas de datos.

¿Qué son los DATOS INDÍGENAS?

Datos, información y conocimiento, en cualquier formato, que genera algún impacto en los Pueblos Indígenas, naciones y comunidades ya sea a nivel colectivo o individual:

DATOS SOBRE RECURSOS Y TERRITORIOS

Tierras, agua, geología, títulos, aire, suelo, sitios sagrados, territorios, plantas, animales, etc.

DATOS COMO INDIVIDUOS

Administrativos, legales, salud, sociales, comerciales, productivos, servicios, etc.

DATOS COMO PUEBLO/NACIÓN

Información sobre tradiciones y cultura, historias orales, literatura y conocimiento de los pueblo, cuentos, etc.



Collaboratory for Indigenous Data Governance

Research, Policy, and Practice for Indigenous Data Sovereignty

INDIGENOUSDATALAB.ORG | [@Indigidatalab](https://twitter.com/Indigidatalab)
USINDIGENOUSDATA.ORG | [@USIDSN](https://twitter.com/USIDSN)
GIDA-GLOBAL.ORG | [@GidaGlobal](https://twitter.com/GidaGlobal)

Cite as:
Global Indigenous Data Alliance. (2022).
'Indigenous Data Sovereignty and
Governance.' Prepared by

Assistant Professor Stephanie R. Carroll, DrPH, MPH (Ahtna – Native Village of Kluti-Kaah)
Associate Director, Native Nations Institute
Director, Collaboratory for Indigenous Data Governance
University of Arizona

Jewel Cummins, Intern

Master of Library and Information Sciences
Master of Legal Studies
University of Arizona

Andrew Martinez, BSBA

Research Specialist, The Native Nations Institute
Citizen of the Salt River Pima-Maricopa Indian Community
Located within the traditional homelands of the Mary's River or Ampinefu Band of Kalapuya

Promoviendo el control Indígena de los datos Indígenas

- Avanzando hacia la Soberanía y Gobernanza de Datos Indígenas
- Afirmando los derechos e intereses de los Pueblos Indígenas en sus datos
- Abogando por el uso de datos para la autodeterminación y bienestar de los Pueblos Indígenas
- Reforzando el derecho de la participación en la toma de decisiones de acuerdo a los usos, costumbres, valores e intereses colectivos de los Pueblos Indígenas



Comisión Económica para América Latina y el Caribe. (2021, julio 1). Gobernanza de datos indígenas: Principios FAIR y CARE [Text]. <https://www.cepal.org/es/notas/gobernanza-datos-indigenas-principios-fair-care>

Principios CREA para la Gobernanza de Datos Indígenas



Carroll, SC, Garba, I, Figueroa-Rodríguez, OL, Holbrook, J, Lovett, R, Materechera, S, Parsons, M, Raseroka, K, Rodriguez-Lonebear, D, Rowe, R, Sara, R, Walker, JD, Anderson, J and Hudson, M. 2020. The CARE Principles for Indigenous Data Governance. Data Science Journal, 19: 43, pp. 1–12. DOI: <https://doi.org/10.5334/dsj-2020-043>



CONTROL

(Autoridad para controlar)



C1. Reconocimiento de derechos e intereses

- C1.1 Reconoce los derechos e intereses de los PI con respecto a sus datos.
- C1.2 Reconoce el derecho de los PI a la consulta libre, previa e informada.



C2. Datos para gobernanza

- C2.1 Relevante y comprendido bajo la cosmovisión de los PI.
- C2.2 Con disponibilidad para el apoyo de la gobernanza indígena.



C3. Gobernanza de datos

- C3.1 Con derecho a desarrollar protocolos de gobernanza cultural de datos.
- C3.2 Acceso y manejo de datos por parte de los PI.



RESPONSABILIDAD



R1. Para las relaciones positivas

R1.1 Respeta la reciprocidad, confianza y el entendimiento mutuo.

R1.2 Mantiene y respeta la dignidad de los colectivos indígenas.



R2. Para ampliar capacidades

R2.1 Incrementa las competencias para el manejo y uso de datos de los PI.

R2.2 Apoya el desarrollo de competencias e infraestructura digital de los PI.



R3. Para los idiomas y cosmovisiones indígenas

R3.1 Provee recursos para generar datos en las lenguas y bajo la cosmovisión de los PI.



APROVECHAMIENTO

para el beneficio colectivo

O



A1. Para el desarrollo inclusivo y la innovación

A1.1 Establece las bases para la innovación y la agregación de valor para PI.

A1.2 Promueve procesos locales y auto-determinados de desarrollo.



A2. Para mejorar la gobernanza y el compromiso ciudadano

A2.1 Mejora la toma de decisiones colectiva (Planeación-Implementación-Evaluación).

A2.2 Reconoce la participación efectiva y empoderadora en las políticas públicas.



A3. Para obtener resultados equitativos

A3.1 El valor creado se distribuye de manera equitativa con los PI.

A3.2 Contribuye a las aspiraciones de bienestar de los PI.



ÉTICA



E1. Para minimizar daños y maximizar beneficios

- E1.1 No estigmatiza a los PI en términos negativos.
- E1.2 Se alinea con los marcos éticos de los PI.
- E1.3 Se alinea con los derechos presentados en la DRIPS.
- E1.4 Valora los beneficios desde la perspectiva de los PI.
- E1.5 Valora los daños potenciales desde la perspectiva de los PI.



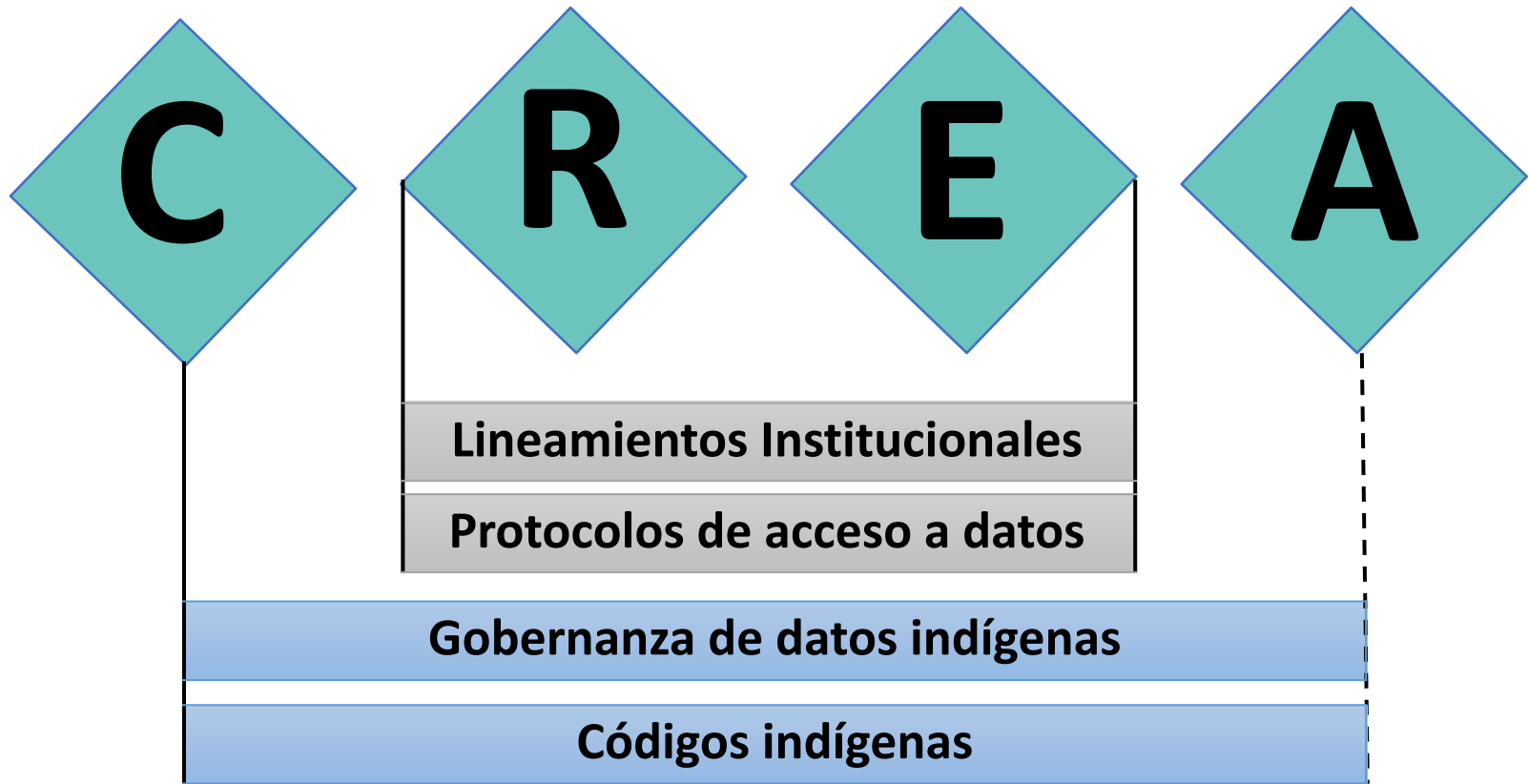
E2. Por la justicia

- E2.1 Atiende desigualdades de poder y acceso/uso de recursos.
- E2.2 Atiende posibles daños a los derechos de los PI.
- E2.3 Los procesos éticos incluyen la representación de los PI.



E3. Para su uso futuro

- E3.1 La gobernanza de datos contribuye al potencial futuro.



SEMINÁRIO INTERNACIONAL DE PRESERVAÇÃO DIGITAL e
ENCONTRO DA REDE BRASILEIRA DE SERVIÇOS DE PRESERVAÇÃO DIGITAL
PROGRAMAS E PROJETOS DE PRESERVAÇÃO DIGITAL
junho de 2023, dias 14, 15 e 16



¡Muchas gracias!

Marisa R. De Giusti

marisa.degiusti@sedici.unlp.edu.ar



Esta obra está bajo una [Licencia Creative Commons](https://creativecommons.org/licenses/by-nc-sa/4.0/)
Atribución-NoComercial-CompartirIgual 4.0 Internacional

