



Leibniz Institute of
Ecological Urban and
Regional Development

Module 3 Knowledge building blocks



RDM Knowledge building blocks

- Repositories
- Persistent Identifiers
- Metadata standards
- Licenses
- Version Control Systems



Data repository

A data repository is an **archiving space** for researchers to deposit research data linked with their research.

“A **trusted** digital repository is one whose mission is to provide reliable, **long-term** access to managed digital resources to its designated community”.



www.coretrustseal.org



Repositories

General-purpose Repositories



Domain/Subject/Discipline Specific Repositories



Institutional Repositories



Zenodo Community

- From **Grey to FAIR** research outputs
- Establish your own community
- Central hub linking research outputs

Zenodo IOER Community

Zenodo: How-To upload guide

The screenshot shows the Zenodo website interface for the Leibniz Institute of Ecological Urban and Regional Development (IOER) community. At the top, there is a blue navigation bar with the Zenodo logo, a search bar, and links for 'Upload' and 'Communities'. The user profile 's.della_chiesa@ioer.de' is visible in the top right corner.

The main content area is titled 'Leibniz Institute of Ecological Urban and Regional Development (IOER)'. Below this, there is a 'Recent uploads' section with a search bar and a 'New upload' button. The uploads are listed as follows:

- September 6, 2019 (1.0) Conference paper Open Access**
Quantitative assessment of public sector eMobility Infrastructure development - A Spatial Perspective in case of Germany
Srijit Kumar
Spatial data-driven approaches are promising for evidence-based public policymaking (e.g. spatial planning) by combining multi-dimensional open data and emerging technologies besides traditional data sources. With this background, this study aims to assess spatial diffusion patterns its potential etc.
Uploaded on November 2, 2022
- August 17, 2022 (v1) Other Open Access**
Plea for more affection in politics - An experiential essay on the 9-Euro-Ticket
Müller, Susanna
9-Euro-Ticket: July 22nd, 2022, 12:00 p.m. - Heavis on DfP: According to the German Federal Minister of Finance Lindner and the opposition parties of the CDU/CSU, the 9-Euro-Ticket should not be extended because 'the concept is not convincing'.
Uploaded on November 1, 2022
- October 12, 2022 (v1) Other Open Access**
Moving from crisis to resonance: a summer cinema
Harms, Philip; Killinger, Mabel; Müller, Susanna; Artmann, Martina
Essay 10. October 2022, URBANANCE Blog (https://urbanance.ioer.info/en/blog/)
Uploaded on October 12, 2022
- October 10, 2022 (1.0) Data management plan Open Access**
Data Management Plan URBANANCE Project
Artmann, Martina; Harms, Philip; Müller, Susanna; Killinger, Mabel; Della Chiesa, Stefano
In the era of exceeding planetary boundaries and rapid urbanisation, (re-)connecting society and people with nature can be considered a major leverage point for sustainability transformation. Surprisingly, we know little about the human-nature connector in an urban and sustainable context. The proto
Uploaded on October 10, 2022
- September 26, 2022 (1.0) Preprint Open Access**
Research Data Strategy 'Anno Zero' - Adopting & Adapting the RISE-DE reference model
Stefano Della Chiesa
Research Data Management (RDM) is a cross-functional, multi-level, multidisciplinary endeavour among various actors. The key in RDM is to maximise the research data value chain, knowledge transfer and impact. This poster was presented at...

On the right side, there is a 'Community' section with the IOER logo and name. Below it, a description of the IOER research centre and its mission is provided. A 'Curated by' section lists Stefano Della Chiesa. A 'Curation policy' section explains the rules for research records in the community. At the bottom, there is a 'Created' date (October 4, 2021) and a 'Harvesting API' link.



The future IOER RDC/ IÖR FDZ Data Repository

Share, archive, and get credit for your data. Find and cite data across all research fields.

Search this dataverse... **Advanced Search**

1 to 10 of 165,320 Results

- Datasets (6,052)
- Files (159,265)
- Files (1,816,193)

Dataverse Category
 Research Project (2,231)
 Researcher (1,895)
 Organization or Institution (487)
 Research Group (420)
 Journal (125) [More...](#)

Metadata Source
 Harvard Dataverse (86,803)
 Harvested (78,517)

Publication Year
 2022 (18,999)
 2021 (23,039)
 2020 (9,614)
 2019 (4,519)
 2018 (4,825) [More...](#)

Subject
 Social Sciences (59,336)
 Arts and Humanities (35,012)
 Earth and Environmental Sciences (11,833)
 Medicine, Health and Life Sciences (8,456)
 Law (5,722) [More...](#)

Author Name
 Master, Daniel M. (28,107)
 Stager, Lawrence E. (28,107)
 Curtis A. Bradley (4,940)
 Oona A. Hathaway (4,938)
 Jack L. Goldsmith (4,270) [More...](#)

Author Affiliation
 Harvard University (30,231)
 Wheaton College (28,171)
 Harvard Law School (4,952)

1 to 10 of 165,320 Results

Predictors of delay in seeking treatment for Pelvic organ prolapse in Wolaita Zone, Southern Ethiopia. Mixed method study
 Nov 1, 2022
 Wana, Ermas Wabeto, 2022, "Predictors of delay in seeking treatment for Pelvic organ prolapse in Wolaita Zone, Southern Ethiopia: Mixed method study", <https://doi.org/10.7910/DV/NFCSXYM>, Harvard Dataverse, V1, UNF:0:CG0k08Soj0ep0jly15A== [fileUNF]
 This is the data set we used during the analysis of the predictors of delay for Pelvic organ prolapse

Digital Marketing Foundations
 Nov 1, 2022
 Zoe, 2022, "Digital Marketing Foundations", <https://doi.org/10.7910/DV/NQ0V5GO>, Harvard Dataverse, V1
 Digital marketing is one of the most important concepts to grasp because it includes all of the most effective marketing strategies and is constantly evolving. A few business owners formerly searched for a simple method of reaching their clients with their goods and services. The...

Replication Data for: Measuring the Impact of LGBTQ Activism on Backlash in the Former Eastern Bloc
 Oct 31, 2022
 Chen, Alex, 2022, "Replication Data for: Measuring the Impact of LGBTQ Activism on Backlash in the Former Eastern Bloc", <https://doi.org/10.7910/DV/NCQ7FSS>, Harvard Dataverse, V1, UNF:0:zlat2KQn4WBVYyJcig== [fileUNF]
 Contains .dta and .do files to replicate the analyses, tables, and figures in the paper, appendix, and online appendix.

Replication Data for: Political Leadership Has Limited Impact on Fossil Fuel Taxes and Subsidies
 Oct 31, 2022
 ROSS, MICHAEL, Hazlett, Chad, Mahdavi, Paasha, Martinez Alvarez, Cesar, 2022, "Replication Data for: Political Leadership Has Limited Impact on Fossil Fuel Taxes and Subsidies", <https://doi.org/10.7910/DV/NJCHDJS>, Harvard Dataverse, V1
 Replication Data for: Political Leadership Has Limited Impact on Fossil Fuel Taxes and Subsidies

Oregon Theater Project Database
 Oct 31, 2022 - Scholars' Bank: Data
 Aronson, Michael; Peterson, Elizabeth; Hayden, Gabriele, 2022, "Oregon Theater Project Database", <https://doi.org/10.7910/DV/NFQ0J23>, Harvard Dataverse, V3, UNF:0:8Z7q8v8XusTTNhwjcdg== [fileUNF]
 This is the data associated with the website "Oregon Theater Project: Exhibition and Moving in the Silent Era" (<https://otp.oregon.edu/>). Website content is overseen and maintained by Michael Aronson and Elizabeth Peterson. The website is a Drupal site built and maintained b...

Replication Data for: A Diffusion Network Event History Estimator
 Oct 31, 2022 - The Journal of Politics Dataverse
 Harden, Jeffrey J., 2022, "Replication Data for: A Diffusion Network Event History Estimator", <https://doi.org/10.7910/DV/MH1VLZ>, Harvard Dataverse, V2
 Research on the diffusion of political decisions across jurisdictions typically accounts for units' influence over each other with (1) observable measures and/or (2) by inferring latent network ties from past decisions. The former approach assumes that interdependence is static...

2021 Canadian Election Study (CES)
 Oct 31, 2022 - Consortium on Electoral Democracy (C-DEM)

IOER Dataverse Search User Guide Support English Sign Up Log In

Root

Metrics 3 Downloads

The root dataverse.

Search this dataverse... **Advanced Search**

1 to 6 of 6 Results

- Datasets (3)
- Files (14)

Publication Year
 2023 (6)

Author Name
 Della Chiesa, Stefano (2)
 Admin, Dataverse (1)

Subject
 Earth and Environmental Sciences (4)

Keyword Term
 Local government districts (1)
 Ortschaften (1)

Deposit Date
 2023 (3)

Test GeoData STD (Leibniz IOER)
 Apr 26, 2023

Social Science Collection (Leibniz IOER)
 Apr 25, 2023

Gemarkungen sachsenweit (Leibniz IOER)
 Apr 21, 2023 - Geospatial Collections
 Della Chiesa, Stefano, 2023, "Gemarkungen sachsenweit", https://doi.org/10.82617/ior-fdz/RDXHV_Root_V1
 Gemarkungen - Shape - EPSG: 25833 (ETRS89_UTM33) - sachsenweit Download 27 10 2022 11:34:05 Nutzungsbedingungen
 Dieser Datensatz kann gemäß den Nutzungsbestimmungen von DatenZentrum Deutschland – Namensnennung – Version 2.0 (<https://www.govdata.de/dl-de/by-2-0>) unter Angabe der...

Ortsstelle Sachsen (Leibniz IOER)
 Apr 21, 2023 - Geospatial Collections
 Della Chiesa, Stefano, 2023, "Ortsstelle Sachsen", https://doi.org/10.82617/ior-fdz/BPE252_Root_V2
 Die Ortsstelle stehen als sachsenweite Abgabe zum kostenfreien Download bereit. Sie erhalten ein Zip-Archiv mit einer Shape-Datei. Diese enthält Polygone der aktuellen Ortsstelle mit beschreibenden Informationen wie Ortsname, Postleitzahl, Schlüssel, die zugehörige Gemeind...

Geospatial Collections (Leibniz IOER)
 Apr 21, 2023
 Collections of Geospatial data sets

DOI Alias pwest (Admin, Dataverse)
 Apr 20, 2023
 Admin, Dataverse, 2023, "DOI Alias pwest", https://doi.org/10.82617/ior-fdz/QLPQ99_Root_V1
 1123hahaha



Registry of research data repositories

Browser resources in the global registry of research data repositories



Filter

[Reset all](#)

Subjects

- Humanities and Social Sciences (1)
- Natural Sciences (2)**
 - Geosciences (including Geography) (2)
 - Atmospheric Science and Oceanography (1)
 - Atmospheric Science (1)
 - Oceanography (1)
 - Geology and Palaeontology (1)
 - Water Research (1)
 - Hydrogeology, Hydrology, Limnology, Urban Water Management, Water Chemistry, Integrated Water Resources Management (1)

Content Types

- Archived data (1)
- Databases (2)**
- Images (2)
- Plain text (2)
- Scientific and statistical data formats (2)
- Standard office documents (2)
- Structured graphics (1)
- Structured text (1)

Countries

- Germany (2)**
- International (1)

API

Certificates

CoreTrustSeal (2)



Persistent Identifiers (PIDs)



- PID is a unique code



- There are various PIDs for various type of objects



- PID are persistent and linked to essential info (Metadata)



- PIDs helps to track objects itself



- PIDs link to other resources



Persistent Identifiers (PIDs)

- PIDs are created by [PID Authorities](#) Providers ([DataCite](#), [Crossref](#), [ROR](#), [ORCID](#), etc.)
- PIDs help in a number of things:
- PIDs make your research FAIR



*Dr. Diana Berger
Dr. Diana Smith*

*Dr. John Smith
Dr. J. Smith
Dr. John S.
PhD. John Smith*

PIDs examples

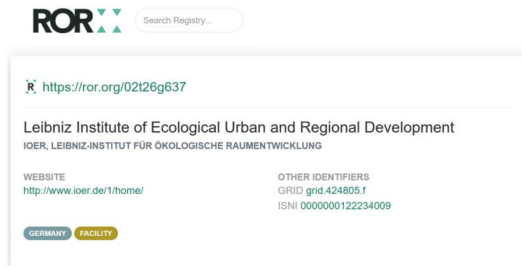
People



Name
Stefano Della Chiesa

<https://orcid.org/0000-0002-6693-2199>

Organization



<https://ror.org/02t26g637>

Research output



April 29, 2022 Technical note Open Access

Research Data Management - Guidance

Della Chiesa, Stefano

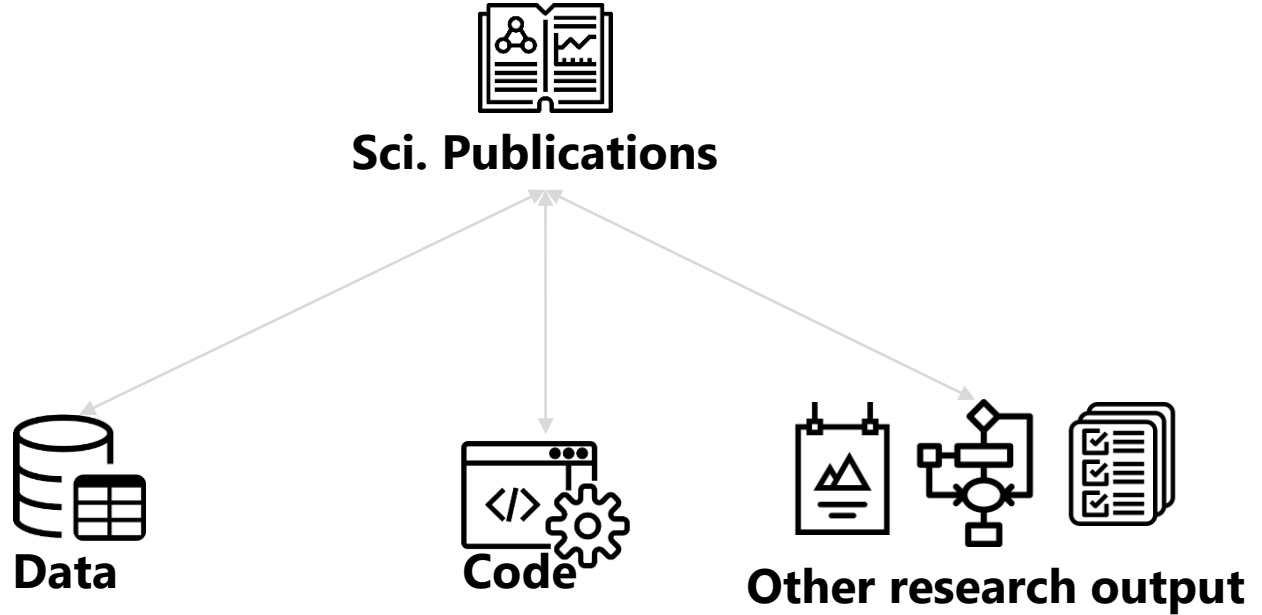
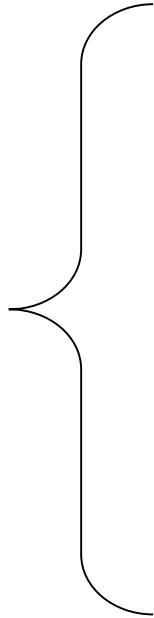
Research Data Management (RDM) provides the strategies, guidelines, and measures to manage research outputs. Like a jigsaw, RDM requires multiple pieces of multidisciplinary knowledge and competencies. RDM Checklist guides the researchers to FAIRLY manage the research outputs along the project lifecycle and beyond it. Research projects are inherently iterative in their try and error phases, and consequently, RDM is a continuous adapting process throughout the research project. Data Management Planning (DMP) is the operational tool to handle this complexity dynamically, in a structured way, ensuring compliance with policies and best practices. This document provides the RDM checklist and guidance for the Leibniz Institute of Ecological Urban and Regional Development (IOER).

<https://doi.org/10.5281/zenodo.6504928>

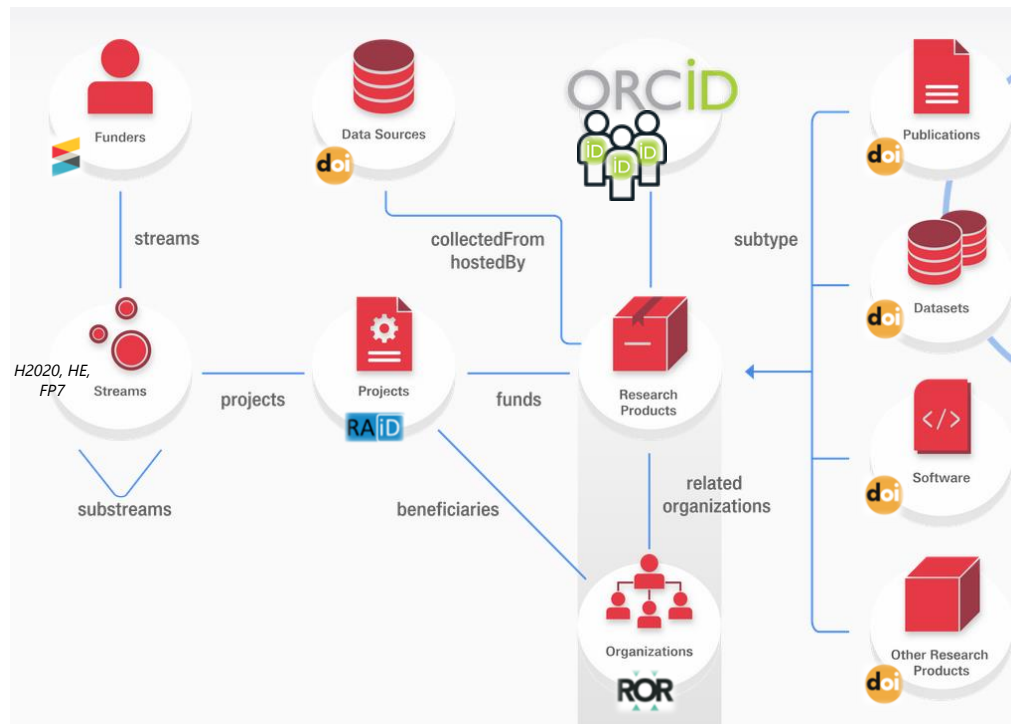
Controlled syntax









DOI (Digital Object Identifier)



PID linked research









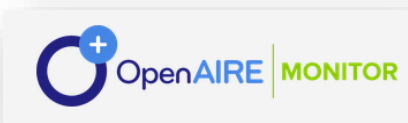
-  Crossref
-  Datacite
-  ORCID
-  Digital Object Identifiers (DOI)
-  Research Activity Identifier (RAiD)
-  Research Organization Registry (ROR)

Graphic adapted from: <https://graph.openaire.eu/>

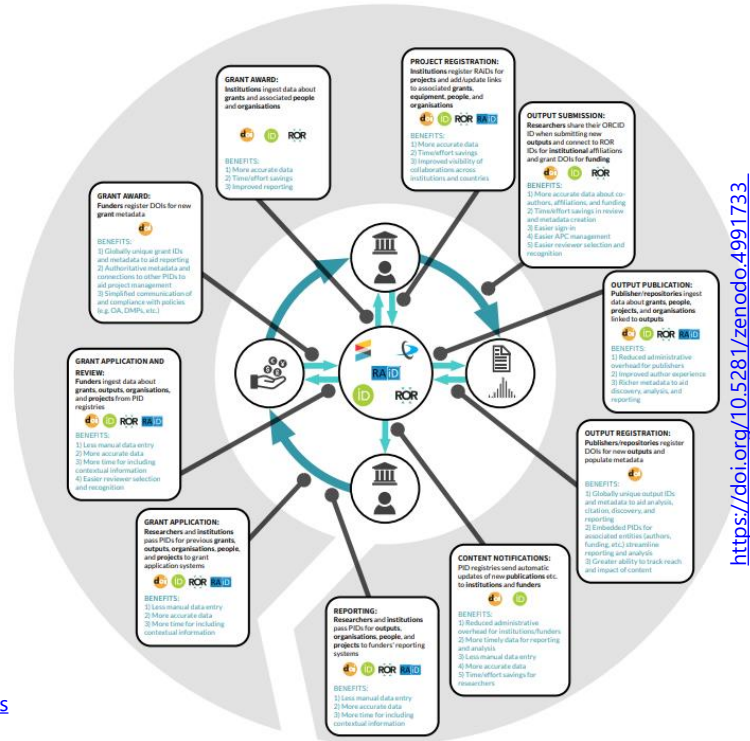
The bigger picture

The PID-optimised research lifecycle

-  Crossref
-  Datacite
-  ORCID
-  Digital Object Identifiers (DOI)
-  Research Activity Identifier (RAID)
-  Research Organization Registry (ROR)



- Example Funding dashboard: <https://monitor.openaire.eu/dashboard/gau/funding/overview>
- Example Publication dashboard: <https://monitor.openaire.eu/dashboard/gau/research-output/publications>
- Example Open Science routes dashboard: <https://monitor.openaire.eu/dashboard/gau/open-science/publications>
- Example Impact dashboard: <https://monitor.openaire.eu/dashboard/gau/impact->



<https://doi.org/10.5281/zenodo.4991733>



The PID Landscape

Heterogeneous Landscape & Maturity

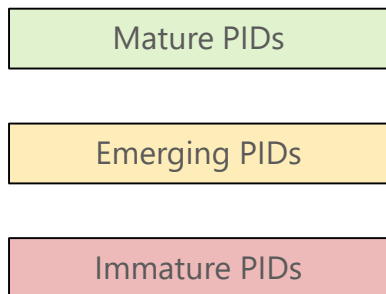


Table 1. A Landscape Analysis of Research Entities, Identifying PID Types and Infrastructure Maturity (as of May 2018)

Research Entity	PID Types Used	Maturity of PID Infrastructure
Publication	DOI, accession number, handle, URN, Scopus EID, Web of Science UID, PMID, PMC, arXiv identifier, BibCode, ISSN, ISBN, PURL	mature
Citation	OCI (secondary aggregation of information)	emerging
Conference	DOI, accession number	emerging
Researcher (or scholar)	ORCID IDs, ISNI (also DAIs, VIAFs, arXivIDs, OpenIDs, ResearcherIDs, ScopusIDs)	mature
Organization	DOI, ISNI, GRID, Ringgold IDs, ROR IDs	emerging
Data	DOI, accession number, handle, PURL, URN, ARK	mature
Data repository	none	immature
Grants	DOI, PURL	emerging
Project	local identifier, accession number, RaID	emerging
Experiment	none	immature
Investigation	DOI, accession number	emerging
Analysis	GitHub gist	immature
Software	DOI, SHA-1 hash	emerging
Computer simulation	UUID	emerging
Software license	none	immature
Equipment		
Instrument, device, sensor, platform, research facility	DOI, RRID, UID	emerging
Archival/storage facility	URI, DOI, UUID	emerging
Field station	none	immature
Sample		
Geological or biological sample	accession number, RRID, DOI, IGSN	emerging
Cultural artifact	DOI, URN, accession number	emerging
Historical or mythical person	URI	emerging
Temporal period and historical place	ARK, URI, accession number	immature
Study Registration		
Clinical trial; non-clinical registration	accession number, DOI	immature
Data management plan	DOI	immature
Workflow	URI, DOI	immature
Protocol	DOI	immature

Adapted from Ferguson et al.^[12]. ARK: Archival Resource Key, BibCode: Bibliographic Codes, DAI: Digital Author Identifier, DOI: Digital Object Identifier, ID: Identifier, IGSN: International Geo Sample Number, ISBN: International Standard Book Number, ISNI: International Standard Name Identifiers, ISSN: International Standard Serial Number, OCI: Open Citation Identifier, ORCID: Open Researcher and Contributor ID, PMID: PubMed ID, PURL: Persistent Uniform Resource Locators, RAID: Research Activity identifier, RRID: Research Resource ID, SHA-1: Secure Hash Algorithm 1, UUID: Universally Unique Identifiers, URI: Uniform Resource Identifier, URN: Uniform Resource Name, VIAF: Virtual International Authority Files

Connected Research: [The Potential of the PID Graph](#)

Metadata & Standards

No Metadata



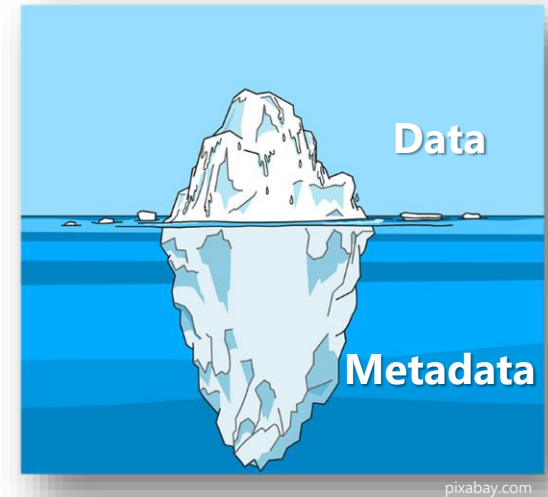
Metadata



hiclipart.com

Data and Metadata (MD)

- Research data includes **measurements, audio-visual information, texts, survey data, samples, questionnaires, algorithms, software, simulations**, etc.
- Metadata represents all the **information that add context and meaning** in a **structured** way to the (research) data.



Metadata Standards

- Metadata is fundamental as makes data findable and reusable.
- Metadata **standards** or **schemas** allow a structured description of your data by using specific MD elements.
- There are general purpose MD schemas.
- There are also disciplines with subject/domain specific metadata standards.



[List of Disciplinary Metadata](#)

[Use cases](#)

Metadata Standards

General purpose MD standard (e.g. Simple Dublin Core)

Table 1. The Fifteen Elements of “Simple Dublin Core”

Identifier	Definition
Title	A name given to the resource.
Creator	An entity primarily responsible for making the content of the resource.
Subject	The topic of the content of the resource.
Description	An account of the content of the resource.
Publisher	An entity responsible for making the resource available.
Contributor	An entity responsible for making contributions to the content of the resource.
Date	A date associated with an event in the life cycle of the resource.
Type	The nature or genre of the content of the resource.
Format	The physical or digital manifestation of the resource.
Identifier	An unambiguous reference to the resource within a given context.
Source	A reference to a resource from which the present resource is derived.
Language	A language of the intellectual content of the resource.
Relation	A reference to a related resource.
Coverage	The extent or scope of the content of the resource.
Rights	Information about rights held in and over the resource.



<https://www.dublincore.org/specifications/dublin-core/usageguide/2001-04-12/generic/#coverage>



Metadata Standards

General purpose MD standard (e.g. DataCite)

Mandatory	Recommended	Optional
Identifier	Subject	Language
Creator	Contributor	Alternate ID
Title	Date	Size
Publisher	Resource Type	Format
Publication year	Related identifier	Version
	Description	Rights
	GeoLocation	



https://schema.datacite.org/meta/kernel-4.4/doc/DataCite-MetadataKernel_v4.4.pdf



Metadata Standards



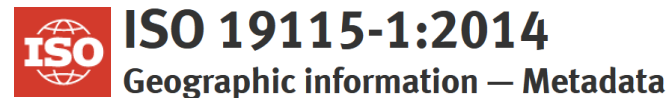
Disciplines MD standard (e.g. Data Documentation Initiative - DDI)

“For describing the data produced by surveys and observational methods in the **social and economic sciences**”.

A screenshot of the gesis website. The header includes the gesis logo (Leibniz Institute for the Social Sciences), a login button, a language selector set to German, and an email icon. Below the header is a search bar containing the text 'data'. A navigation menu below the search bar includes 'Services', 'Research', and 'Institute'. The main content area shows a '< Back' link and a search result for 'Mobile Data Collection - Incentive Experiment' by Florian Keusch. The result includes the text 'GESIS Data Archive, Cologne. ZA6978 Data file Version 1.0.0, https://doi.org/10.4232/1.13247' and an abstract describing the study's goal to measure the influence of incentive schemes on smartphone data collection among German users.



Metadata Standards



Disciplines MD standard (e.g. ISO 19115)

“For describing for
Geographical data”.



Resource title: Corine Land Cover Change (CHA) 2012 - 2018, Version 2020_20u1

Resource abstract: CHA1218 is one of the Corine Land Cover (CLC) datasets produced within the frame the Copernicus Land Monitoring Service referring to land cover / land use status between the years 2012 and 2018. CLC service has a long-time heritage (formerly known as “CORINE Land Cover Programme”), coordinated by the European Environment Agency (EEA). It provides consistent and thematically detailed information on land cover and land cover changes across Europe.

CLC datasets are based on the classification of satellite images produced by the national teams of the participating countries - the EEA members and cooperating countries (EEA39). National CLC inventories are then further integrated into a seamless land cover map of Europe. The resulting European database relies on standard methodology and nomenclature with following base parameters: 44 classes in the hierarchical 3-level CLC nomenclature; minimum mapping unit (MMU) for status layers is 25 hectares; minimum width of linear elements is 100 metres. Change layers have higher resolution, i.e. minimum mapping unit (MMU) is 5 hectares for Land Cover Changes (LCC), and the minimum width of linear elements is 100 metres. The CLC service delivers important data sets supporting the implementation of key priority areas of the Environment Action Programmes of the European Union as e.g. protecting ecosystems, halting the loss of biological diversity, tracking the impacts of climate change, monitoring urban land take, assessing developments in agriculture or dealing with water resources directives. CLC belongs to the Pan-European component of the Copernicus Land Monitoring Service (<https://land.copernicus.eu/>), part of the European Copernicus Programme coordinated by the European Environment Agency, providing environmental information from a combination of air- and space-based observation systems and in-situ monitoring. Additional information about CLC product description including mapping guides can be found at https://land.copernicus.eu/user-corner/technical-library/clc2018technicalguidelines_final.pdf. CLC class descriptions can be found at <https://land.copernicus.eu/user-corner/technical-library/corine-land-cover-nomenclature-guidelines/html/>.

Resource type: Dataset

Resource locator: <https://land.copernicus.eu/pan-european/corine-land-cover/lcc-2012-2018>

Topic of category: Environment, Imagery/Base Maps/Earth Cover

Keyword: Copernicus Land Satellite Image Interpretation 2013 Corine Corine Land Cover Change 2012

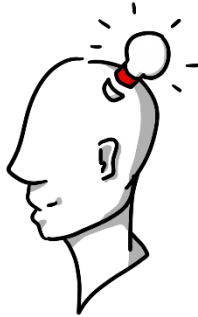
Bounding Box: West = -31.561261
East = 44.820775
North = 71.409109
South = 27.405827

Bounding Box: West = -61.906047
East = -60.905616
North = 16.607552
South = 15.736333





Intellectual Property (IP)



*“Creation of the mind”
(inventions, literary & artistic
work, designs, names)*

Intellectual Property Right (IPR)



“Legal tool that protect IP”

IP & IPR

Copyright is the common IPR for research output

* Corresponding author.

E-mail address: stefano.dellachiesa@eurac.edu (S. Della Chiesa).



<https://doi.org/10.1016/j.geoderma.2019.02.010>

Received 19 June 2018; Received in revised form 1 February 2019; Accepted 4 February 2019
Available online 16 February 2019

0016-7061 © 2019 The Authors. Published by Elsevier B.V. This is an open access article under
(<http://creativecommons.org/licenses/by/4.0/>).

*Work protected from the moment is created
([No registration](#) authority)*

Database IPR [sui generis database right](#)

Note: RAW data is generally not protected by IPR.

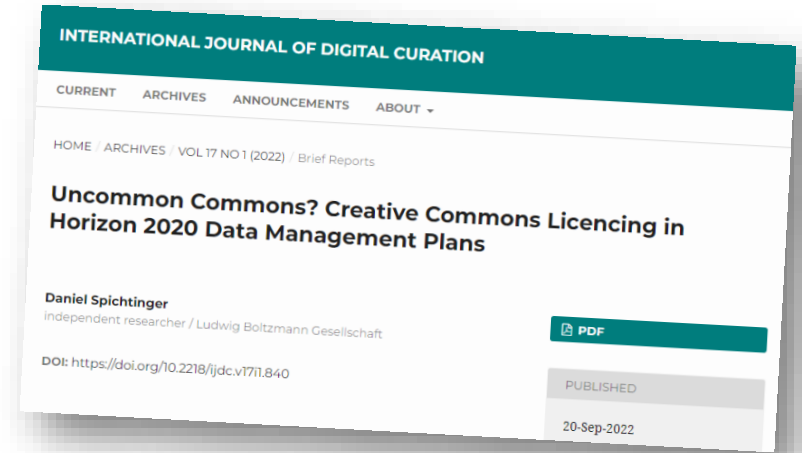
**Copyright protect you by default.
License provides the use conditions.**

External resources: [The EU IP-Helpdesk](#)

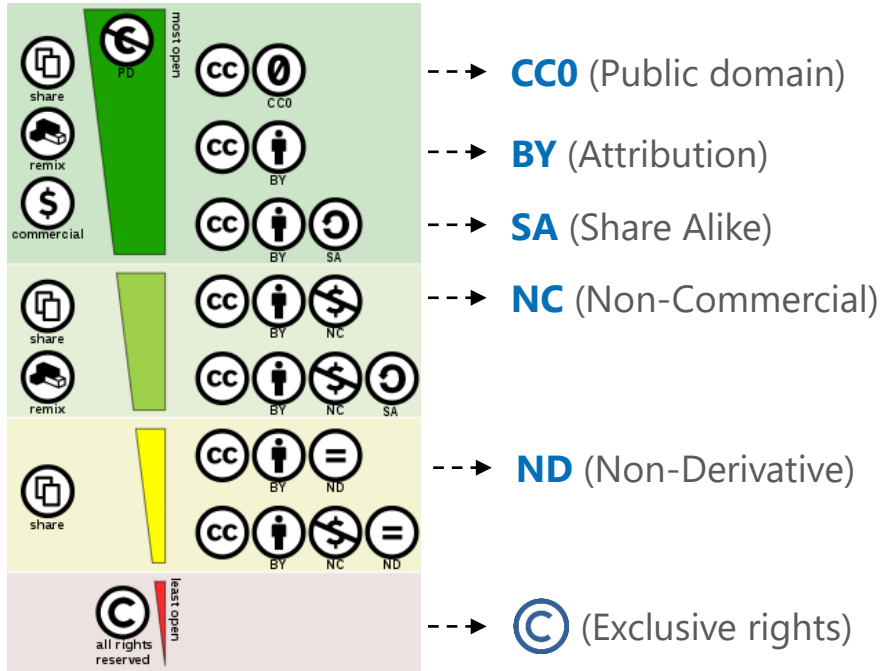
Research Output Licenses

Still lack of knowledge

- Users **lack of understanding** what they can do with the data in question
- In Horizon 2020 projects still lack of sufficient **comprehension of licensing**
- License are fundamental for the **downstream reuse** of data



Research Output Licenses



CREATIVE COMMONS LICENSES

	COPY & PUBLISH	ATTRIBUTION REQUIRED	COMMERCIAL USE	MODIFY & ADAPT	CHANGE LICENSE
PUBLIC DOMAIN	✓	✗	✓	✓	✓
CC BY	✓	✓	✗	✓	✓
CC BY-SA	✓	✓	✓	✓	✗
CC BY-ND	✓	✓	✗	✗	✓
CC BY-NC	✓	✓	✗	✓	✓
CC BY-NC-SA	✓	✓	✗	✓	✗
CC BY-NC-ND	✓	✓	✗	✗	✓

✓ You can redistribute (copy, publish, display, communicate, etc.)
 ✓ You have to attribute the original work
 ✓ You can use the work commercially
 ✓ You can modify and adapt the original work
 ✓ You can choose license type for your adaptations of the work.

Shaddim; [original CC license symbols by Creative Commons](#), CC BY 4.0

Creative Commons licenses, [by Foter](#) CC-BY-SA

License: Inbound vs Outbound



Compatibility Checker

	PUBLIC DOMAIN	PUBLIC DOMAIN	CC BY	CC BY SA	CC BY NC	CC BY ND	CC BY NC SA	CC BY NC ND
PUBLIC DOMAIN	✓	✓	✓	✓	✓	✗	✓	✗
PUBLIC DOMAIN	✓	✓	✓	✓	✓	✗	✓	✗
CC BY	✓	✓	✓	✓	✓	✗	✓	✗
CC BY SA	✓	✓	✓	✓	✗	✗	✗	✗
CC BY NC	✓	✓	✓	✗	✓	✗	✓	✗
CC BY ND	✗	✗	✗	✗	✗	✗	✗	✗
CC BY NC SA	✓	✓	✓	✗	✓	✗	✓	✗
CC BY NC ND	✗	✗	✗	✗	✗	✗	✗	✗

https://wiki.creativecommons.org/License_Compatibility



Browse resources by subject



[Standards](#)

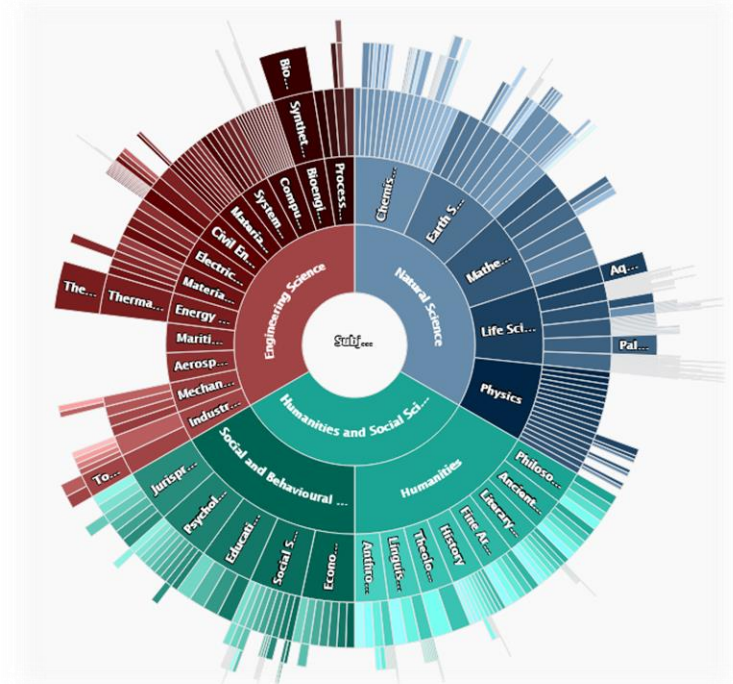


[Database/Repositories](#)



[Policies](#)

freicons.io/



<https://fairsharing.org/browse/subject>

Data organization

Naming conventions

No way!

Outline.docx

My.conf~abstrc#qwertz > <

Presentation 1 .ppt

figure.jpg

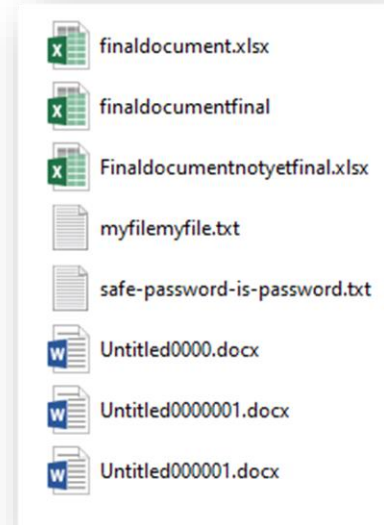
Better!

2022-05-20_RDM_Seminar_Outline_v01.docx

2022-05-22_Abstract_Nexus_Conference.docx

01_DMP_Introduction.pptx

Fig01_ScatterPlot_QuercusPetrea_Height_vs_Age.jpg



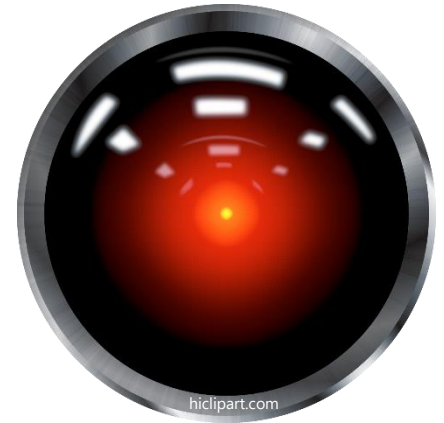
**Is this really
someone else's folder?**

Naming convention

Machine readable

(is not their fault if you are human)

- **No spaces, or special characters** `{[("{$%&?*+ #,.;`^ < > °]}`
programs might not be able to read the files
- **Use delimiters “_” or “-” to frame keywords and metadata**
(makes search easier, listing and extract info from file names)
- **File ordering, left pad numbers with zeros** (01_Draft_*
02_Draft_* and not 1_Draft_*, 10_Draft_*)



Naming convention

Human readable

(pepl mght undrst ths bt isnt tht practcl)

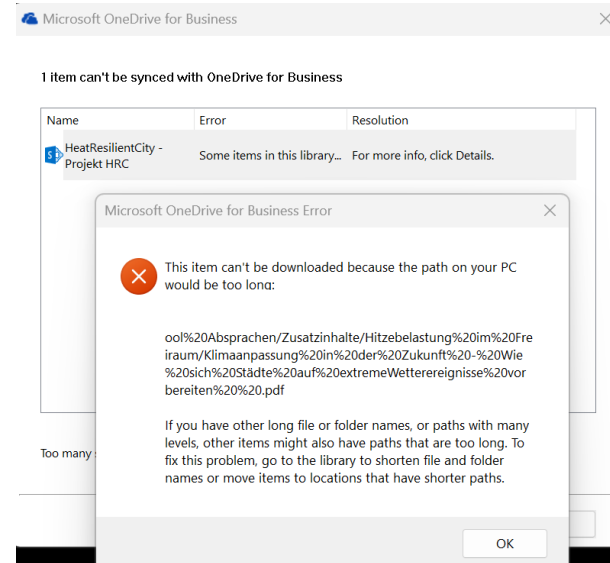
- **Use the standard ISO 8601** strings for dates (*2022-05-20* or *20220520*)
- **Left pad numbers with zeros** for logical ordering (*01_Introduction.txt* *02_StudyArea*) or use the dates (*20220520_Kickoff_Presentation.pptx*)
- Use **Capital letters** *20220520_MeetingNotes_ProjectKickoff*)
- **File versioning** (*01_Introduction_v01.txt*)
- **Encode other Metadata** (*Author/Revision John Smith*
01_Introduction_v01_Rev_JoS.txt)



*Think you search for a file
one hour before a deadline*

File and Folder's name length

Less is more



Naming convention

Document your Naming Convention

- Some naming best practices are project independent
- Some other are specific
- Think about what better fits your needs/project
- Attach your documentation as appendix to the DMP

Additional resources

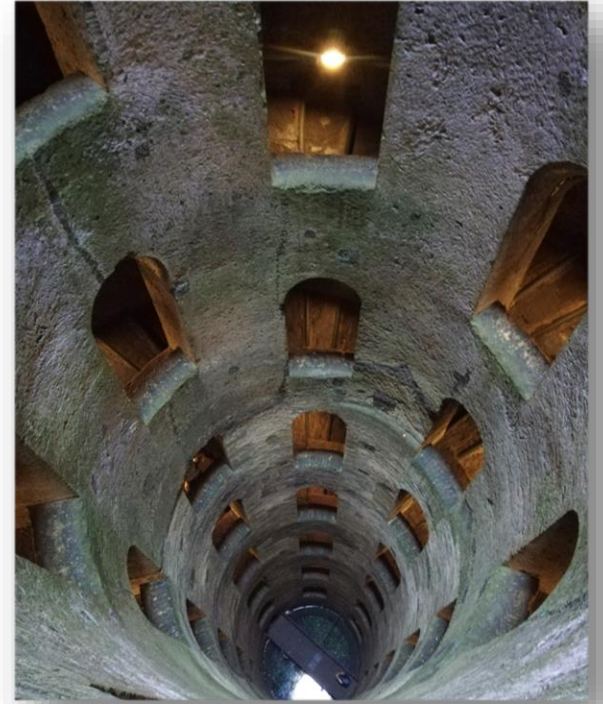
[Harvard naming convention](#)

[OstData Ordner strukturieren und Dateien benennen](#)



Data organization

Storage, Backup & Recovery Strategy



Data organization

Storage, Backup & Recovery Strategy

Organizations → 3-2-1 Backup rule



**Primary
data**



3 copies



**2 types
Storage Media**



1 Off-Site



Research Data Management
Storage, Backup & Recovery Strategy



<https://doi.org/10.5281/zenodo.5792643>



Cloud and collaboration

- Never save you data on your local PC!
- Your organization can provide you cloud storage for collaboration, with versioning and backup (FILR/[MS SharePoint](#))
- IOER Intranet info ([FILR/MS SharePoint](#))



Blue screen of death

Version Control Systems

Where is my **FAIR** code?

```
17 string sInput;
18 int iLength, iN;
19 double dblTemp;
20 bool again = true;
21
22 while (again) {
23     iN = -1;
24     again = false;
25     getline(cin, sInput);
26     system("cls");
27     stringstream(sInput) >> dblTemp;
28     iLength = sInput.length();
29     if (iLength < 4) {
30         again = true;
31         continue;
32     } else if (sInput[iLength - 3] != '.') {
33         again = true;
34         continue;
35     } while (++iN < iLength) {
36         if (isdigit(sInput[iN])) {
37             continue;
38         } else if (iN == (iLength - 3)) {
39             continue;
40         }
41     }
42 }
```

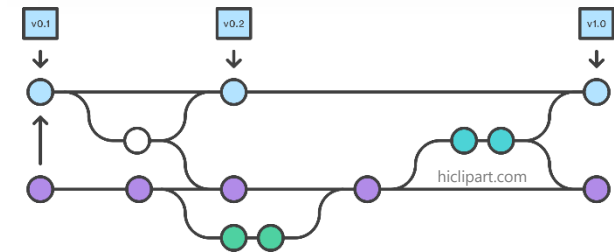
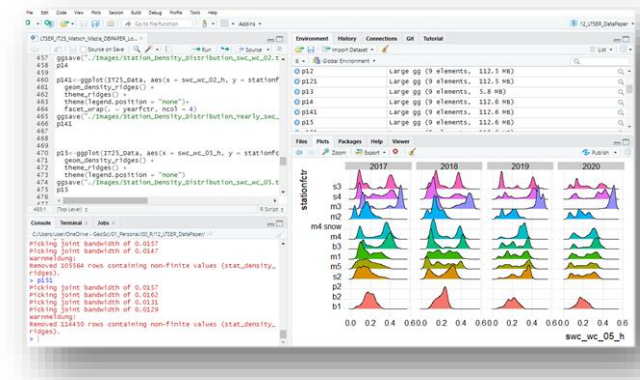
pixabay.com

Nowadays many journals require FAIR algorithms that reproduce results described in a scientific publication.

Management of algorithm and software

Data processing, model building are commonly executed in a scripting environment using various high-level programming languages (e.g. [R](#), [Python](#), [MATLAB](#))

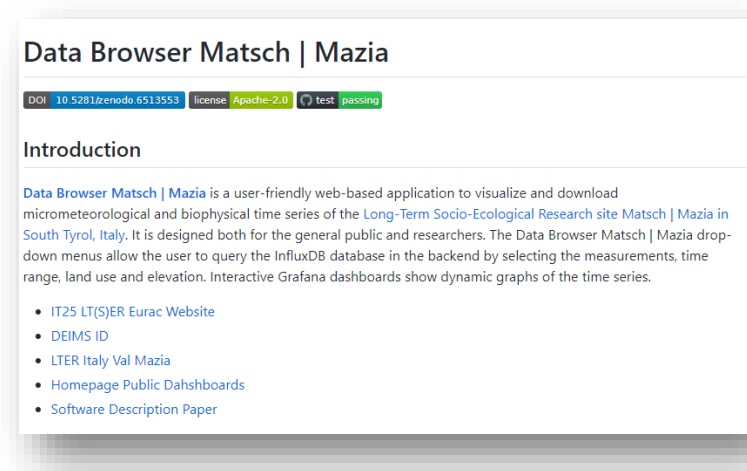
Researcher that work with code and scripts use version control systems to manage changes and collaborate. [GitHub](#) and [GitLab](#) are web-based hosting platform to manage and publish code.



Github-Zenodo Integration

- Easy to publish on Zenodo your Github repository
- Ensure reproducibility and long-term archiving
- Zenodo version management by issuing a new DOI each time you create a new GitHub release

[How to link Github and Zenodo](#)



Coding best practices

Make sure you can understand and reuse your code (**six months later!**)

[Checkout some simple best practices](#)

When your code is working...



...but it is a mess



Good habits rely
on practising