



Mandatory Open Science Practices: Tips for implementation

Webinar „How to address Open Science in EU projects”, 9th June 2022

Barbara Sánchez, TU Wien

Agenda

- Open Science in EU proposal forms
- Tips: Put Open Science into practice
- Recap



Required at proposal stage

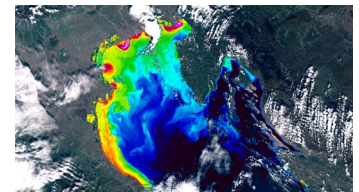
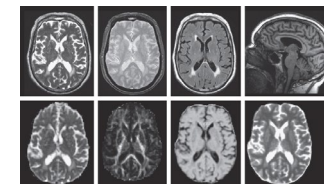
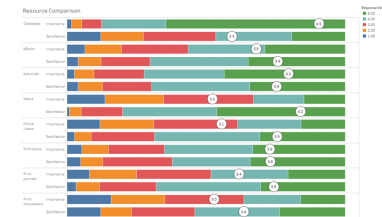
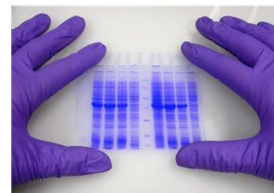
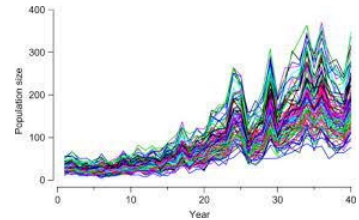
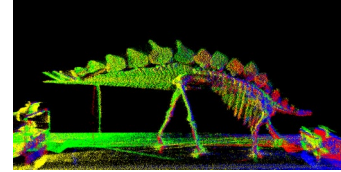
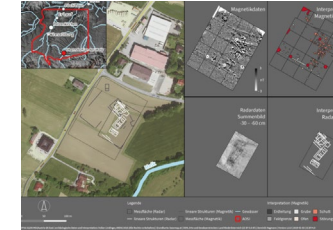
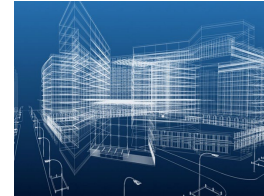


- How appropriate open science practices are implemented as an integral part of the proposed methodology. Show how the choice of practices and their implementation are adapted to the nature of your work, in a way that will increase the chances of the project delivering on its objectives. If you believe that none of these practices are appropriate for your project, please provide a justification here.
 - ⚠ *Open science is an approach based on open cooperative work and systematic sharing of knowledge and tools as early and widely as possible in the process. Open science practices include early and open sharing of research (for example through preregistration, registered reports, pre-prints, or crowd-sourcing); research output management; measures to ensure reproducibility of research outputs; providing open access to research outputs (such as publications, data, software, models, algorithms, and workflows); participation in open peer-review; and involving all relevant knowledge actors including citizens, civil society and end users in the co-creation of R&I agendas and contents (such as citizen science).*
- **Research data management and management of other research outputs:** Applicants generating/collecting data and/or other research outputs (except for publications) during the project must provide maximum 1/2 page on how the data/ research outputs will be managed in line with the FAIR principles (Findable, Accessible, Interoperable, Reusable).
 - ⚠ *For guidance on open science practices and research data management, please refer to the relevant section of the [HE Programme Guide](#) on the Funding & Tenders Portal.*

- Horizon Europe grant application form, Part B, Section „Excellence“, subtopic „Methodology“
- https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/temp-form/af/af_he-ria-ia-stage-1_en.pdf

Data – natural component of research activities

- Instrument measurements
- Experimental observations
- Still images, video and audio
- Text documents, spreadsheets, databases
- Quantitative data (e.g. survey data)
- Survey results & interview transcripts
- Simulation data, models & software
- Slides, artefacts, specimens, samples
- Questionnaires
- Code
- Sketches, diaries, lab notebooks ...
- Data from public registries
- Data from companies
- ...



1/2-1 page of Research Data Management

Describe:

- **Types of data/research outputs** (e.g. experimental, observational, images, text, numerical) and their estimated size; if applicable, combination with, and provenance of, existing data.
- **Findability of data/research outputs:** Types of persistent and unique identifiers (e.g. digital object identifiers) and trusted repositories that will be used.
- **Accessibility of data/research outputs:** IPR considerations and timeline for open access (if open access not provided, explain why); provisions for access to restricted data for verification purposes.
- **Interoperability of data/research outputs:** Standards, formats and vocabularies for data and metadata.
- **Reusability of data/research outputs:** Licenses for data sharing and re-use (e.g. Creative Commons, Open Data Commons); availability of tools/software/models for data generation and validation/interpretation /re-use.
- **Curation and storage/preservation costs;** person/team responsible for data management and quality assurance.

EU Grant: Data Management Template (H2020-01-05-00-2021)
 The Horizon Europe Model Grant Agreement requires that a data management plan ("DMP") is established and regularly updated. The use of this template is recommended for Horizon Europe beneficiaries. In completing the sections of the template the requirements for research data management of Horizon Europe as described in article 17 and analysed in the Annexated Grant Agreement, article 17, must be addressed.

1. Data Summary
 Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.
 What types and formats of data will the project generate or re-use?
 What is the purpose of the data generation or re-use and its relation to the objectives of the project?
 What is the expected size of the data that you intend to generate or re-use?
 What is the origin/provenance of the data, either generated or re-used?
 To whom might your data be useful ("data utility"), outside your project?

2. FAIR data
2.1. Making data findable, including provisions for metadata
 Will data be identified by a persistent identifier?
 Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.
 Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?
 Will metadata be offered in such a way that it can be harvested and indexed?

2.2. Making data accessible
 Repository:
 Will the data be deposited in a trusted repository?
 Have you explored appropriate arrangements with the identified repository where your data will be deposited?
 Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?
 Data:
 Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints, as per the Grant Agreement.
 If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.
 Will the data be accessible through a free and standardised access protocol?
 If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?
 How will the identity of the person accessing the data be ascertained?
 Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?
 Metadata:
 Will metadata be made openly available and licensed under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?

Version 3.0
05 May 2021

HORIZON EUROPE TEMPLATE

CORE REQUIREMENTS FOR DATA MANAGEMENT PLANS

When developing solid data management plans, researchers are required to deal with the following topics and answer the following questions:

- 1. Data description and collection or re-use of existing data**
 - a. How will new data be collected or produced and/or how will existing data be re-used?
 - b. What data (for example the kinds, formats, and volumes) will be collected or produced?
- 2. Documentation and data quality**
 - a. What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany data?
 - b. What data quality control measures will be used?
- 3. Storage and backup during the research process**
 - a. How will data and metadata be stored and backed up during the research process?
 - b. How will data security and protection of sensitive data be taken care of during the research?
- 4. Legal and ethical requirements, codes of conduct**
 - a. If personal data are processed, how will compliance with legislation on personal data and on data security be ensured?
 - b. How will other legal issues, such as intellectual property rights and ownership, be managed? What legislation is applicable?
 - c. How will possible ethical issues be taken into account, and codes of conduct followed?
- 5. Data sharing and long-term preservation**
 - a. How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?
 - b. How will data for preservation be selected, and where will data be preserved long-term (for example a data repository or archive)?
 - c. What methods or software tools will be needed to access and use the data?
 - d. How will the application of a unique and persistent identifier (such as a Digital Object Identifier (DOI)) to each data set be ensured?
- 6. Data management responsibilities and resources**
 - a. Who (for example role, position, and institution) will be responsible for data management (i.e. the data steward)?
 - b. What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?

PRACTICAL GUIDE TO THE INTERNATIONAL ALIGNMENT OF RESEARCH DATA MANAGEMENT 9

Science Europe Practical Guide & Template: <https://doi.org/10.5281/zenodo.4915862>



Required at proposal stage



- How appropriate open science practices are implemented as an integral part of the proposed methodology. Show how the choice of practices and their implementation are adapted to the nature of your work, in a way that will increase the chances of the project delivering on its objectives. If you believe that none of these practices are appropriate for your project, please provide a justification here.

⚠ *Open science is an approach based on open cooperative work and systematic sharing of knowledge and tools as early and widely as possible in the process. Open science practices include early and open sharing of research (for example through preregistration, registered reports, pre-prints, or crowd-sourcing); research output management; measures to ensure reproducibility of research outputs; providing open access to research outputs (such as publications, data, software, models, algorithms, and workflows); participation in open peer-review; and involving all relevant knowledge actors including citizens, civil society and end users in the co-creation of R&I agendas and contents (such as citizen science).*

- **Research data management and management of other research outputs:** Applicants generating/collecting data and/or other research outputs (except for publications) during the project must provide maximum 1/2 page on how the data/ research outputs will be managed in line with the FAIR principles (Findable, Accessible, Interoperable, Reusable).

⚠ *For guidance on open science practices and research data management, please refer to the relevant section of the [HE Programme Guide](#) on the Funding & Tenders Portal.*

- Horizon Europe grant application form, Part B, Section „Excellence“, subtopic „Methodology“
- https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/temp-form/af/af_he-ria-ia-stage-1_en.pdf

Mandatory Open Science practices

Mandatory Open Science practices for all beneficiaries per the Grant Agreement (GA)

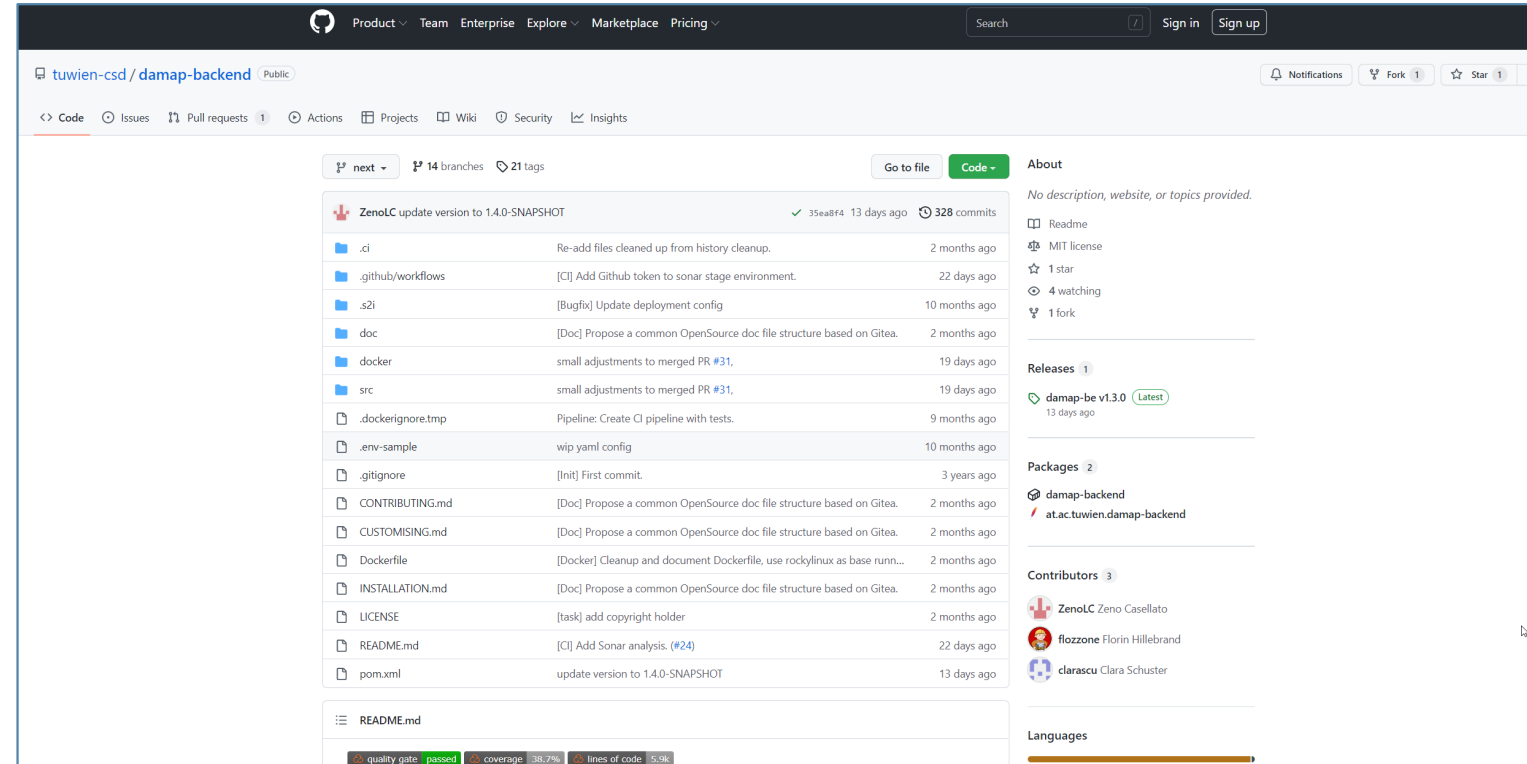
- **Open Access to scientific publications** (required by the GA)
- **FAIR principles applied in responsible management of research data** and open access to research data under the **principle of “as open as possible, as closed as necessary”**. Data Management Plans compulsory for all grants, by M6
- **Information about research outputs, tools and instruments** needed to validate conclusions of scientific publications or to validate and re-use research data
- **Digital or physical access to results** needed to validate conclusions of scientific publications (unless exceptions apply)
- **Public emergency** requires immediate open access to all research outputs under open licenses if requested by the granting authority, or, if exceptions apply, access under fair and reasonable conditions to legal entities needing research outputs to address public emergency



TECHNISCHE
UNIVERSITÄT
WIEN

Tips: Put Open Science into practice

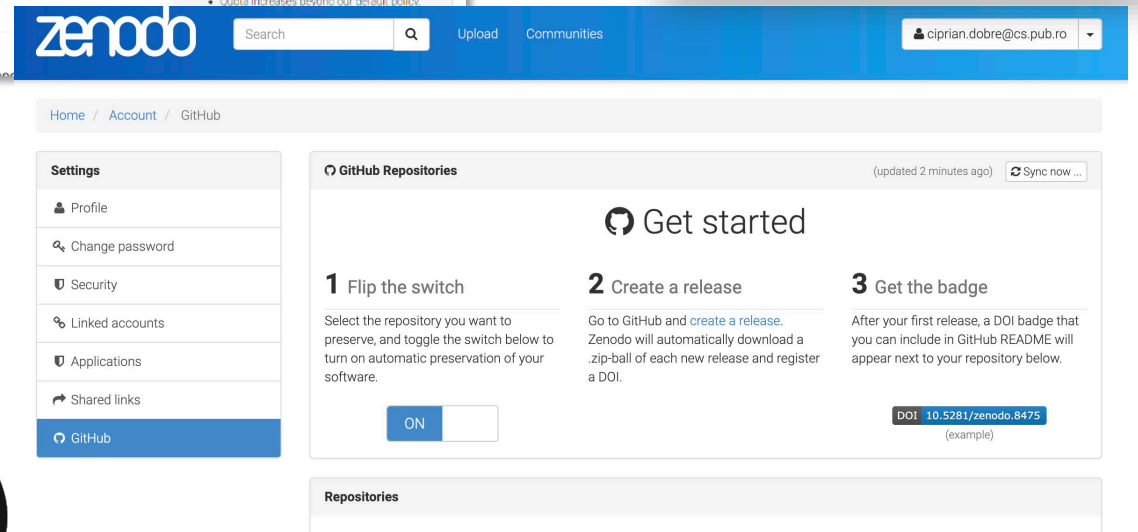
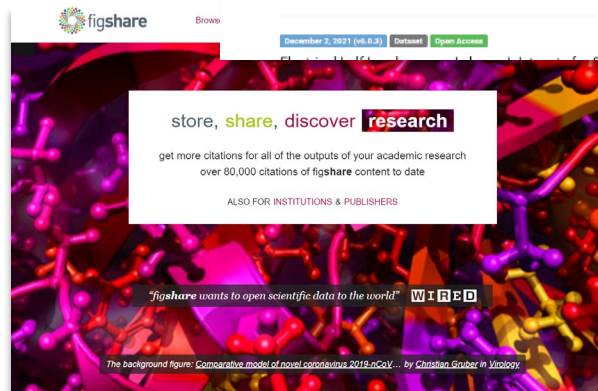
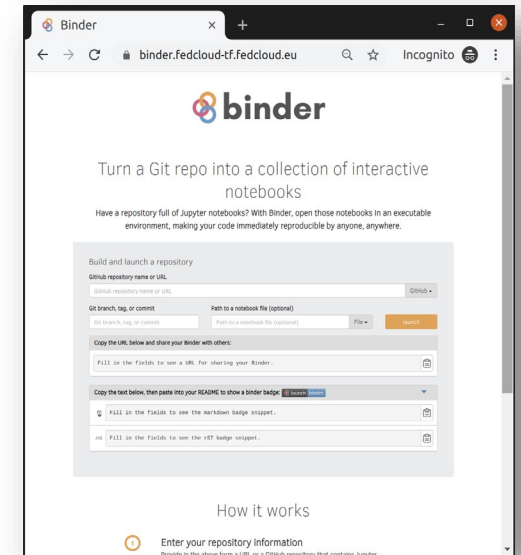
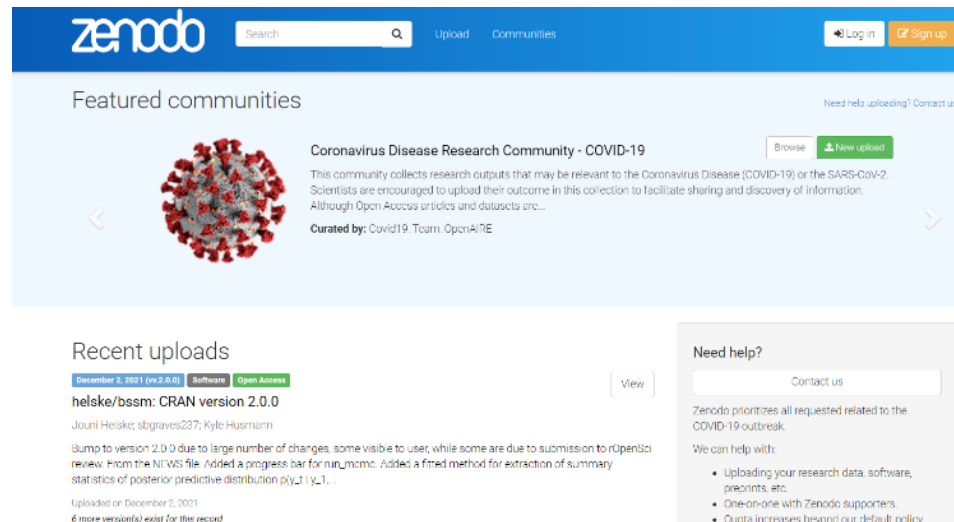
Use collaborative platforms during research process



Use repositories for sharing and publishing



re3data: Registry, over 2.000 data repositories
<https://www.re3data.org/>



What's on a dataset landing page

January 19, 2021 | Version 1.0

European Sentinel-1 Forest Type and Tree Cover Density Maps

Description of the dataset

This dataset was generated by the French components of forestry and geo-information. European Sentinel-1 forest type and tree cover density maps represent five consecutive decade forest layers based on Sentinel-1 C-Band Synthetic Aperture Radar (SAR) backscatter data. For the year 2017, these cover the majority of European continent with 10 m and 300 m temporal resolution, type and tree cover density, respectively. The maps were derived using the method described in <https://www.sciencedirect.com/science/article/pii/S0034425720379785>.

The forest type map shows the dominant forest type class (coniferous, broadleaf). The tree cover density map shows the percentage of forest canopy cover within the 100 m pixel.

Users are referred to our peer-reviewed article <https://doi.org/10.1016/j.isprsar.2020.10.005> for details and primary 2D remote access Europe.

Dataset Record

The forest type and tree cover density maps are sampled at 10 m and 100 m pixel spacing respectively, georeferenced to the equator and central meridian square (10° of 200m extent) 7171°E with this setup, the forest maps consist of 420 files over the European continent, with data volumes of 1.2 GB per one 100 m pixel.

The files (Erdos) are 1.2GB compressed GeoTIFF having 15-bit integer values, with biggest metadata on encoding and georeferencing. Contact us by web (www.geogis.tuwien.ac.at) or via email (geogis@tuwien.ac.at) and geogis@tuwien.ac.at.

In this repository, we provide each forest map as files, whereas two zipped datasets collections are available for download below.

Code Availability

For the usage of the EpiGrid we provide data and tools via the python package available on GitHub at <https://github.com/epi-grid/epi-grid>. More details on the git reference can be found in <https://www.sciencedirect.com/science/article/pii/S0034425720379785>.

Acknowledgements

The computational results presented have been achieved using the Vienna Scientific Cluster (VSC).

Access regulation

European Sentinel-1 Forest Type and Tree Cover Density Maps

Dataset Record

Code Availability

Acknowledgements

Versions

Version 1.0 | Jan 19, 2021

View all 1 versions

Digital Object Identifier

DOI: 10.48436/tkkfs-11b75

Export

JSON

Cite As

Dostalova, Alena, Cao, Senmao & Wagner, Wolfgang. (2021), European Sentinel-1 Forest Type and Tree Cover Density Maps (Version 1.0) [Dataset]. TU Data. <https://doi.org/10.48436/tkkfs-11b75>

Persistent identifier (DOI):
10.48436/tkkfs-11b75

Citation snippet

License: CC-BY-SA 4.0

Details

Licenses cc-by-sa-4.0 | Creative Commons Attribution-ShareAlike 4.0 International

Resource type Dataset

Formats application/x-geoiff

Related works

Isreferencedby <https://doi.org/10.3390/rs13030337> (doi)

Issupplementto <https://doi.org/10.5281/zenodo.35115933> (doi)
<https://github.com/TUW-GEOIT/epiGrid> (url)

references <https://doi.org/10.1089/0131161.2018.1479786> (doi)
<https://doi.org/10.1016/j.isprsar.2020.10.005> (doi)

Upload information Created: January 19, 2021 | Modified: March 29, 2022

Reference to related works

Files for download

ForestType.zip

- 2017_FOREST_CLASSES_BUSION_E029009T1.tif | 1.0 MB
- 2017_FOREST_CLASSES_BUSION_E030010T1.tif | 670.2 KB
- 2017_FOREST_CLASSES_BUSION_E030011T1.tif | 1.0 MB
- 2017_FOREST_CLASSES_BUSION_E030012T1.tif | 7.5 MB
- 2017_FOREST_CLASSES_BUSION_E030013T1.tif | 8.0 MB
- 2017_FOREST_CLASSES_BUSION_E030014T1.tif | 5.5 MB

Files for download

Name	Size	Preview	Download
ForestType.zip	3.1 GB		
treecoverdensity.zip	376.1 MB		

Assign usage licenses

- Once intellectual creative output has been invested, data is licensable because **copyright** applies



- Non-licensed material (if it is protected by copyright) is **not** freely available

- Only **owners** (= usually data creators) are entitled to distribute licenses for the use of data
- Several creators → all creators are in principle jointly entitled to make decisions

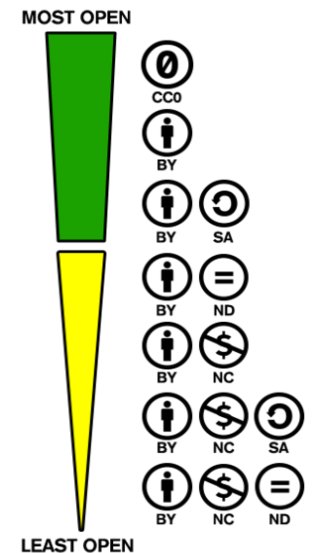
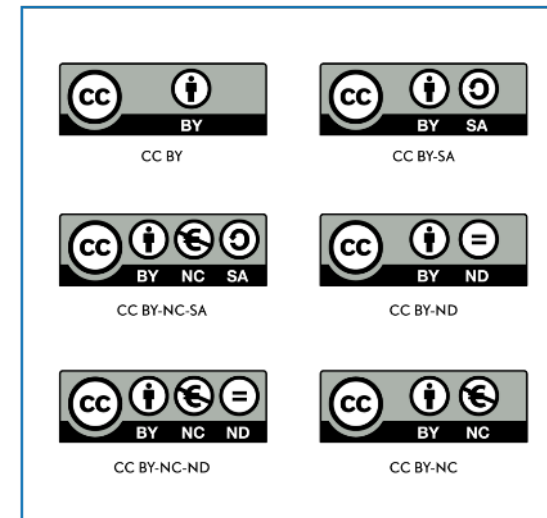


Creative Commons Licenses

- Recommendation for data: **Creative Commons**
 - Standardized license text → Reduces effort and creates legal certainty
 - No transfer of exclusive exploitation rights → Researchers can continue to work freely
 - Internationally distributed / compatible with a wide range of legal systems compatible → helps avoid problems in an international context



<https://creativecommons.org/>



Assign licenses for code

- Tools for selecting a software license:
 - <https://ufal.github.io/public-license-selector/>
 - <https://choosealicense.com/>

Choose an open source license

An open source license protects contributors and users. Businesses and savvy developers won't touch a project without this protection.

Which of the following best describes your situation?



I need to work in a community.

Use the **license preferred by the community** you're contributing to or depending on. Your project will fit right in.

If you have a dependency that doesn't have a license, ask its maintainers to **add a license**.



I want it simple and permissive.

The **MIT License** is short and to the point. It lets people do almost anything they want with your project, like making and distributing closed source versions.

Babel, **.NET**, and **Rails** use the MIT License.



I care about sharing improvements.

The **GNU GPLv3** also lets people do almost anything they want with your project, *except* distributing closed source versions.

Ansible, **Bash**, and **GIMP** use the GNU GPLv3.

What if none of these work for me?

My project isn't software.

There are licenses for that.

I want more choices.

More licenses are available.

I don't want to choose a license.

Here's what happens if you don't.

Choose a License


Answer the questions or use the search to find the license you want

Is your code used directly as an executable or are you licensing a library (your code will be linked)?

Search for a license...



GNU General Public License 2 or later (GPL-2.0)

You may copy, distribute and modify the software as long as you track changes/dates of in source files and keep all modifications under GPL. You can distribute your application using a GPL library commercially, but you must also disclose the source code.

Publicly Available 



GNU General Public License 3 (GPL-3.0)

You may copy, distribute and modify the software as long as you track changes/dates of in source files and keep modifications under GPL. You can distribute your application using a GPL library commercially, but you must also provide the source code. GPL 3 tries to close some loopholes in GPL 2.

Publicly Available  

Affero General Public License 3 (AGPL-3.0)

The AGPL license differs from the other GNU licenses in that it was built for network software. You can distribute modified versions if you keep track of the changes and the date you made them. As per usual with GNU licenses, you must license derivatives under AGPL. It provides the same restrictions and freedoms as the GPLv3 but with an additional clause which makes it so that source code must be distributed along with web publication. Since web sites and services are never distributed in the traditional sense, the AGPL is the GPL of the web.

Publicly Available  

Assign the FAIR principles

CLUSTER FORSCHUNGSDATEN

FAIR DATA AUSTRIA

LET'S MAKE OUR DATA FAIR!

19.10.2021

HOSTED BY TU WIEN,
CENTER FOR RESEARCH DATA MANAGEMENT

TECHNISCHE UNIVERSITÄT WIEN

Tomasz Miksa. Let's make our data FAIR. DOI: 10.25365/phaidra.306. CC-BY4.0

Machine-actionability - example

Not machine-actionable

Day	Lowest Temperature (°C)
Saturday, 13 November 2010	-2
Sunday, 14 November 2010	-4
Monday, 15 November 2010	-7

Machine-actionable

Day	Lowest Temperature (°C)
Saturday, 13 November 2010	-2
Sunday, 14 November 2010	-4
Monday, 15 November 2010	-7

CSV

Interoperable – simplified examples

Yes

- XML following known XSD Schema
- MP3 for audio recordings

No

- Custom XML without any documentation
- M4P (Apple) for audio recordings

Findable – simplified examples

Yes

Data repository

No

Personal website

FAIR explained with many practical **examples**:

- **Video** recording of the presentation Let's make our data FAIR. Handle: [11353/10.1246920](https://hdl.handle.net/11353/10.1246920).
- **Slides**: Tomasz Miksa (19.10.2021). Let's make our data FAIR. DOI: [10.25365/phaidra.306](https://doi.org/10.25365/phaidra.306). CC BY 4.0.

Control access to your data

- FAIR ≠ Open
- FAIR through access control

such a large and complex data set. Last but not of the colleagues from the TU Wien Center for R

Sorry, the files are restricted!

Details

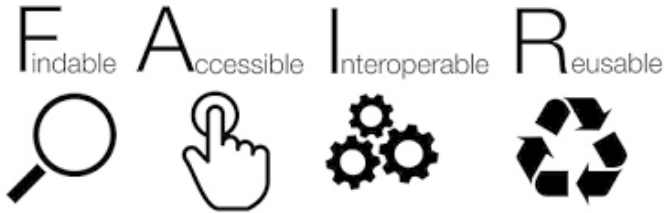
Licenses [cc-by-4.0](#)



TECHNISCHE
UNIVERSITÄT
WIEN

Recap

3 concepts – overlapping, not the same



Open Data



Research Data Management

- is about metadata and findability
- is about standards and interoperability
- is about machine-to-machine communication
- sensitive data (privacy, IPR, ...)
- data visiting instead of data sharing
- FAIR data is **not equivalent to open data** (publicly available to everyone to access and reuse). Data can, and should be FAIR even when access is restricted

- is available on the internet
- can be freely used, modified and shared by anyone for any purpose
- research data should be **open as a default**, unless there are legitimate reasons for keeping them closed
- as a general rule, open access to other research outputs such as software, models, algorithms, workflows, protocols, simulations, electronic notebooks and others is not required but strongly recommended

- refers to the entire life cycle of research data
- from planning, generation, analysis, processing and backup, to documentation, archiving and publication, and finally, reuse by third parties
- enables FAIR and Open
- RDM is mandatory in Horizon Europe for projects generating or reusing data
- DMPs are useful planning instruments (and a Deliverable in M6)



Good to know

- Failure to sufficiently address open science practices will result in a lower evaluation score.
- A clear explanation of how applicants will adopt recommended practices will result in a higher evaluation score.
- If proposers believe that none of the open science practices (mandatory or recommended) apply to their project, then they have to provide a justification.
- Under '**capacity of participants and consortium as a whole**', proposers should describe how the consortium brings together the necessary disciplinary and inter-disciplinary knowledge. Proposers should show how this includes expertise and/or track record in open science practices, relevant to what is planned for the project.
- Finally, in **part A** of their proposals, proposers are asked to list up to five relevant publications, widely used datasets or other achievements of consortium members that they consider significant for the action proposed. Open access is expected for publications, in particular journal articles, while datasets are expected to be FAIR and 'as open as possible, as closed as necessary'.

MORE INFO IN JJTHE PROGRAMME GUIDE: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide_horizon_en.pdf



TECHNISCHE
UNIVERSITÄT
WIEN

Contact

Barbara Sánchez

Center for Research Data Management
Favoritenstraße 16 (top floor), 1040 Vienna

barbara.sanchez@tuwien.ac.at
research.data@tuwien.ac.at

www.tuwien.at/researchdata