

RESEARCH DATA MANAGEMENT AT THE UPC

Sharing research data is one of the mandates of open science, which aims to increase cooperation and the transparency of research in all its stages and to increase its visibility and impact. The result is more robust and replicable research that allows for increased returns on public funds dedicated to research. At the same time, it favours cross-disciplinarity and the opening up of new fields of research. In some cases, access to or use of data may be restricted to safeguard individuals' privacy, protect confidentiality or allow the commercial exploitation of research results. But it is always necessary to properly manage research data following FAIR principles, and this needs to be planned in advance.

1. What is meant by research data

Research data are the numerical, textual, image or sound records that are used as primary sources for scientific research and are commonly accepted as validating the conclusions and results of the research by the scientific community. Research data may be experimental data, observational data, operational data, third party data, public sector data, monitoring data, processed data or reused data. Source: [Codata](#)

2. How these data should be managed. Data management plans

A data management plan (DMP) is a living document that must be drawn up at the start of the research and updated during the course of the project. It includes aspects such as:

- The types of data to be collected, their format, the name of files and the management of different versions.
- Where the data will be stored and how they will be shared.
- Data security measures.
- Who will be responsible.
- The criteria for their selection and preservation.
- How the application of FAIR principles will be guaranteed.
- Ethical and intellectual property aspects.

In the case of funded projects, it is very possible that you will already have to outline what your DMP will look when you apply, although it will not have to be fully drafted until the beginning of the project, as one more deliverable. It is important that the applications state that the data will be well managed and FAIR principles followed.

There are tools and applications available that help give the DMP all the necessary sections, such as [CORA-eiNa DMP](#). UPC libraries provide support for preparing data management plans.

3. FAIR principles. What they are and how you can fulfil them

FAIR, an acronym for **F**indable, **A**ccessible, **I**nteroperable, **R**eusable, data principles are a set of characteristics that data objects (metadata and data) must have in order to be recoverable and reusable by humans and machines. Source: Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Sci Data* 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>

To comply with the FAIR principles the data must be:

Findable

The dataset's metadata are abundant and varied enough to describe the data and have a unique and persistent identifier (DOI, handle, etc.) that makes the data and metadata easy for others to find.

Accessible

The data are recoverable by people and machines using a standardised communication protocol, with authentication and authorisation when necessary. Metadata must be accessible even when the data are no longer available.

Interoperable

The data and associated metadata use a "formal, accessible, shared and widely applicable language for knowledge representation"¹. This involves using community-accepted languages, formats and vocabularies in the data and metadata. Metadata should refer to and describe relationships with other data, metadata and information using identifiers.

Reusable

The associated metadata provide accurate and relevant information and the data include a clear data licence and detailed information on their provenance. Metadata standards and data specific to the subject area are used.

FAIR principles must be applied to the dataset. All previous data must be properly managed to achieve a FAIR dataset, but not all the principles must be fulfilled at all times.

¹ FAIR Principles, 2017. GO FAIR. Online. [Accessed 14 May 2024]. Available from: <https://www.go-fair.org/fair-principles/>

4. Personal data collection and processing

What if my research includes processing personal data, which is understood to be “any operation or set of operations carried out on personal data or sets of personal data, whether by automated procedures or not, such as collection, registration, organization, structuring, conservation, adaptation or modification, extraction, consultation, use, communication by transmission, dissemination or any other form of enabling access, confrontation or interconnection, limitation, deletion or destruction”²?

There are different situations in which you may need to collect personal data during the course of your research project. In these cases, the DMP must specify the measures taken to protect people's rights and freedoms and ensure compliance with both the [General Data Protection Regulation](#) (GDPR) and [Organic Law 3/2018, of 5 December, on personal data protection and the safeguarding of digital rights](#).

Situation 1: Collection of personal data solely for the purposes of project communication and dissemination

In this case, it is not exactly research data that you will have to process according to FAIR criteria, nor will you generate datasets to be published, but you will have to define very clearly what the data will be used for and how they will be stored.

When you collect personal data, you must always ask for the subject's express consent and state the purpose of processing the data, who will be responsible for processing them and how subjects can exercise their rights.

Click on the [link](#) for examples in Catalan, Spanish and English of the information that you must add to your data collection forms.

Choose the purpose of the data processing from among the codes established by the University in the [RAT](#) (the register of data processing activities of the UPC). In this case (Situation 1) you could choose F05.12 Actions to promote research and innovation, for example.

Once the data has been collected, how should I process them?

The personal data collected must be kept in such a way that the right of access, the right to rectification or erasure and the right to restrict processing are guaranteed. Access to the data must be limited to persons who legitimately have to use them and must be protected by a password. The data must only be used for the purpose stated when they were collected.

What do I do after the processing of personal data is completed?

Once the activity for which the personal data was being processed has ended, the data controller is obliged to block the data. Blocking the data consists in identifying and retaining them and adopting technical and organisational measures to prevent them being processed,

² Generalitat de Catalunya (eds.), 2023. Unitat 3. Quan podem tractar les dades personals? - EAPC Wiki, Escola d'Administració Pública de Catalunya. Online. 17 May 2023. [Accessed 14 May 2024]. Available from: <https://formaciooberta.eapc.gencat.cat/contingutsdels cursos/abpd/030/inici.html>

including being displayed, except in cases in which the data must be made available to judges, the courts, the Public Prosecutor's Office, the competent public administrations and particularly the data protection authorities due to a formal requirement for potential liabilities stemming from the processing and only for the limitation period. After this period, the data must be destroyed.

Situation 2: My research includes collecting data in surveys or participation in experiments, but the participants' data does not need to be processed.

Research data should be collected without the need to specify the data that allow persons to be identified. Age groups and other data that allow participants to remain anonymous may be collected. In this case, no special procedure must be followed. The minimum necessary data must always be collected.

If personal data must be collected, they must be anonymised or pseudonymised before they can be processed. Find information and tutorials on anonymisation by clicking on the [link](#). The best practice is to anonymise the data when they are collected. This means that personal data are not processed and the GDPR does not apply.

Informed consent must be requested and recorded. It consists of two parts: an information sheet on the project, the project's purpose and the implications the project might have for the participant, and a consent form. Find more information and templates on the [UPC Ethics Committee website](#).

Once the data have been anonymised, the GDPR no longer applies, but if you still have the original data they must be treated according to the necessary security protocols.

Situation 3: In my research I need to process personal data to achieve the research objectives, but the data do not contain special category data (health status, political opinion, sexual orientation, ethnic origin, etc.) or the data of minors or vulnerable people.

In this case, you must apply for an ethics review of the project by the UPC Ethics Committee before collecting the data. Without the Committee's prior approval, you will not be able to publish the results of your research under any circumstances. Find additional information on the process and the documents to be submitted to the Committee on its [website](#).

Situation 4: In my research I need to process personal data to achieve the research objectives, and the data contain special category data (health status, political opinion, sexual orientation, ethnic origin, etc.) or the data of minors or vulnerable people.

In addition to the project review by the UPC Ethics Committee, in this case security measures must be increased in the storage of the data. Details must be given of all the technical and organisational measures that will be followed to guarantee their safety.

Situation 5: Reuse of personal data collected in previous research projects or publicly available data

In the case of data collected in another project, you must state the method followed to collect the data originally and the informed consent procedure, and you must have the permission of the dataset's owner.

If publicly available data are used, you must provide details of the source and ensure that they can be used for research. The fact that the data are public does not mean that there are no limits on their use. You must ensure that your use complies with the terms and conditions published by the data controller.

If you have any questions, always consult the [data protection officer](#).

Finally, this [decision tree](#) drawn up by the European Commission will help you decide what protocols you must follow in each case.

5. Open-access data publishing

One of the premises of open science is open-access publication of the data that validates the research results. In accordance with the fundamental lemma underlying the FAIR principles, that is, "as open as possible and as closed as necessary"³, publishing the data is not compulsory, but FAIR principles must always be followed. In addition, if the research has received funding it must comply with the obligations established by the funding body regarding open access. Ask your library for advice on the repository best suited to your case. The repositories may be multidisciplinary, such as the CSUC's [CORA.RDR](#) and Zenodo, or specialised.

The latest version of a Creative Commons Attribution International ([CC BY](#)) public licence or a licence that grants equivalent rights is compulsory for Horizon Europe project data.

APPENDICES

- a. Documents for the project proposal stage: [Information to be included in the data management plan](#)
- b. Informed consent form (for personal data): [Form](#) on the Ethics Committee website
- c. Additional [consent forms](#) in the "Templates and forms" section of the Ethics Committee website
- d. [Information on anonymisation and pseudonymisation](#) from the Spanish Data Protection Agency

Document prepared by the UPC research data management work group
Contact: ciencia.oberta@upc.edu
Version 1 (March 2023)

³ European Commission (ed.), [no date]. Open Science. , Research and innovation. Online. [Accessed 14 May 2024]. Available from: https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en
