

Make your data FAIR with IPGP Research Collection

Hands-on training – IPGP 2023-03-09

What is FAIR data?

What is FAIR Data?

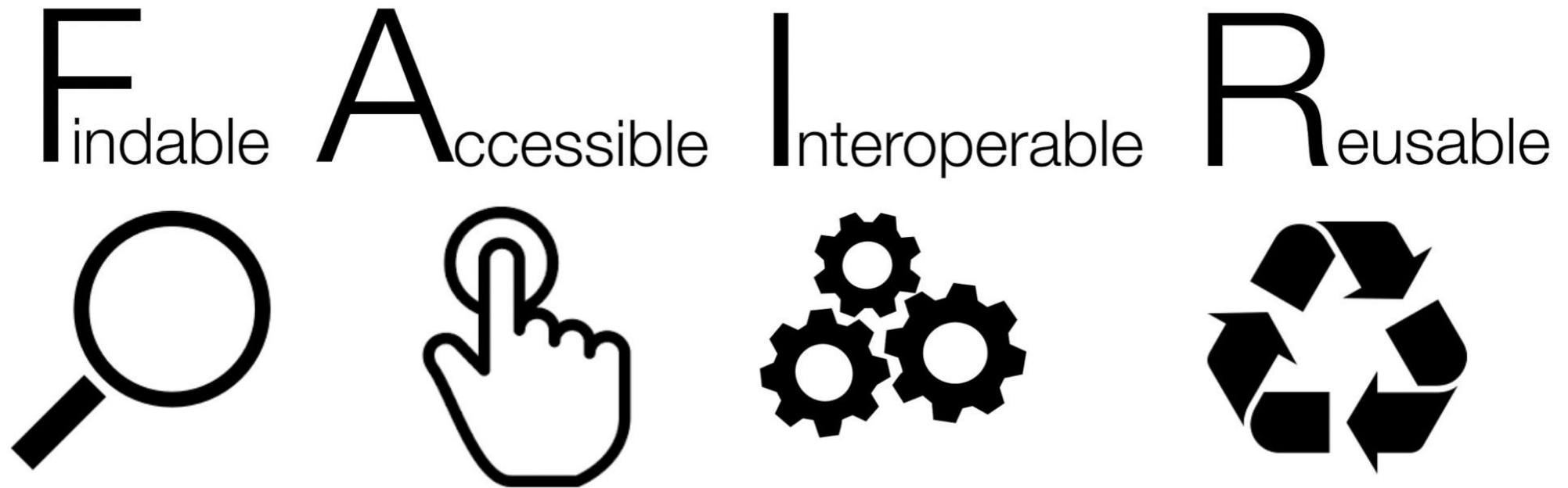


Image credit: https://commons.wikimedia.org/wiki/File:FAIR_data_principles.jpg (CC-BY-SA)

What is FAIR Data? (Highlighted items are up to you!)

F
indable



- Data is published through a repository (institutional, thematic or general-purpose).
- Unique and permanent identifiers (e.g., DOI).
- Data is described with rich metadata.
- Clearly named and organized files and folders.

Sources:

[Données de la recherche
– SciencesPo](#)
[Open Science – Institut
Pasteur](#)

A
ccessible



- Standard communication protocol to access data (user interface, application interface).
- Well defined data access conditions (including authentication and authorization, where necessary).
- Data is "as open as possible, as closed as necessary".
- Metadata is always accessible, even for restricted data, to indicate the existence of the data.

I
nteroperable



- Standard metadata schema and vocabularies.
- Known, open, documented, shared formats.
- The source code of the software necessary to read, process, analyze the data is available.
- Linked to other resources (other data, publication, etc.)

R
eusable



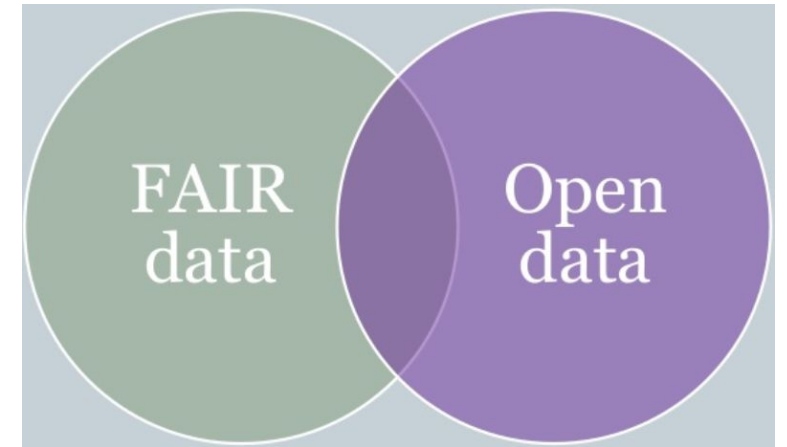
- Data has clear license terms for sharing and reusing.
- Compliance to laws and regulations (e.g., GDPR, French "loi numérique", Intellectual Property Law).
- Documentation to describe the data in detail, contextualize it, make it understandable (e.g., data paper).

FAIR is not equal to Open!

The 'A' in FAIR stands for 'Accessible under well-defined conditions'.













There may be legitimate reasons to shield data and services generated with public funding from public access.






These include personal privacy, national security, and competitiveness.



Data (and metadata) licenses

Data 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.

The most used licenses are the **Creative Commons copyright licenses**

<https://creativecommons.org/>

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

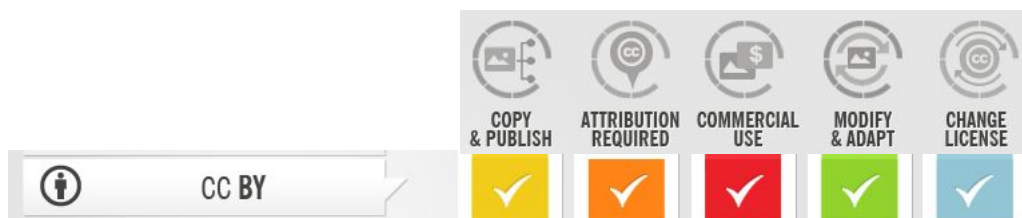
Data licenses

The official French data license is the
Licence Ouverte / Open Licence



LICENCE OUVERTE
OPEN LICENCE

Compatible with CC **BY** :



The “Reuser” is free to reuse the “Information”:

- To reproduce it, copy it.
- To adapt, modify, retrieve and transform it in order to create “derived information”, products and services.
- To share, disseminate, redistribute, publish and transmit it.
- To exploit it for commercial purposes, e.g., by combining it with other information, or by including it in his/her own product or application.

Subject to:

An acknowledgement of the authorship of the “Information”: its source (at least, the name of the “Grantor”) and the date of the most recent update of the reused “Information”.

[...]

“Derived information”: any new data or information created directly from the “Information” or from a combination of the “Information” and other data or information not subject to this licence.

Source: www.etalab.gouv.fr/licence-ouverte-open-licence/

Data licenses available on IPGP-RC

The only licenses legally accepted in France for the reuse of public data (<https://www.data.gouv.fr/fr/pages/legal/licences/>)



LICENCE OUVERTE
OPEN LICENCE

Licence Ouverte / Open Licence 2.0

(Attribution – Commercial use – Derivative Work: *can change license*)



Open Knowledge
Foundation

Open Data Commons Open Database License (ODbL) v1.0

(Attribution – Commercial use – Derivative Work: *share alike*)

But also CC BY 4.0 (fully compatible with the Open Licence and more widespread internationally)



CC BY 4.0

(Attribution – Commercial use – Derivative Work: *can change license*)

Data licenses available on IPGP-RC



LICENCE OUVERTE
OPEN LICENCE

Licence Ouverte / Open Licence 2.0

(Attribution – Commercial use – Derivative Work: *can change license*)



Open Knowledge
Foundation

Open Data Commons Open Database License (ODbL) v1.0

(Attribution – Commercial use – Derivative Work: *share alike*)



CC BY 4.0

(Attribution – Commercial use – Derivative Work: *can change license*)

CUSTOM
TERMS

Custom Terms

e.g., for closed data (embargo, sensitive data, etc...)

Open access to FAIR Data

“As open as possible, as closed as necessary,,

European Commission, Open access & Data management, [Horizon 2020 Online Manual](#)

Some reasons for not opening research data:

- Incompatibility with the need for confidentiality in connection with security issues
- Incompatibility with rules on protecting personal data
- Data embargoes (time limited)

Metadata license on IPGP-RC

The Dataset metadata on IPGP Research Collection is distributed under the CC0 Public Domain Dedication license



<https://creativecommons.org/publicdomain/zero/1.0/deed.en>

No Copyright

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.

Why no copyright for metadata?

To make it easy for other web portals and/or applications to index IPGP Research Collection datasets ("harvesting") without incurring into legal problems → **Interoperability** (the "I" in "FAIR" !)

Institut de physique du globe de Paris, IPGP Data Center

INSTITUT
de PHYSIQUE
du GLOBE RESEARCH COLLECTION
de PARIS

Make your data FAIR with

The IPGP Research Collection

IPGP 
Centre de données



The IPGP Research Collection

The institutional repository of the Institut de physique du globe de Paris (IPGP), for sharing and preserving **Research and Observatory data** produced, collected and curated at IPGP.

<https://research-collection.ipgp.fr>

- Launched on November 2022
- Hosted and operated by the IPGP Data Center
- Official repository + a Sandbox for training

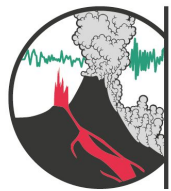


IPGP-RC in the IPGP data landscape



Centre de données

<https://datacenter.ipgp.fr>



<https://ws.ipgp.fr>

<http://volobsis.ipgp.fr>

<http://geoscope.ipgp.fr>

<https://www.seis-insight.eu/science>

- Continuous time series (seismology, GNSS)
- Real-time earthquake catalogs
- Volcano reports



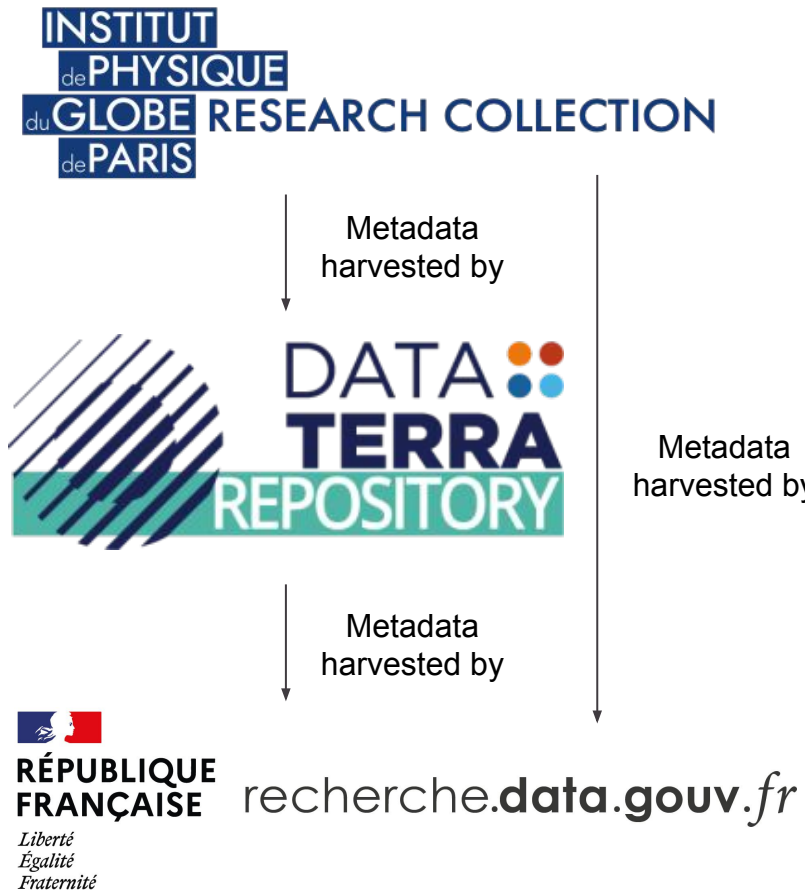
<https://research-collection.ipgp.fr>

- Scientific publication datasets
- Research project datasets
- Observatory data products



Magnetic
Observatories

IPGP-RC in the national data landscape



The IPGP Institutional research and observatory data repository

<https://research-collection.ipgp.fr>

- Research and observatory data from IPGP

National thematic research data repository (*not yet open*)

<https://www.data-terra.org/en/data-terra-repository/>

- Earth system and environment scientific communities
- Geo-referenced data

National multidisciplinary research data repository

<https://recherche.data.gouv.fr/en>

- Research data for which no thematic or institutional solution exists

IPGP Research Collection and FAIR

F
indable



- A single web portal for IPGP Research and Observatory data
- A DOI (Digital Object Identifier) for each Dataset
 - Datasets can be cited in the same way as scientific publications
- Searchable metadata (using the web user interface or the web application interface)
- *Metadata will be harvested by other national and international data portals (to be realized)*

A
ccessible



- Interactive web user interface + scriptable web application interface
 - Based a worldwide used open source research data repository software ([Dataverse](#))
- Metadata always accessible (CC0 license) even when data is closed

I
nteroperable



- Standard metadata formats (Dublin Core, DDI, DataCite, etc.)
- Metadata under public domain (CC0 license)
- Standard harvesting protocol (OAI-PMH)

R
eusable



- Mandatory choice of a license (French licenses, CC-BY, custom terms)
- The web user interface requires users to describe the Dataset and strongly encourages adding rich metadata (e.g., keywords, related publications, geospatial information)

IPGP Research Collection: what?

From the [Terms of Use](#):

The Resources that can be deposited on IPGP Research Collection are:

- Documents associated with the activity of the IPGP **Observatories**.
- Documents associated with a **research project** funded by a French (e.g., ANR, CNES, CNRS, CPR, PIA), European (e.g., ERC) or foreign (e.g., NSF) research project in which IPGP participates.
- Documents associated with a **scientific publication** in a peer-reviewed journal of which the Contributor is a co-author.



Documents:

texts, images, videos, measurement data, calculation results

IPGP Research Collection: who?


From the [Terms of Use](#):

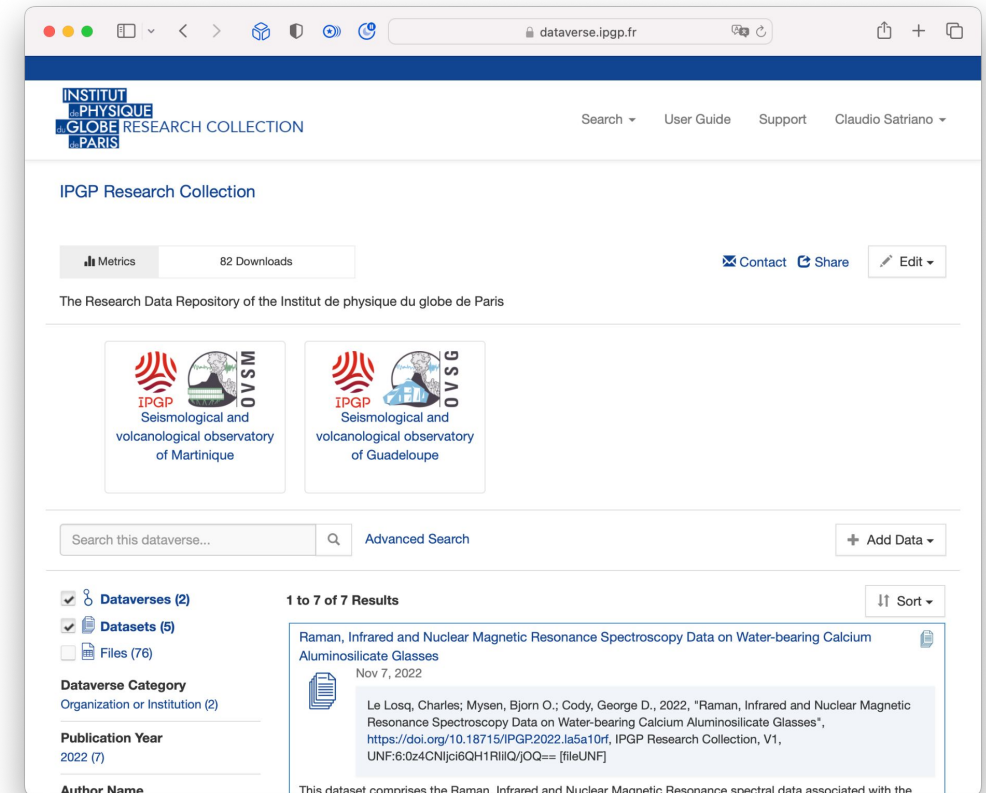
IPGP Research Collection requires the creation of a user account to upload Resources as a Contributor.

The following are entitled to create an account on IPGP Research Collection:

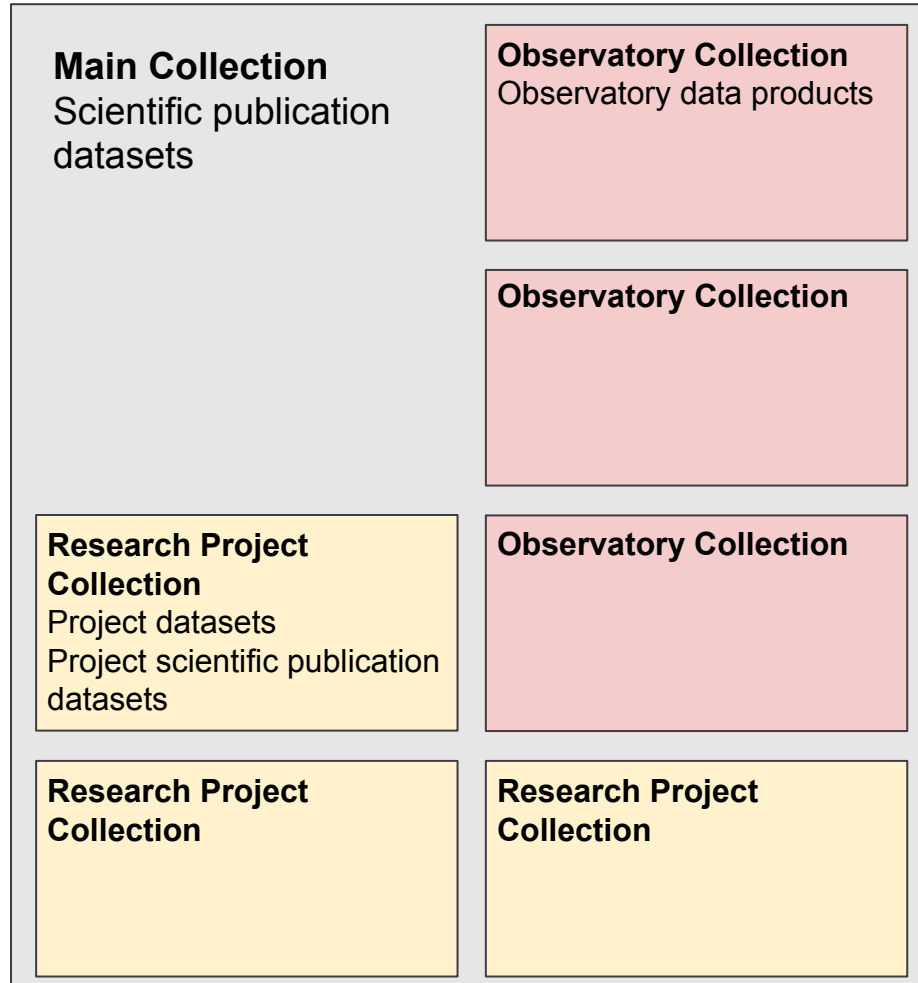
- The agents (permanent or contractual) of the Institut de physique du globe de Paris (IPGP).
- The agents (permanent or contractual) of Université Paris Cité assigned to IPGP.
- The agents (permanent or contractual) of CNRS, IGN and Université de La Réunion assigned to the UMR 7154 (joint research unit) or the UAR 3454 (support and research unit).
- Individuals with "Associate Researcher" status at IPGP.
- The doctoral students at the Doctoral School 560 STEP'UP attached to IPGP.

IPGP Research Collection: how?

- Accessible via: <https://research-collection.ipgp.fr>
 - The data repository
 - A sandbox for training
- Browsing, searching and downloading Datasets does not require a user account
- Uploading a Dataset requires:
 - Choosing the collection for the dataset (main collection, observatory collection or project collection)
 - Creating a user account (ORCID ID required)
 - Asking for authorization to upload Datasets by clicking on the Contact icon of the chosen collection  **Contact**
- Creating a new Collection requires:
 - Having a user account
 - Asking for authorization to create a new Collection by clicking on the Contact icon in the main collection.



IPGP Research Collection: structure



- Each Collection has its contact address (accessible via the "Contact" icon)
- General user support is available from rc-support@ipgp.fr
- Example Observatory Collection: <https://dataverse.ipgp.fr/dataverse/OVSM>

Liability

From the [Terms of Use](#):

The Contributor guarantees that the Resources are disseminatable and that they comply with the applicable legal provisions in force and in particular the provisions relating to personal data, sensitive data, the protection of privacy and intellectual property rights.

The Institut de physique du globe de Paris reserves the right, without warning, in its own judgment and without incurring any liability, to change and/or remove content that is inappropriate or unrelated to the IPGP (see section "Authorized Resources") and to limit or remove access to a Contributor if it considers that its use of IPGP Research Collection violates these Terms, applicable laws in force or the rights of third parties.

Resources

- [Getting Started with IPGP Research Collection](#)
- [National Plan for Open Science](#)
- [The FAIR principles](#)
- [Research Data Guide – SciencesPo](#)
- [List of the licences that can be legally used in France](#) (in French)
- [Recherche Data Gouv](#)
- [Data Terra Repository](#)
- [Dataverse User Guide](#)

Make your data FAIR with IPGP Research Collection