

El Impacto de los Repositorios de Datos en la Investigación Científica

CAPACITACIÓN

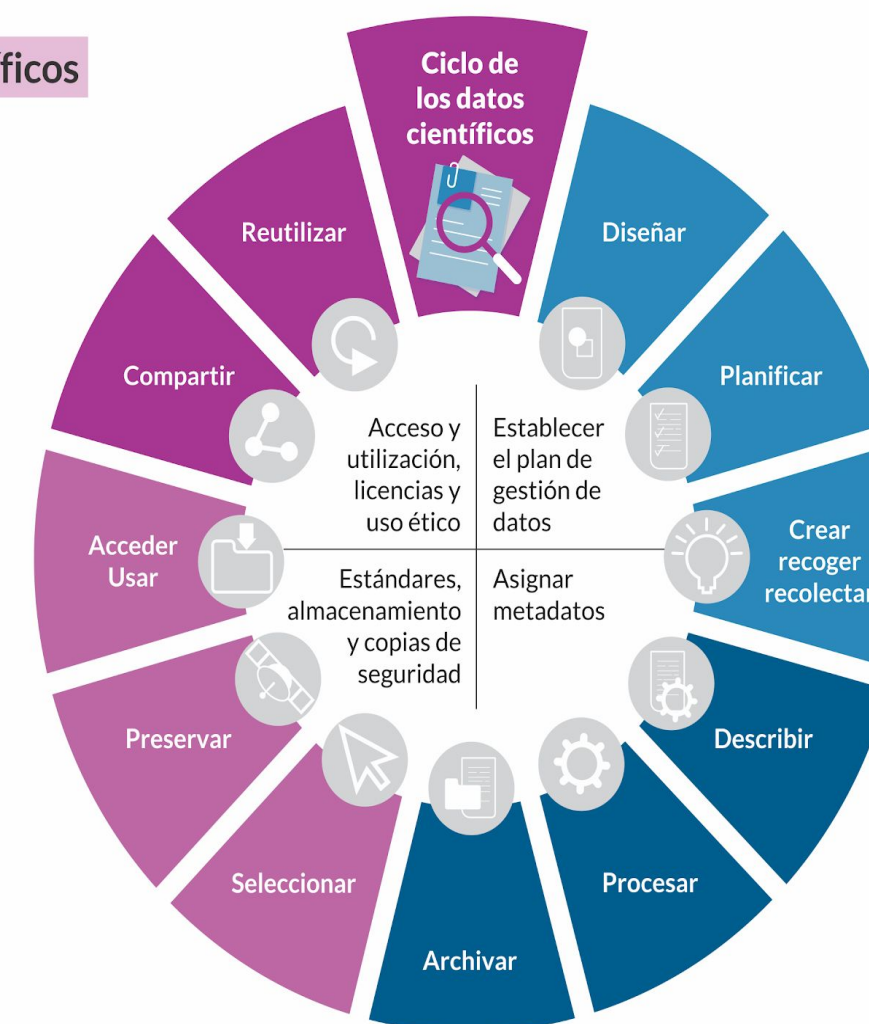
Curación de Datos en el

REPOSITORIO DE DATOS ACADÉMICOS UNR

Plan de Gestión de Datos (PGD)

26 de septiembre de 2023

El ciclo de los datos científicos



Basado en REBIUN



Recomendación de Ciencia Abierta de UNESCO (2021)

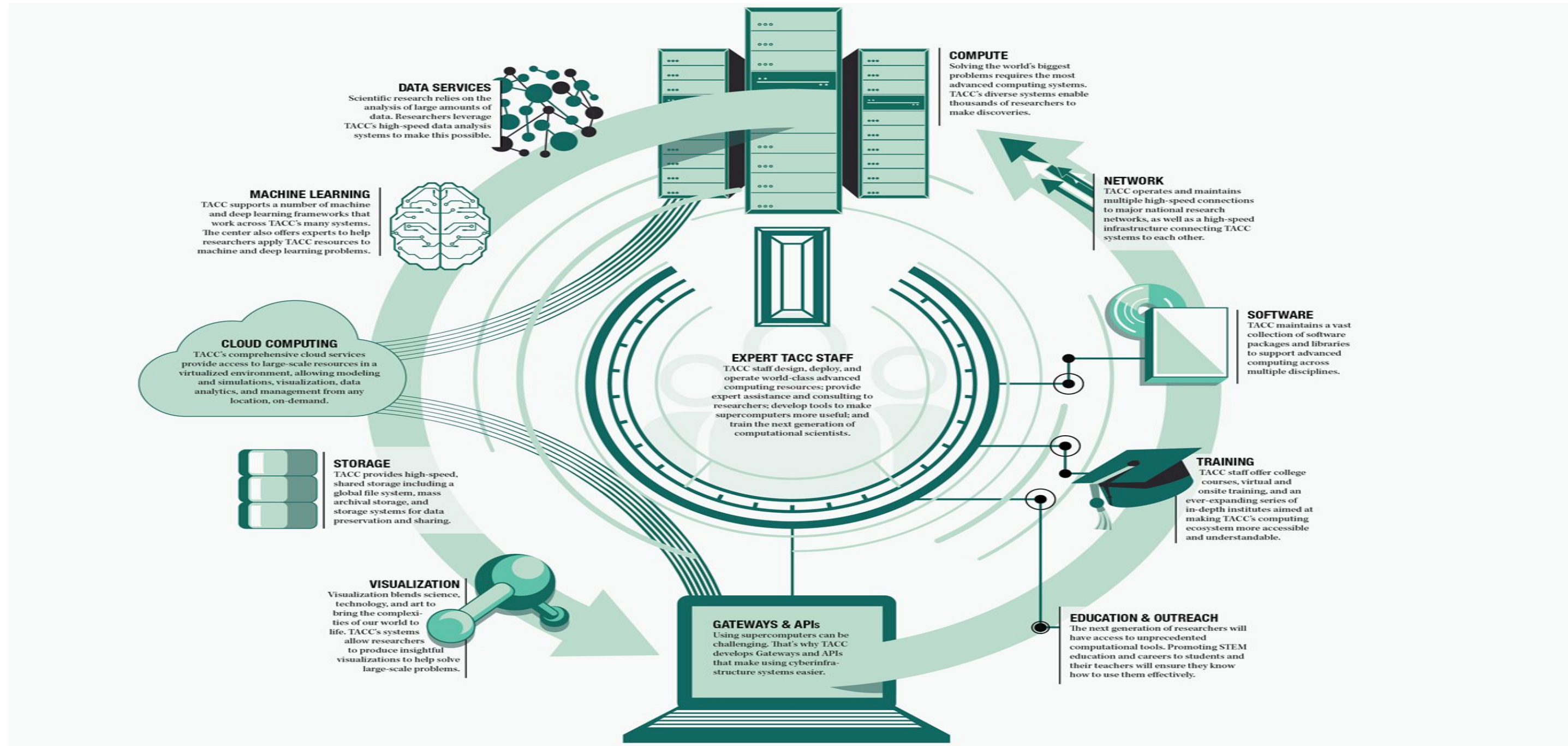
MINCyT (2022). Diagnóstico y lineamientos para una política de ciencia abierta en Argentina

¿Qué son los datos de investigación?



Información recolectada, observada o generada durante el proceso de investigación, que fundamenta las afirmaciones de la investigación y sirve para **validar los resultados**.

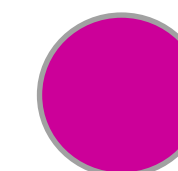
Todo dato sobre el que se basa una investigación y que fundamenta un nuevo conocimiento (**Ley 26.899**)



Ejemplo de infraestructura de investigación TACC, Texas Advanced Computing Center
Universidad de Texas en Austin

Los repositorios de datos abiertos

- ✓ Componente de **la infraestructura para investigación**
- ✓ Las bibliotecas académicas aceptaron el desafío
- ✓ Plataforma para datos: Software y almacenamiento seguro
- ✓ Gobernanza: compromiso institucional y personal especializado
- ✓ Estándares y buenas prácticas
- ✓ Distribución global de datasets y Marketing
- ✓ Preservación indefinida de los datos

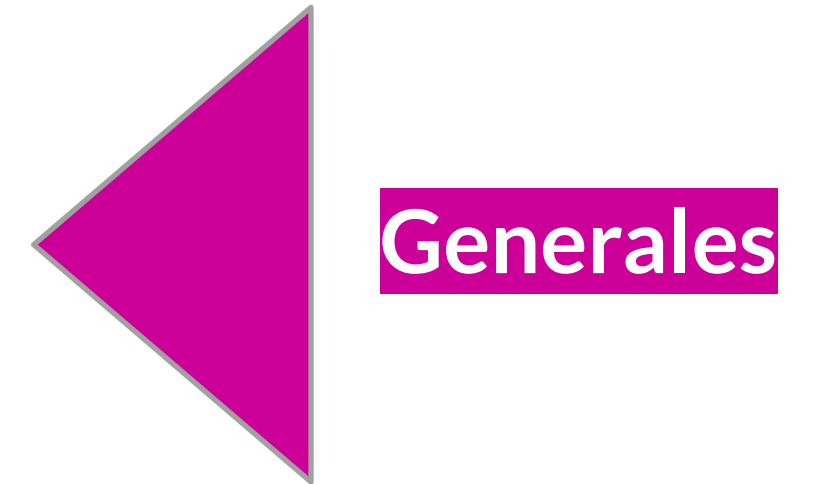


Implementación de repositorios de datos

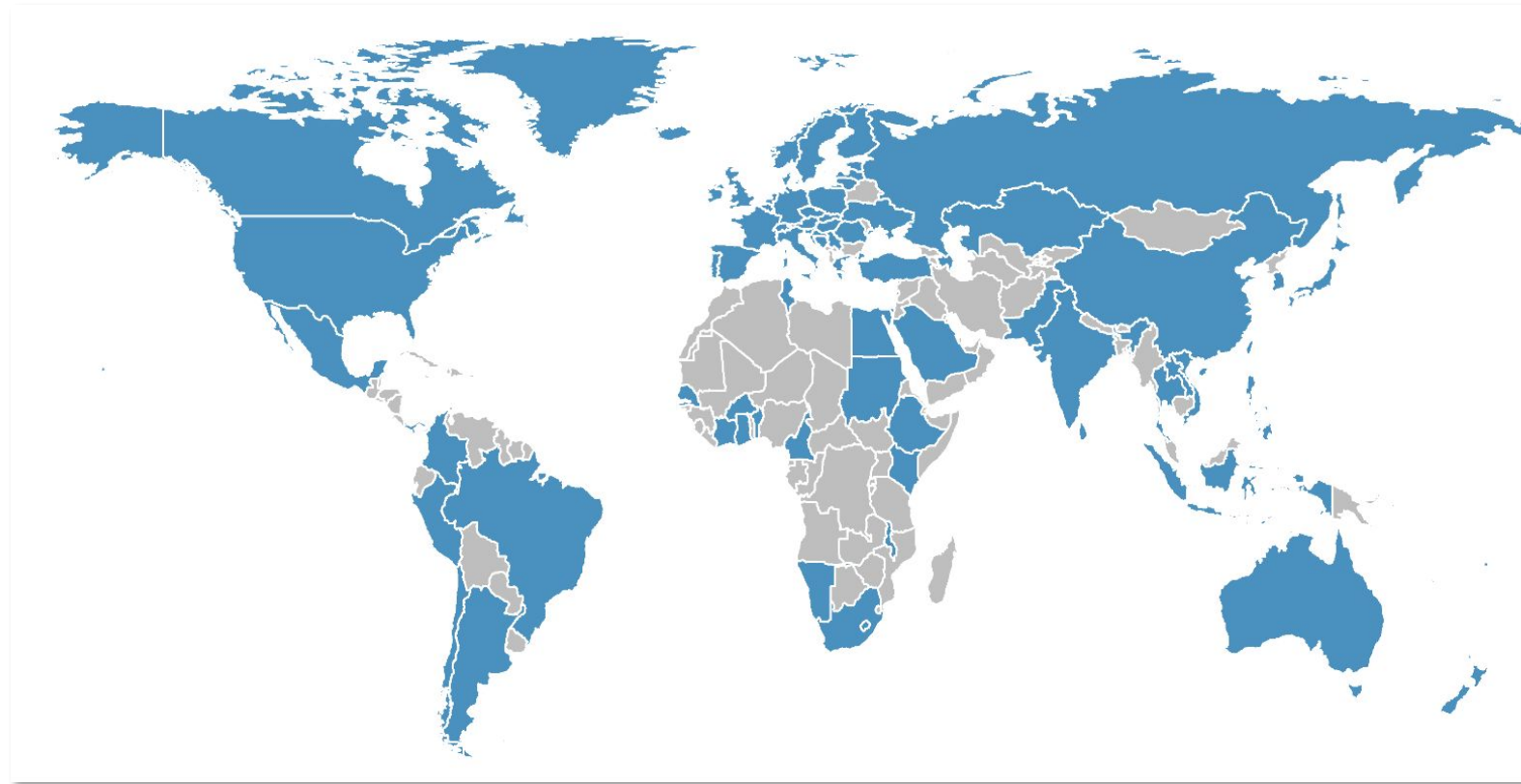
- Bibliotecas académicas
- Repositorios institucionales
- Modelo de auto publicación con guía

Agencias gubernamentales

- Curación especializada
- Colecciones especializadas
- Financiamiento para colecciones de datos



Repositorios de datos en diferentes países y disciplinas



re3data.org Search Browse Suggest Resources Conta

Filter
Reset all

Subjects ▾
Humanities and Social Sciences (1168)
Humanities (352)
Ancient Cultures (75)
Prehistory (8)
Classical Philology (1)
Ancient History (6)
Classical Archaeology (20)
Egyptology and Ancient Near Eastern Studies (7)
History (118)
Medieval History (9)
Early Modern History (7)
Modern and Current History (20)
History of Science (10)
Fine Arts, Music, Theatre and Media Studies (80)
Art History (22)
Musicology (19)
Theatre and Media Studies (7)

Search... Search

Toggle short help

← Previous 1 2 3 4 5 6 7 ... 47 Next → Sort by ▾

Found 1168 result(s)

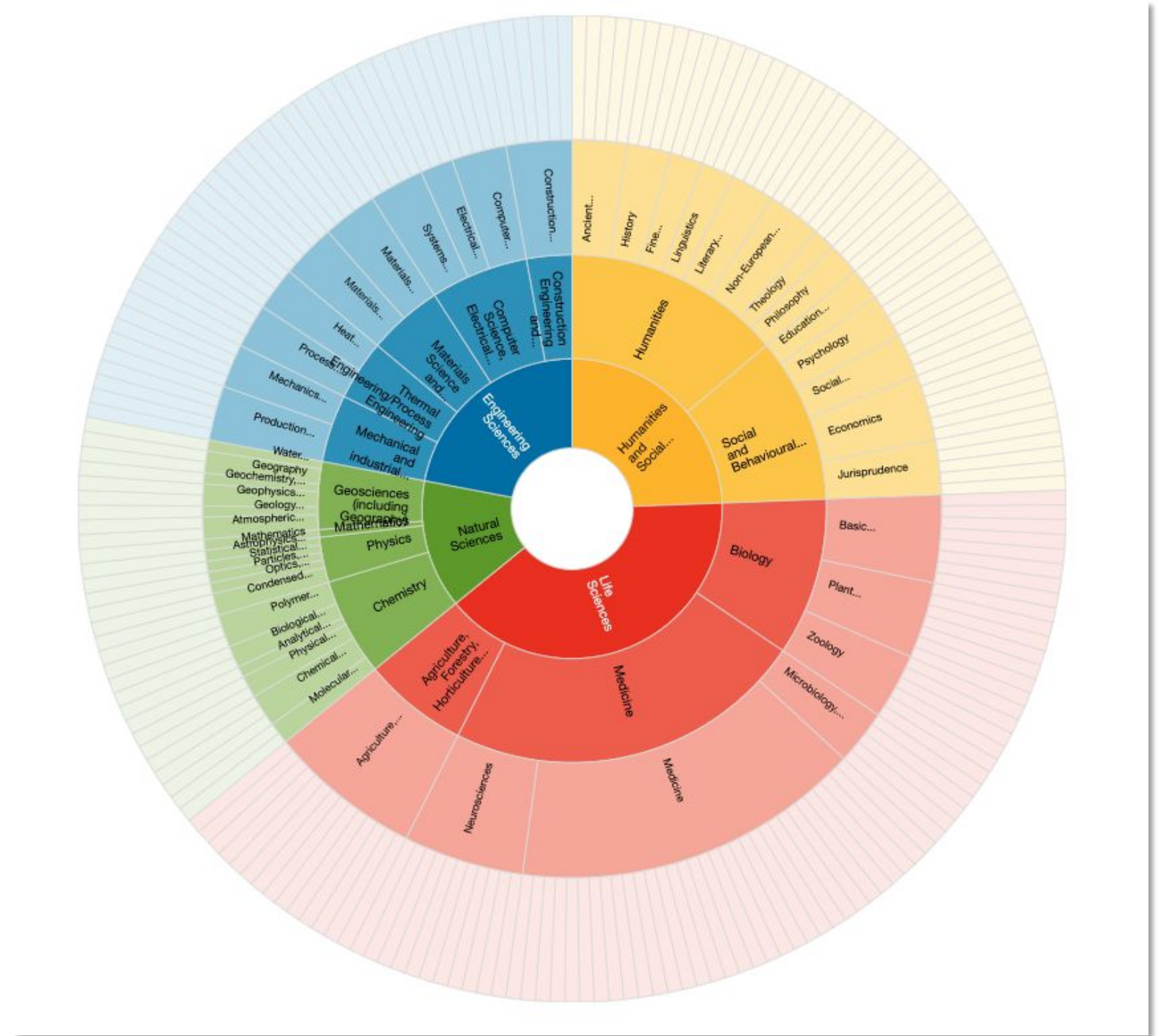
Visual Arts Data Service
VADS

Subject(s) Fine Arts, Music, Theatre and Media Studies Humanities Humanities and Social Sciences

Content type(s) Audiovisual data Images other

Country United Kingdom

VADS is the online resource for visual arts. It has provided services to the academic community for 12 years and has built up a considerable portfolio of visual art collections comprising over 100,000 images that are freely available and copyright cleared for use in learning, teaching and research in the UK. VADS provides: expert guidance and help for digital projects in art education; resource development and hosting for art education; project management and



Fuente: re3data <https://www.re3data.org/>

Repositorios Específicos: OFA Records

Order DNA Tests | Pay My Bill | Log In to OFA Online | Search OFA Health-Tested Dogs

ABOUT | CHIC PROGRAM | DISEASES | BROWSE BY BREED | APPLICATIONS | HEALTH CLINICS | ADVANCED SEARCH

DACHSHUND CHIC

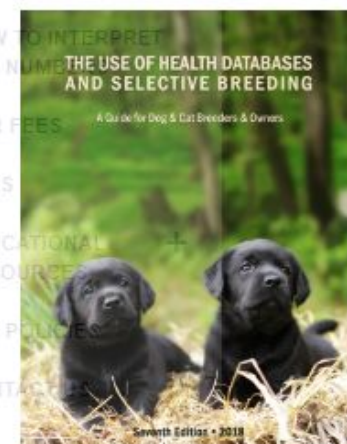
Health Screenings | Testing Statistics | Download Reports | Health Survey | Print Page | Club Web Page

Testing Statistics
Download

Test/Registry	Total Tested	Normal/Clear		Abnormal/Dysplastic		Carrier		Equivocal		Rank	% Trend by Birth Year (% indicates Normal if not further broken down)							
		#	%	#	%	#	%	#	%		<=1990	1991-1995	1996-2000	2001-2005	2006-2010	2011-2015	2016-2020	>=2021
EYES	3,560	3,416	96.0	144	4.0	-	-	-	-	125	.0	.0	50.0	72.3	89.1	95.8	97.1	98.0
PATELLA	2,928	2,808	95.9	120	4.1	-	-	-	-	46	.0	100.0	85.0	95.4	95.6	96.8	95.5	96.9
BASIC CARDIAC	1,154	1,148	99.5	4	.3	-	-	2	.2	9	.0	.0	.0	100.0	88.9	99.9	99.7	
CONGENITAL CARDIAC	952	949	99.7	2	.2	-	-	1	.1	64	.0	100.0	75.0	100.0	100.0	99.7	99.8	.0
HIPS	373	353	94.6	17	4.6	-	-	3	.8	175								
HIPS - EXCELLENT		61	16.4								19.0	23.1	15.8	7.4	9.1	12.5	18.4	25.0
HIPS - GOOD		242	64.9								66.7	61.5	57.9	66.7	68.2	66.1	64.7	62.5
HIPS - FAIR																		
HIPS - ALL DYSPLASTIC																		
HIPS - MILD																		
HIPS - MODERATE																		
HIPS - SEVERE																		
LEGG-CALVE-PERTHES																		
THYROID																		



Ocular Disorders in Purebred Dogs



Monograph: The Use of Health Databases and Selective Breeding



Collecting and Utilizing Phenotypic Data to Minimize Disease: A Breeder's Practical Guide



How the Orthopedic Foundation for Animals (OFA) is tackling inherited disorders in the USA: Using hip and elbow dysplasia as examples

- ✓ <https://ofa.org/chic-programs/browse-by-breed/>
Gestionado por Orthopedic Foundation for Animals (OFA)
- ✓ Tiene como objetivo Recopilar y difundir información sobre enfermedades ortopédicas y genéticas de los animales.
- ✓ Keller, G. G., Dziuk, E., & Bell, J. S. (2011). How the Orthopedic Foundation for Animals (OFA) is tackling inherited disorders in the USA: Using hip and elbow dysplasia as examples. The Veterinary Journal, 189(2), 197-202. <https://doi.org/10.1016/j.tvjl.2011.06.019>

Buscar datasets: DataCite Commons

The screenshot shows the DataCite Commons search results for the query 'dogs'. The search bar at the top contains 'dogs' and a search icon. Below the search bar are navigation tabs for 'Works', 'People', 'Organizations', and 'Repositories', with 'Works' selected. The search results show 5,112 Works. The first result is titled 'Data from: Genetic evidence for multiple events of hybridization between wolves and domestic dogs in the Iberian Peninsula' by Raquel Godinho, Luís Llaneza, Juan Carlos Blanco, Susana Lopes, Francisco Álvares, Emilio J. García, Vicente Palacios, Yolanda Cortés, Javier Talegón & Nuno Ferrand. The result includes a version number (1), a publication year (2011), and a link to the dataset in DRYAD. The abstract text is partially visible, discussing hybridization between wild species and their domestic counterparts. On the left side, there is a filter section for 'Creators & Contributors' and 'Publication Year'. The 'Creators & Contributors' section lists several authors with checkboxes and their respective work counts. The 'Publication Year' section is partially visible at the bottom.

DataCite Commons

dogs

Pages Support Sign In

Works People Organizations Repositories

5,112 Works

Data from: Genetic evidence for multiple events of hybridization between wolves and domestic dogs in the Iberian Peninsula

Raquel Godinho, Luís Llaneza, Juan Carlos Blanco, Susana Lopes, Francisco Álvares, Emilio J. García, Vicente Palacios, Yolanda Cortés, Javier Talegón & Nuno Ferrand

Version 1 of Dataset published 2011 in [DRYAD](#)

Hybridization between wild species and their domestic counterparts may represent a major threat to natural populations. However, high genetic similarity between the hybridizing taxa makes the detection of hybrids a difficult task and may hinder attempts to assess the impact of hybridization in conservation biology. In this work, we used a combination of 42 autosomal microsatellites together with Y-chromosome microsatellite-defined haplotypes and mtDNA sequences to investigate the occurrence and dynamics of wolf-dog hybridization in the Iberian Peninsula. To do this, we applied a variety of Bayesian analyses and a parallel set of simulation studies to evaluate (i) the differences between Iberian wolves and dogs, (ii) the frequency and geographical distribution of hybridization and (iii) the directionality of hybridization. First, we show that Iberian wolves and dogs form two well-differentiated genetic entities, suggesting that introgressive hybridization is not a widespread phenomenon shaping both gene pools. Second, we found evidence for the existence of hybridization that is apparently restricted to more peripheral and recently expanded wolf populations. Third, we describe compelling evidence suggesting that the dynamics of hybridization in wolf populations is mediated by crosses between male dogs and female wolves. More

Creators & Contributors

<input type="checkbox"/>	Mickelson, James R.	21
<input type="checkbox"/>	Drögemüller, Cord	21
<input type="checkbox"/>	Minor, Katie M.	21
<input type="checkbox"/>	Letko, Anna	21
<input type="checkbox"/>	Jagannathan, Vidhya	21
<input type="checkbox"/>	Seefried, Franz R.	19
<input type="checkbox"/>	Eads, David A.	15
<input type="checkbox"/>	Béguin, Jérémy	9
<input type="checkbox"/>	Martinez, Gina Lorena García	8
<input type="checkbox"/>	Pedraza, Luz Natalia	8

Publication Year

2022 701

FEEDBACK

<https://commons.datacite.org/doi.org?query=dogs&resource-type=dataset>

dataverse.unr.edu.ar

Primer repositorio de datos

RDA-UNR
dataverse-info.unr.edu.ar

Única universidad
en Argentina
con Repositorio
Institucional
dedicado a datos
de investigación.

The screenshot shows the homepage of the dataverse.unr.edu.ar website. At the top, there is a navigation menu with links: Inicio, Acerca de, Equipo, Sobre datos abiertos, Historias de datos, and Novedades. A search icon is located in the top right corner. The main content area features a large infographic with the text: "¿Qué es el Repositorio de Datos Académicos UNR?" and "dataverse.unr.edu.ar" with the UNR logo. Below the infographic is the word "Infografía". To the right of the infographic is a vertical list of blue buttons: "Buscar datos", "Depositar datos", "Guía de uso", "Buenas prácticas", and "Políticas". At the bottom right, there is a purple button that says "CALENDARIO - Consultas de curación powered by Calendly".

dataverse-info.unr.edu.ar/

Equipos de trabajo + asesoría externa



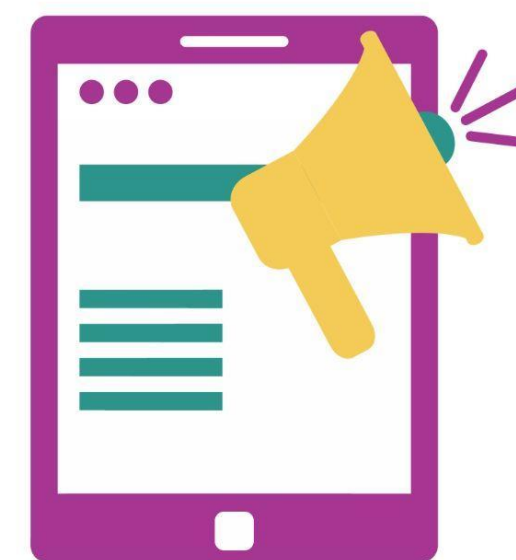
Manual de buenas prácticas de curación de datos



Capacitaciones



Estadísticas



Comunicación

¿POR QUÉ PUBLICAR MIS DATOS?

- ✓ **Importancia para la investigación, la educación y la práctica profesional**
- ✓ **Registro científico (archivo de ciencias)**
- ✓ **Evidencia**
- ✓ **Antecedentes académicos**
- ✓ **Citas de datos y de artículos**
- ✓ **Servicio público**
- ✓ **Economía de recursos**
- ✓ **Formación de nuevos investigadores**



Reduce



Reuse



Recycle

¿QUÉ DATOS VOY A PUBLICAR?

- ✓ Mis datos tienen un **valor de reutilización (validar, integrar) y reproducción (valor evidencial)**
- ✓ **Analizar, con ayuda de un curador, el proceso de investigación y determinar cuáles son los fundamentales para reutilizar y/o reproducir.**
- ✓ Los datos deben ser **funcionalmente utilizables**. ¿Se pueden leer y utilizar los datos? ¿Están disponibles las descripciones y son suficientes para permitir que los futuros usuarios comprendan sus datos?
- ✓ ¿Existen aspectos legales o éticos que impidan la publicación de los datos?
- ✓ De acuerdo a la Ley 26.899 hay que publicar los datos luego de 5 años salvo excepciones.

¿Qué piensan los investigadores?

Principales motivaciones para compartir sus datos (n=6.104)

- ✓ **Citación de sus artículos de investigación (67 %)**
- ✓ **Mayor impacto y visibilidad de sus artículos (61 %)**
- ✓ **Alguna forma de beneficio público (56 %)**
- ✓ **Mandato de la revista/editor (56 %).**
- ✓ **Citación de los datos (54%)**
- ✓ **Mayor transparencia y reutilización (52%)**



Science, Digital; Goodey, Gregory; Hahnel, Mark; Zhou, Yuanchun; Jiang, Lulu; Chandramouliswaran, Ishwar; et al. (2022):
The State of Open Data 2022. Digital Science. Report. <https://doi.org/10.6084/m9.figshare.21276984.v1>

Cada vez más revistas requieren datos abiertos



Las revistas también dan reconocimiento a los artículos que proveen datos abiertos
<https://bmcmicrobiol.biomedcentral.com/open-data-badge-articles>

Data Availability

The following policy applies to all PLOS journals, unless otherwise noted.

Introduction

PLOS journals require authors to make all data necessary to replicate their study's findings publicly available without restriction at the time of publication. When specific legal or ethical restrictions prohibit public sharing of a data set, authors must indicate how others may obtain access to the data.

When submitting a manuscript, authors must provide a Data Availability Statement describing compliance with PLOS' data policy. If the article is accepted for publication, the Data Availability Statement will be published as part of the article.

Acceptable data sharing methods are listed below, accompanied by guidance for authors as to what must be included in their Data Availability Statement and how to follow [best practices in research reporting](#).

PLOS believes that sharing data fosters scientific progress. Data availability allows and facilitates:

- › Validation, replication, reanalysis, new analysis, reinterpretation or inclusion into meta-analyses;
- › Reproducibility of research;
- › Efforts to ensure data are archived, increasing the value of the investment made in funding scientific research;
- › Reduction of the burden on authors in preserving and finding old data, and managing data access requests;
- › Citation and linking of research data and their associated articles, enhancing visibility and ensuring recognition for authors, data producers and curators.

Open data badge



BMC Microbiology participated in a pilot project in which all accepted papers were evaluated for eligibility to receive an Open data badge.

Open data badges were assessed based on the content of the article's data availability statement. All relevant information regarding data which has been used or reused to support the central findings for the study should be described in the data availability statement within the manuscript.

The criteria for receiving a badge were as follows:

1. A data availability statement is included with the manuscript, stating how the data can be accessed.
2. The dataset (or part of the dataset) is deposited in a public repository.
3. A DOI, Accession Number or other appropriate persistent identifier is supplied for the dataset.
4. The dataset provided is relevant to the related paper.

Other references to data in the citations or references, or within the body of the article, were not assessed for an Open data badge.

<https://journals.plos.org/plosone/s/data-availability>

Compartir los datos aumenta las citas



Publicar los datos en repositorios es el único método de compartir datos correlacionado significativamente con el impacto de las citas a los artículos

<https://researchdata.springernature.com/posts/how-sharing-your-data-could-increase-your-citations>

Colavizza G, Hrynaszkiewicz I, Staden I, Whitaker K, McGillivray B (2020) The citation advantage of linking publications to research data. PLoS ONE 15(4): e0230416. <https://doi.org/10.1371/journal.pone.0230416>

THE IMPACT OF DATA REUSE

Citations of DesignSafe (updated on March 31st 2023)

Another way of measuring the impact of DesignSafe is by identifying research papers that cite the use of DesignSafe or the data available at DesignSafe. Table 1 lists DesignSafe citations since 2018 as determined from papers identified via Google Alerts. The first column represents papers that make any reference to DesignSafe through citation of the DesignSafe marker paper (Rathje et al. 2017) or through the acknowledgements. The next column represents papers in which a researcher cites their own data in DesignSafe as a part of the original research project, and the third column represents papers that re-use data available in DesignSafe after the original project is over. Note that a paper may contribute to multiple columns in Table 1. For instance, a data re-use paper may also reference the marker paper, or a paper may cite more than one dataset. There is a meaningful number of total citations that reference the use of DesignSafe and the data published in DesignSafe. While Google Alerts may not capture all of the citations and mentions of DesignSafe datasets that are available in the literature, the positive trend highlights the value of publishing data, the importance of citing data in the references using DOIs, and the types of research being conducted using data published in DesignSafe.

Year	DesignSafe Citation	Primary Data Use	Subsequent Data Reuse	Totals
2023	19	42	37	98
2022	65	107	105	277
2021	42	89	60	191
2020	52	74	61	187
2019	21	25	30	76
2018	26	31	13	70

Funte: Repositorio DesignSafe <https://www.designsafe-ci.org/rw/impact-of-data-reuse/>



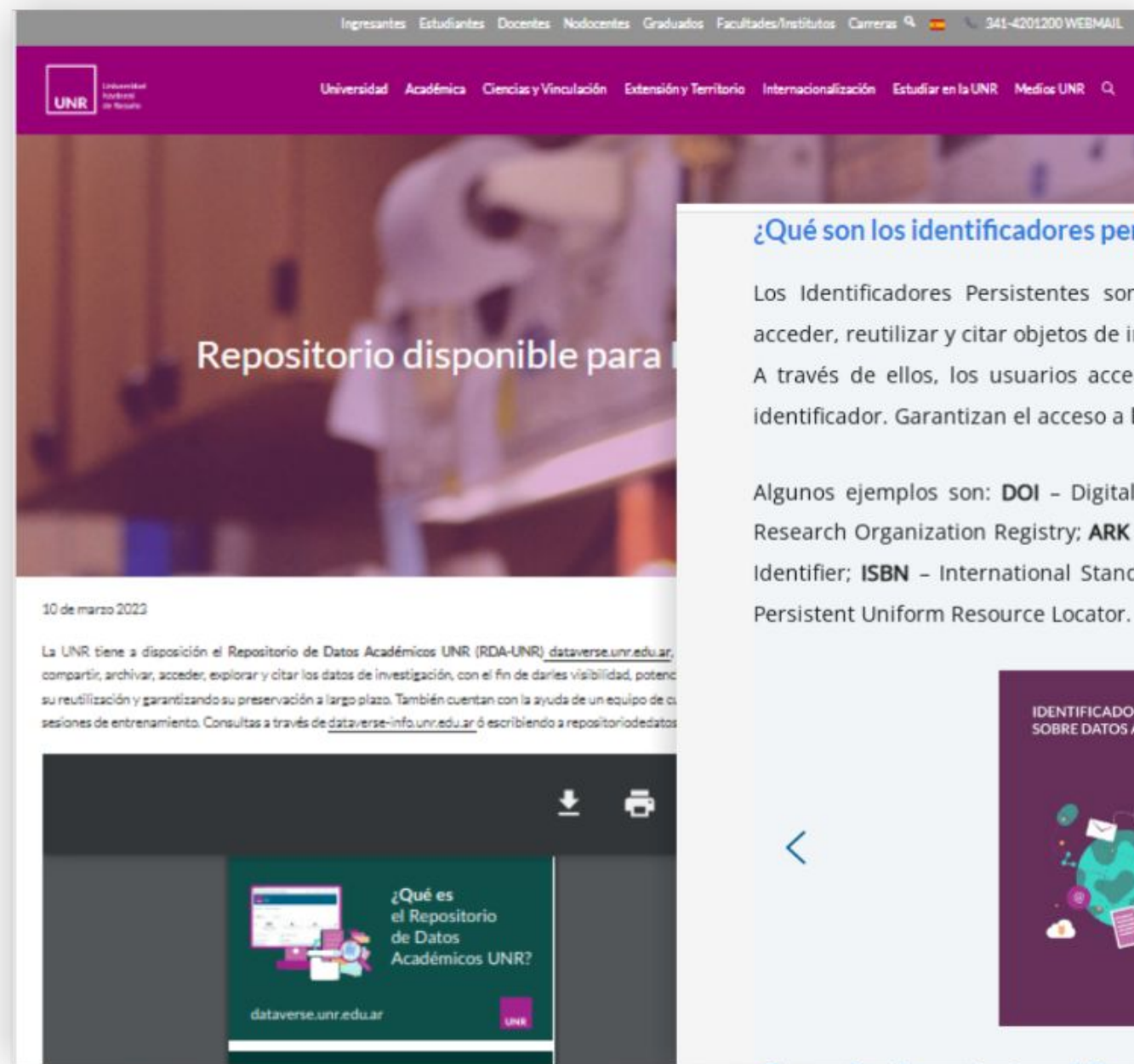
El Repositorio de Datos Académicos RDA-UNR CUMPLE 1 AÑO
dataverse.unr.edu.ar

 **15.438 VISITAS**

 **2.371 ACCESOS**

 **+ 30 PAÍSES**





¿Qué son los identificadores persistentes?

Los Identificadores Persistentes son referencias digitales únicas y permanentes que permiten encontrar, acceder, reutilizar y citar objetos de información digital de cualquier tipo en la web. A través de ellos, los usuarios acceden a una URI (PID) permanente y son redirigidos a la URL asociada al identificador. Garantizan el acceso a los contenidos aunque el sitio web cambie de dirección web.

Algunos ejemplos son: **DOI** – Digital Object Identifier; **ORCID** – Open Researcher and Contributor ID; **ROR** – Research Organization Registry; **ARK** – Archival Resource Key; **Handle** – Handle System; **PMID** – PubMed Unique Identifier; **ISBN** – International Standard Book Number; **ISSN** – International Standard Serial Number; **PURL** – Persistent Uniform Resource Locator.



Web UNR
unr.edu.ar
unr.edu.ar/servicio-disponible-para-investigadores-as/

Web RDA-UNR Info
dataverse-info.edu.ar
dataverse-info.unr.edu.ar/?page_id=129#identificadores

Instagram
[@academicaunr_oficial](https://www.instagram.com/academicaunr_oficial)
[instagram.com/academicaunr_oficial](https://www.instagram.com/academicaunr_oficial)

Correo Electrónico
repositoriodatos@unr.edu.ar

The screenshot shows a Google Dataset Search interface. At the top, the search bar contains 'Gobierno Abierto Argentina'. Below the search bar are several filter buttons: 'Última actualización', 'Formato de descarga', 'Derechos de uso', 'Tema', 'Proveedor', 'Gratis', and 'Conjuntos de datos gu'. The search results section on the left shows three datasets. The first result, highlighted, is 'Respuestas de iniciativas de Gobierno Abierto en Municipi...' from 'dataverse.unr.edu.ar', with a red circular icon containing the letter 'R'. It lists file formats 'pdf +2' and a last update date of 'Jul 27, 2022'. The second result is 'Tercer Plan de Acción Nacional de Gobierno Abierto' from 'datos.gob.ar' and 'data.amerigeoss.org', with a blue circular icon containing the letter 'D' and a last update date of 'Sep 13, 2023'. The main content area on the right displays details for the first result: the title 'Respuestas de iniciativas de Gobierno Abierto en Municipios de Argentina', a bookmark icon, the source 'Ver en: RDA UNR | dataverse.unr.edu.ar', file formats 'text/comma-separated-values(46903), pdf(213418), xlsx(23264)', the unique identifier 'https://doi.org/10.57715/UNR/UVRYFZ', the update date 'Jul 27, 2022', the provider 'Conjunto de datos proporcionado por RDA UNR', and the license 'Licencia Attribution 4.0 (CC BY 4.0)'. A note at the bottom states 'Se ha obtenido la información de la licencia automáticamente'.

Google

Gobierno Abierto Argentina

Última actualización Formato de descarga Derechos de uso Tema Proveedor Gratis Conjuntos de datos gu

Se han encontrado 3 conjuntos de datos

R Respuestas de iniciativas de Gobierno Abierto en Municipi...
dataverse.unr.edu.ar
pdf +2
Última actualización: Jul 27, 2022

D Tercer Plan de Acción Nacional de Gobierno Abierto
datos.gob.ar
data.amerigeoss.org
xlsx
Última actualización: Sep 13, 2023

Respuestas de iniciativas de Gobierno Abierto en Municipios de Argentina

Ver en: [RDA UNR | dataverse.unr.edu.ar](https://dataverse.unr.edu.ar)

text/comma-separated-values(46903), pdf(213418), xlsx(23264)

Identificador único
<https://doi.org/10.57715/UNR/UVRYFZ>

Fecha de actualización del conjunto de datos
Jul 27, 2022

Conjunto de datos proporcionado por
RDA UNR

Licencia
[Attribution 4.0 \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)
Se ha obtenido la información de la licencia automáticamente

Plan de Gestión de Datos (PGD)



- ✓ Es un documento formal, elaborado por investigadores, que se presenta con el proyecto de investigación.
- ✓ Describe los aspectos de la gestión de datos, es decir, qué datos se van a generar y qué se hará con los datos durante y después del proyecto de investigación.
- ✓ Objetivo: asegurar que estos datos puedan ser publicados para su reutilización en un plazo no mayor a cinco (5) años a partir del momento de su recolección.
- ✓ En el sitio se puede generar el PGD.

<https://dataverse-info.unr.edu.ar/pgd/>

ACTIVIDAD

- ✓ ¿Conocías el RDA-UNR? <https://www.menti.com/alywpzsef3mu>
- ✓ ¿Qué beneficios tiene publicar los datos? Código 3741 9423
- ✓ Regístrate en <https://dataverse.unr.edu.ar/>



RDA-UNR | REPOSITORIO DE DATOS ACADÉMICOS

<https://dataverse-info.unr.edu.ar/>

<https://dataverse.unr.edu.ar/>

<http://doi.org/10.17616/R31NJN9K>

 [@academicaunr_oficial](#)



Íconos de <https://www.freepik.com/>

UNR



¡Muchas gracias!

repositoriodatos@unr.edu.ar