

# FAIR Data

¿Cómo hacer que los datos sean FAIR?: buenas prácticas para datos (abiertos) de investigación

Paula Martinez Lavanchy – Research data officer –TU Delft | 4TU.ResearchData

# Principios FAIR - Origen



[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](#)

- Findable
- Accessible
- Interoperable
- Reusable

# Principios FAIR

- Encontrable – otros (humanos y computadores) pueden descubrir los datos
- Accesible – los datos están disponible para otros usuarios (y computadores)
- Interoperable – los datos pueden ser integrados fácilmente con otros o en distintos procesos de análisis
- Re-utilizable – los datos pueden ser usados en nuevas investigaciones

## Box 2 | The FAIR Guiding Principles

### To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

### To be Accessible:

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
  - A1.1 the protocol is open, free, and universally implementable
  - A1.2 the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

### To be Interoperable:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

### To be Reusable:

- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
  - R1.1. (meta)data are released with a clear and accessible data usage license
  - R1.2. (meta)data are associated with detailed provenance
  - R1.3. (meta)data meet domain-relevant community standards

# Principios FAIR – NO son

- Un estándar
- Todo o nada
- Datos Abiertos
- Aplicados utilizando una única tecnología o herramienta

## Box 2 | The FAIR Guiding Principles

### To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

### To be Accessible:

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
  - A1.1 the protocol is open, free, and universally implementable
  - A1.2 the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

### To be Interoperable:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

### To be Reusable:

- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
  - R1.1. (meta)data are released with a clear and accessible data usage license
  - R1.2. (meta)data are associated with detailed provenance
  - R1.3. (meta)data meet domain-relevant community standards

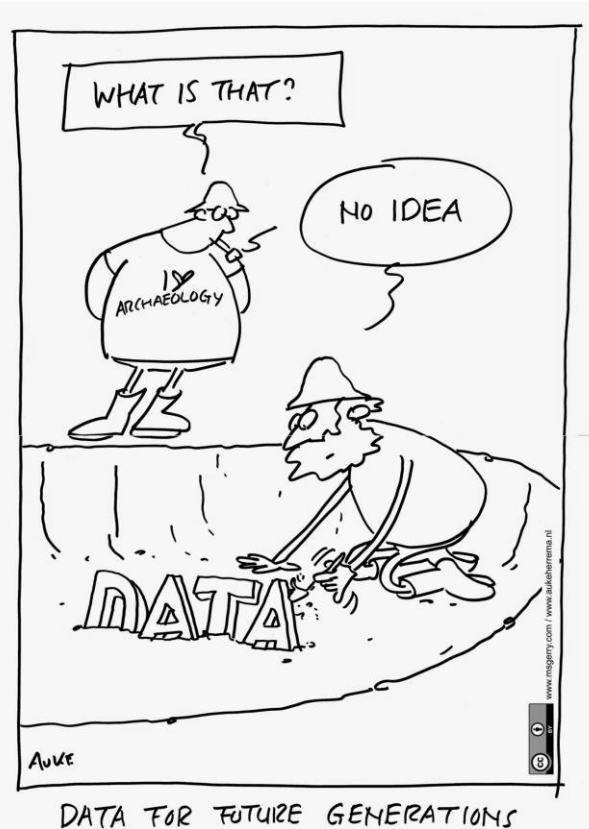
# Desgranando los principios FAIR

Elementos relevantes:

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



# Documentación y Metadatos



- Encontrable
- Re-utilizable
- Interoperable\*

**Bonus - Reproducible**



# Documentación y Metadatos

- Descripción del proceso – p. ej. recolección de datos
- Descripción de la procedencia e historial de los datos – p.ej. reporte del uso y modificaciones de un set de datos, registro de distintas versiones
- Descripción de los datos mismos o Metadatos – p.ej. tamaño, formato, software para leer los datos, autor, etc.

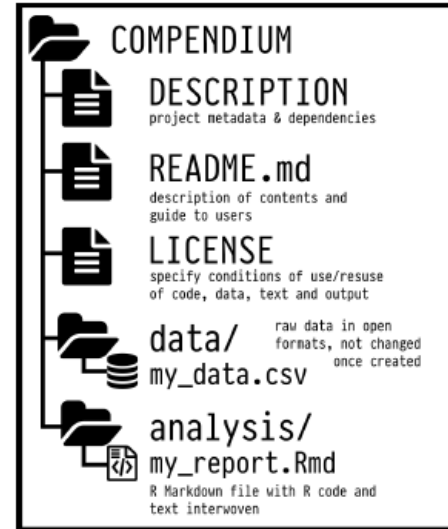
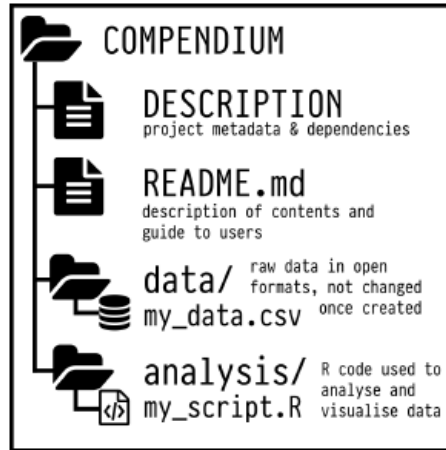
# Documentación y Metadatos

- Estructura del proyecto y convenciones
- Descripción del proceso – p. ej. recolección de datos
- Descripción de los datos mismos o Metadatos – p.ej. tamaño, formato, software para leer los datos
- Descripción de la procedencia de los datos – p.ej. reporte del uso y modificaciones de un set de datos, registro de distintas versiones

# Estructura de proyecto

- Determinar una estructura de archivos desde el principio de un proyecto
- Datos, métodos y resultados deben estar claramente separados
- El ambiente computacional del software para analizar los datos debe ser especificado
- Preparar un archivo 'ReadMe' para el Proyecto y si es posible cada una de las carpetas de archivos
- **Ser consistente**

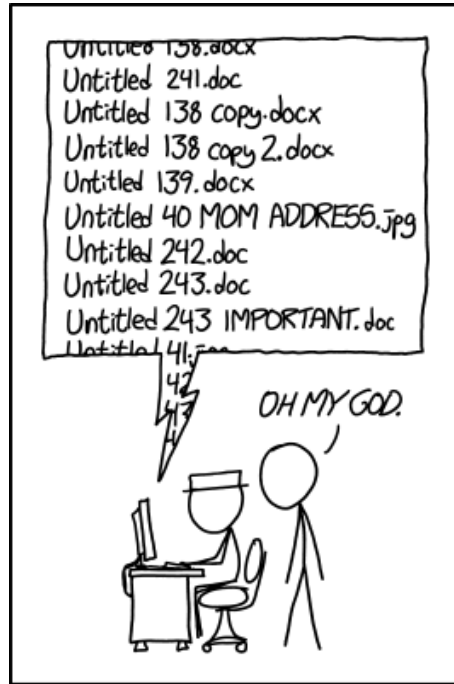
# Estructura de proyecto – compendio de investigación



Marwick B, Boettiger C, Mullen L. 2018. Packaging data analytical work reproducibly using R (and friends) PeerJ Preprints 6:e3192v2 <https://doi.org/10.7287/peerj.preprints.3192v2>

# Convenciones – nombre y versiones de archivos

(Image source: <http://sxnkd.com/1459/>)



PRO TIP: NEVER LOOK IN SOMEONE ELSE'S DOCUMENTS FOLDER.



# Convenciones – nombres de archivos

2020-08-26\_Reactor1\_oxígeno\_día\_1.csv

2020-08-26\_Reactor1\_oxígeno\_día\_1\_v1.0.csv

2020-08-26\_Reactor1\_oxígeno\_día\_1\_v1.1.csv

# Documentación - ReadMe

- Se puede utilizar para un Proyecto, para un set de datos o para un archivo específico
- Debe tener un nombre de archivo que lo asocie fácilmente con los archivos datos que describe
- Escribirlo en un formato abierto p.ej. .txt
- Estructurarlo con secciones definidas (y espaciadas)
- Utilizar 'templates' cuando sea posible
- Utilizar términos o vocabularios estándares disciplinarios cuando sea posible

	A	B	C	D	E	F	G	H	I	J
1	sample	timepoint	cont	pore_size	MHC	Naph	Phen	Anth	Fluor	Pyr
2	T1_0.1-A	1	0	0.1	0	0	0	0	0	0
3	T1_0.1-B	1	0	0.1	0	0	0	0	0	0
4	T1_0.1-C	1	0	0.1	0	0	0	0	0	0
5	T1_0.1+A	1	1	0.1	3600	0.21	6.1	4.5	6.4	5.8
6	T1_0.1+B	1	1	0.1	3600	0.16	6.1	4.5	6.2	5.7
7	T1_0.1+C	1	1	0.1	3100	0.14	5.3	3.9	5.2	4.8
8	T1_0.1-A	1	0	0.1	0	0	0	0	0	0

ReadMe\_full\_set\_data\_T1+T2+T3\_example - Notepad

File Edit Format View Help

Timepoint: 1: six weeks; 2: three months; 3: six months

cont: 0: non-contaminated, 1:contaminated

pore\_size: in micrometers (filter)

All chemical data are in mg/ Kg soil

MHC: Mineral Hydrocarbons

Naph: Naphtalin

Phen: Phenanthren

Anth: Anthracen

Fluor: Fluoranthen

Pyr: Pyren

Bact: 16S rRNA copies/ gr soil; primers: XXXX (Ref)

Nomenclature Sample: T1\_0.1-A; T1: time point 1; 0.1: 0.1 um filter; (-): non-contaminated A:



# Documentación - ReadMe

Guide to writing "readme" style metadata - Cornell University:  
<https://data.research.cornell.edu/content/readme#bestpractices>

4TU.ResearchData - Guidelines:  
[https://researchdata.4tu.nl/fileadmin/user\\_upload/Documenten/Guidelines\\_for\\_creating\\_a\\_README\\_file.pdf](https://researchdata.4tu.nl/fileadmin/user_upload/Documenten/Guidelines_for_creating_a_README_file.pdf)

# Documentación - Herramientas

- Cuadernos (Electrónicos) de Laboratorio:
  - Ten Simple Rules for a Computational Biologist's Laboratory Notebook:  
<https://doi.org/10.1371/journal.pcbi.1004385>
  - Electronic Lab Notebooks - for prospective users:  
<https://www.gurdon.cam.ac.uk/institute-life/computing/elnguidance>
- GitHub/GitLab:
  - 1.1: Introduction - Git and GitHub for Poets:  
<https://www.youtube.com/watch?v=BCQHnlnPusY> (14 min video)
  - Ten Simple Rules for Taking Advantage of Git and GitHub:  
<https://doi.org/10.1371/journal.pcbi.1004947>

# Documentación - Herramientas

- RMarkdown:
  - R Markdown from Rstudio: <https://rmarkdown.rstudio.com/lesson-1.html>
  - R Markdown with RStudio - Roger Peng - <https://www.youtube.com/watch?v=DNS7i2m4sB0>
- Jupyter Notebooks:
  - Why Jupyter is data scientists' computational notebook of choice: <https://www.nature.com/articles/d41586-018-07196-1>
  - Ejemplos de Jupyter Notebooks en distintas disciplinas: <https://github.com/jupyter/jupyter/wiki/A-gallery-of-interesting-Jupyter-Notebooks>

# Metadatos

- Son relevantes para acceder, entender y procesar datos (por humanos y computadores/procesadores)
- El quién, qué, por qué, dónde y cuándo de los datos
- Pueden estar incluidos o ser externos a un archivo de datos



pexels-pixabay-209389 Properties

General Security Details Previous Versions

Property	Value
<b>Image</b>	
Image ID	
Dimensions	1280 x 853
Width	1280 pixels
Height	853 pixels
Horizontal resolution	72 dpi
Vertical resolution	72 dpi
Bit depth	24
Compression	
Resolution unit	
Color representation	
Compressed bits/pixel	
<b>Camera</b>	
Camera maker	
Camera model	
F-stop	
Exposure time	
ISO speed	

[Remove Properties and Personal Information](#)

OK Cancel Apply

# Estándares de Metadatos

- Proveen una estructura que describen 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 del los datos

Fuente: <https://sos598-rdm.github.io/rdm-lecture-metadata/#1>

## Sample GenBank Record

This page presents an annotated sample GenBank record (accession number [U49845](#)) in its *GenBank Flat File* format. You can see the corresponding [live record for U49845](#), and see [examples of other records](#) that show a range of biological features.

```
LOCUS          SCU49845      5028 bp      DNA          PLN          21-JUN-1999
DEFINITION    Saccharomyces cerevisiae TCP1-beta gene, partial cds, and Axl2p
                (AXL2) and Rev7p (REV7) genes, complete cds.
ACCESSION    U49845
VERSION      U49845.1  GI:1293613
KEYWORDS     .
SOURCE       Saccharomyces cerevisiae (baker's yeast)
ORGANISM     Saccharomyces cerevisiae
                Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
                Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE    1 (bases 1 to 5028)
AUTHORS      Torpey,L.E., Gibbs,P.E., Nelson,J. and Lawrence,C.W.
TITLE        Cloning and sequence of REV7, a gene whose function is required for
                DNA damage-induced mutagenesis in Saccharomyces cerevisiae
JOURNAL      Yeast 10 (11), 1503-1509 (1994)
PUBMED       7871890
REFERENCE    2 (bases 1 to 5028)
AUTHORS      Roemer,T., Madden,K., Chang,J. and Snyder,M.
TITLE        Selection of axial growth sites in yeast requires Axl2p, a novel
                plasma membrane glycoprotein
JOURNAL      Genes Dev. 10 (7), 777-793 (1996)
PUBMED       8846915
REFERENCE    3 (bases 1 to 5028)
AUTHORS      Roemer,T.
TITLE        Direct Submission
JOURNAL      Submitted (22-FEB-1996) Terry Roemer, Biology, Yale University, New
                Haven, CT, USA
FEATURES     Location/Qualifiers
```



2019 September 24: Hubble Catalog of Variables released. Close x  
 Find out what's new in the [Hubble Source Catalog](#).

# Hubble Legacy Archive

Orion    [advanced search](#)  
 Examples: M101, 14 03 12.6 +54 20 56.7 r=0.2d, more...  
 Requires Firefox, Safari, IE, or compatible browser

Retrieve	RA	DEC	Level	Target	Detector	Aperture	Spectral_Elt	NExposures	ExpTime	StartTime	Dataset	PropID
<a href="#">FITS</a>	05:34:56.48	-05:31:36.1	5	ORION	ACS/WFC		B		0	1900-01-01 00:00:00	hlsp_orion_hst_acs_colorimage_b_v1_drz	
<a href="#">FITS</a>	05:34:56.48	-05:31:36.1	5	ORION	ACS/WFC		G		0	1900-01-01 00:00:00	hlsp_orion_hst_acs_colorimage_g_v1_drz	
<a href="#">FITS</a>	05:34:56.48	-05:31:36.1	5	ORION	ACS/WFC		R		0	1900-01-01 00:00:00	hlsp_orion_hst_acs_colorimage_r_v1_drz	
<a href="#">FITS</a>	05:34:56.48	-05:31:36.1	4	ORION	ACS/WFC		B/G/R		0	1900-01-01 00:00:00	hlsp_orion_hst_acs_colorimage_b_g_r_v1_drz	
<a href="#">FITS</a>	05:35:15.6	-05:24:18.0	5	ORIMOS-13	ACS/WFC	WFCENTER	F435W	14	2940	2005-04-05 04:35:48	hlsp_orion_hst_acs_strip0l_f435w_v1_drz	10246
<a href="#">FITS</a>	05:35:15.6	-05:24:18.0	5	ORIMOS-13	ACS/WFC	WFCENTER	F555W	14	2695	2005-04-05 05:04:13	hlsp_orion_hst_acs_strip0l_f555w_v1_drz	10246
<a href="#">FITS</a>	05:35:15.6	-05:24:18.0	5	ORIMOS-13	ACS/WFC	WFCENTER	F658N	14	2380	2005-04-05 04:27:01	hlsp_orion_hst_acs_strip0l_f658n_v1_drz	10246
<a href="#">FITS</a>	05:35:15.6	-05:24:18.0	5	ORIMOS-13	ACS/WFC	WFCENTER	F775W	14	2695	2005-04-05 04:55:14	hlsp_orion_hst_acs_strip0l_f775w_v1_drz	10246
<a href="#">FITS</a>	05:35:15.6	-05:24:18.0	5	ORIMOS-13	ACS/WFC	WFCENTER	F850LP	14	2695	2005-04-05 04:46:08	hlsp_orion_hst_acs_strip0l_f850lp_v1_drz	10246
<a href="#">FITS</a>	05:35:15.6	-05:24:18.0	4	ORIMOS-13	ACS/WFC	WFCENTER	F850LP/F658N/ F435W	70	13405	2005-04-05 04:27:01	hlsp_orion_hst_acs_strip0l_f850lp_f658n_f435w_v1_drz	10246
<a href="#">FITS</a>	05:35:16.8	-05:27:00.0	5	ORIMOS-26	ACS/WFC	WFCENTER	F435W	12	2520	2005-04-09 04:20:48	hlsp_orion_hst_acs_strip1l_f435w_v1_drz	10246
<a href="#">FITS</a>	05:35:16.8	-05:27:00.0	5	ORIMOS-26	ACS/WFC	WFCENTER	F555W	14	2695	2005-04-09 01:37:16	hlsp_orion_hst_acs_strip1l_f555w_v1_drz	10246
<a href="#">FITS</a>	05:35:16.8	-05:27:00.0	5	ORIMOS-26	ACS/WFC	WFCENTER	F658N	12	2040	2005-04-09 03:23:13	hlsp_orion_hst_acs_strip1l_f658n_v1_drz	10246
<a href="#">FITS</a>	05:35:16.8	-05:27:00.0	5	ORIMOS-26	ACS/WFC	WFCENTER	F775W	14	2695	2005-04-09 01:28:17	hlsp_orion_hst_acs_strip1l_f775w_v1_drz	10246
<a href="#">FITS</a>	05:35:16.8	-05:27:00.0	5	ORIMOS-26	ACS/WFC	WFCENTER	F850LP	14	2695	2005-04-09 01:19:11	hlsp_orion_hst_acs_strip1l_f850lp_v1_drz	10246
<a href="#">FITS</a>	05:35:16.8	-05:27:00.0	4	ORIMOS-26	ACS/WFC	WFCENTER	F850LP/F658N/ F435W	66	12645	2005-04-09 01:19:11	hlsp_orion_hst_acs_strip1l_f850lp_f658n_f435w_v1_drz	10246
<a href="#">FITS</a>	05:34:37.2	-05:25:30.0	5	ORIMOS-19	ACS/WFC	WFCENTER	F435W	12	2520	2005-04-04 07:45:05	hlsp_orion_hst_acs_strip1r_f435w_v1_drz	10246



## HOW STANDARDS PROLIFERATE: (SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC.)



<https://xkcd.com/927/>

# Estándares de Metadatos – Dónde buscar?

The screenshot displays the FAIRsharing.org website. At the top left is the logo for FAIRsharing.org, with the tagline "standards, databases, policies". To the right of the logo is a search bar containing the text "Search all of FAIRsharing". Further right are several navigation buttons: "Standards", "Databases", "Policies", "Collections", "Add/Claim Content", "Stats", and "PeteMcQ".

Below the navigation bar are three main content cards:

- Find Recommendations:** Features a thumbs-up icon and the text "Standards and/or databases recommended by journal or funder data policies."
- Discover Collections:** Features an icon of three circles with dots inside and the text "Standards and/or databases grouped by domain, species or organization."
- Learn Educational:** Features a graduation cap icon and the text "About standards, their use in databases and policies, and how we can help you."

At the bottom of the screenshot is a dark navigation bar. On the left is a search bar with the text "Search all of FAIRsharing" and a "Search" button. Below the search bar are four checked checkboxes: "Standards", "Databases", "Policies", and "Collections/Recommendations". To the right of the search bar are two sections: "Advanced Search" with a magnifying glass icon and the text "Fine grained control over your search.", and "Search Wizard" with a speech bubble icon and the text "Let us guide you to your results."

# Metadatos - Herramientas

- 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 embebido (mínimo o extendido)
  - 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/dfs/>

# Acceso a los Datos

- Datos FAIR no 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 (Non-Disclosure Agreement - NDA)

# 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/publicar:
  - Proveer acceso a los metadatos
  - Indicar en los metadatos cómo se puede acceder a los datos

Cuando los datos pueden ser de acceso abierto...

# Identificadores Persistentes

- Referencia duradera a un objeto digital
- Hace el objeto Encontrable y Accesible
- Hace el objeto citable de una manera precisa



# Identificadores Persistentes

- Ejemplos:



Libros - International Standard Book Number (ISBN)



Para Investigadores - ORCID



Publicaciones y Datos - Digital Object Identifiers (DOIs)

# Identificadores Persistentes - DOI

PubMed.gov

m chloride concentrations in North American and European freshw

[Advanced](#) [Create alert](#) [Create RSS](#)

Found 1 result for *Long-term chloride concentrations in North Am...*

› [Sci Data. 2017 Aug 8;4:170101. doi: 10.1038/sdata.2017.101.](#)

## Long-term chloride concentrations in North American and European freshwater lakes

Hilary A Dugan<sup>1 2</sup>, Jamie C Summers<sup>3</sup>, Nicholas K Skaff<sup>4</sup>, Flora E Krivak-Tetley<sup>5</sup>, Jonathan P Doubek<sup>6</sup>, Samantha M Burke<sup>7</sup>, Sarah L Bartlett<sup>8</sup>, Lauri Arvola<sup>9</sup>, Hamdi Jarjanazi<sup>10</sup>, János Korponai<sup>11 12</sup>, Andreas Kleeberg<sup>13</sup>, Ghislaine Monet<sup>14</sup>, Don Monteith<sup>15</sup>, Karen Moore<sup>16</sup>, Michela Rogora<sup>17</sup>, Paul C Hanson<sup>1</sup>, Kathleen C Weathers<sup>2</sup>

Affiliations [+ expand](#)

PMID: 28786983 PMID: PMC5548073 DOI: 10.1038/sdata.2017.101

[Free PMC article](#)

Explore our content Journal information 

nature &gt; scientific data &gt; data descriptors &gt; article

Open Access | Published: 08 August 2017

# Long-term chloride concentrations in North American and European freshwater lakes

Hilary A. Dugan , Jamie C. Summers, Nicholas K. Skaff, Flora E. Krivak-Tetley, Jonathan P. Doubek, Samantha M. Burke, Sarah L. Bartlett, Lauri Arvola, Hamdi Jarjanazi, János Korponai, Andreas Kleeberg, Ghislaine Monet, Don Monteith, Karen Moore, Michela Rogora, Paul C. Hanson & Kathleen C. Weathers

## Data Citations

- 1 Dugan, H. *Environmental Data Initiative* <https://doi.org/10.6073/pasta/b473b048846875b721d139416b8fd882> (2017)

Download PDF



## Associated Content

*Environmental Science and Pollution Research* | Article  
Recent trends in chloride and sodium concentrations in the deep subalpine lakes (Northern Italy)

Michela Rogora, Rosario Mosello[...] & Fabio Buzzi



enter search terms



▶ ADVANCED SEARCH

## Data Package Summary [View Full Metadata](#)

*Title:* Global Lake Ecological Observatory Network: Long term chloride concentration from 529 lakes and reservoirs around North America and Europe: 1940-2016

*Creators:* Dugan, Hilary; University of Wisconsin-Madison  
Bartlett, Sarah; University of Wisconsin-Milwaukee  
Burke, Samantha; University of Waterloo  
Doubek, Jonathan; Virginia Tech  
Krivak-Tetley, Flora; Dartmouth Co...  
[Show more >](#)

*Publication Date:* 2017-05-10

# 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**.
- **Citación** apropiada es siempre **buena práctica**



The person who associated a work with this deed has **dedicated** the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law.

You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission.



**Attribution** — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

**No additional restrictions** — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.



# Repositorios de Datos

- Depositar datos en un repositorio ayuda a hacer los datos Encontrables, Accesibles y Re-utilizables\*
- 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
- Permite relacionar los datos con otros datos, publicaciones y referencias relevantes mejorando interoperabilidad

# Recordar!

- Si los datos no se pueden compartir abiertamente/publicar:
  - Proveer **acceso a los metadatos**
  - Indicar en los metadatos cómo se puede acceder a los datos



# Repositorio de Datos - Selección

Un repositorio de datos debería:

- Ser reconocido en la comunidad de investigación
- Tener Condiciones y Términos de uso claros
- Garantizar sustentabilidad
- Utilizar estándares de metadatos usados comúnmente
- Proveer un identificador único y persistente a los (meta)datos
- Ofrecer licencias estándares para datos

# Repositorio de Datos - Selección

Podría:

- Ofrecer un servicio de revisión de los datos
- Permitir períodos embargo para set de datos o acceso restringido a los datos (acceso abierto a metadatos)
- Proveer estadísticas (vistas, descargas, citación)

Cite

Download (1.07 MB)

29

views

7

downloads

0

citations

## Air Cargo Transport Network (ACTN) Dataset

Dataset posted on 23.01.2020, 01:00 by [Alessandro Bombelli](#), [Bruno F. Santos](#), L.A. (Lori) Tavasszy

This dataset contains (i) a set of worldwide airports that are relevant for the global air cargo transport (ii) a dataset containing aircraft-specific yearly recorded frequencies (referring to the year 2014) for all passenger and cargo airlines (integrators such as FedEx are excluded) for all different origin destination (OD) airport pairs (iii) a dataset for each integrator FedEx, UPS, DHL with yearly estimated cargo capacity (expressed in tonnes) referring to the year 2019 for every OD airport pair. The estimation was based on a dataset containing all recorded flights for each OD airport pair of interest, which was filtered to extrapolate only flights operated by the integrators

DataCite

Bombelli, Alessandro; Santos, Bruno F.; Tavasszy, L.A. (Lori) (2020): Air Cargo Transport Network (ACTN) Dataset. 4TU.ResearchData. Dataset. <https://doi.org/10.4121/uuid:5725add4-7fe8-41d1-a452-b1fc011e0bae>

<https://doi.org/10.4121/uuid:5725add4-7fe8-41d1-a452-b1fc011e0bae>

### CATEGORIES

- Aerospace Transport
- Transportation and Freight Services
- Aerospace Engineering

### KEYWORD(S)

Air cargo network

Air cargo traffic

Air cargo transportation

Freight services

### LICENCE

MIT

### CONTRIBUTORS

TU Delft, Faculty of Aerospace Engineering, Department of Control and Operations, Air Transport & Operations section;  
TU Delft, Faculty of Technology, Policy and Management, Department of Engineering, Systems and Services

### PUBLISHER

4TU.Centre for Research Data

### TIME COVERAGE

Passenger and cargo airlines: 2014; Integrators: second half of 2019

### GEOLOCATION

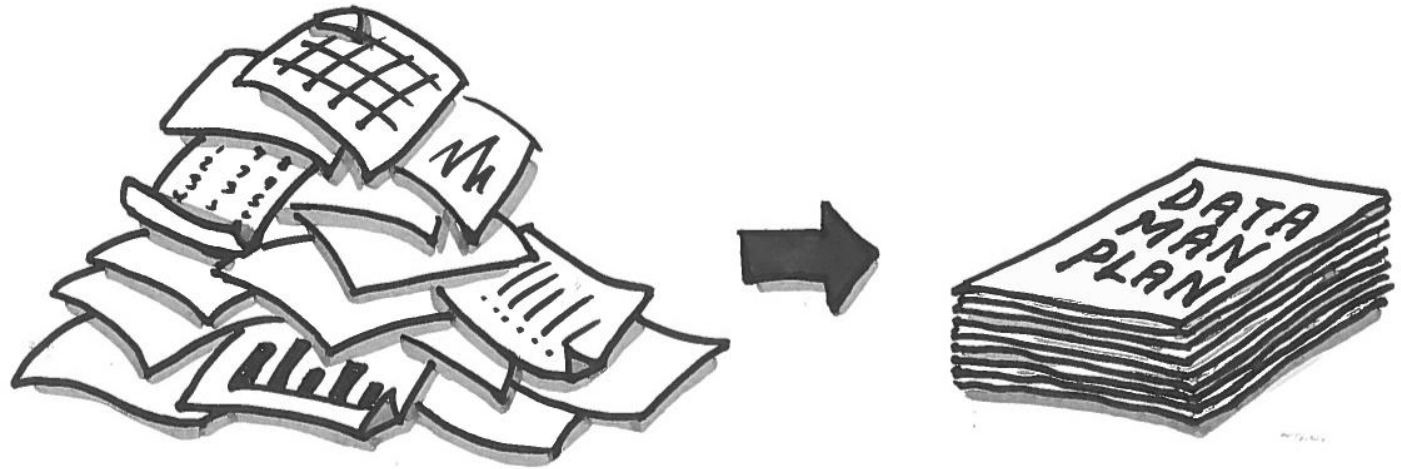
The World

### FORMAT

media types: application/vnd.openxmlformats-officedocument.spreadsheetml.sheet, application/x-7z-compressed, text/plain



# Plan de gestión de datos



# Plan de gestión de datos

- Ofrece una excelente oportunidad de:
  - Desarrollar una estrategia de documentación y comunicarla entre los miembros de un proyecto
  - Clarificar qué estándares se usarán en un proyecto y que políticas o legislación son relevantes de seguir dentro de un proyecto
  - Mantener al equipo/colaboradores informados y alineados acerca de las prácticas en gestión de datos del proyecto

# Qué se viene?


















## Towards FAIR principles for research software

Cite

**Issue title:** FAIR Data, Systems and Analysis

**Guest editors:** Paul Groth and Michel Dumontier

**Article type:** Position Paper

**Authors:** Lamprecht, Anna-Lena<sup>a,\*</sup>  | Garcia, Leyla<sup>b</sup>  | Kuzak, Mateusz<sup>c, d</sup>  | Martinez, Carlos<sup>e</sup>   
| Arcila, Ricardo<sup>f</sup>  | Martin Del Pico, Eva<sup>g</sup>  | Dominguez Del Angel, Victoria<sup>h</sup>  | van de Sandt, Stephanie<sup>i</sup>  | Ison, Jon<sup>j</sup>  | Martinez, Paula Andrea<sup>k</sup>  | McQuilton, Peter<sup>l</sup>  | Valencia, Alfonso<sup>m;</sup>   
<sup>n</sup>  | Harrow, Jennifer<sup>o</sup>  | Psomopoulos, Fotis<sup>p</sup>  | Gelpi, Josep LL.<sup>q; r</sup>  | Chue Hong, Neil<sup>s; t</sup>  |  
Goble, Carole<sup>u</sup>  | Capella-Gutierrez, Salvador<sup>v; \*\*</sup> 

Publicación: <https://content.iospress.com/articles/data-science/ds190026>

Recomendaciones: <https://fair-software.nl/>

Preguntas?  
Comentarios?  
Sugerencias?





# Contactos

Paula Martinez Lavanchy

Research Data Officer

**TU Delft Library** | **4TU.ResearchData**

Email: [p.m.martinezlavanchy@tudelft.nl](mailto:p.m.martinezlavanchy@tudelft.nl)

Twitter: [@paulammartinez2](https://twitter.com/paulammartinez2)

ORCID: 0000-0003-1448-0917