INFORMATION ON OPEN SCIENCE

Markéta Máliková, Jan Vališ

updated: 2026-06-03

The aim of this overview is to summarize information on the principles of Open Science (OS) and to provide basic guidance for IMC employees on how to make research results and continuous outputs (e.g., data) accessible in the OS regime.

Guidelines for Open Access (OA) publishing are based on Directive No. 12 of 2022. Information on the publication of research data considers the conditions of individual project providers.

Content:

1 What Open Science means 2
1.1 Basic elements of open science: 2
1.2 Advantages and disadvantages of OS 2
1.3 Licenses and copyright 3
1.4 OS requirements in research projects 5
2 Publishing via OA 6
2.1 Stages of manuscripts 6
2.1.1 Forms of OA 6
2.2 Modes of open access 6
2.2.1 Green OA 6
2.2.2 Gold OA 7
2.2.3 Platinum/Diamond OA 7
2.3 Options for financing the costs of publishing OA articles 8
2.4 How to select an open journal 8
3 Research Data Management 9
3.1 FAIR principles 9
3.2 Open data 10
3.2.1 Storing open data 11
3.2.2 Data repositories 12
3.3 Data Management Plan 12
3.4 How to publish open research data 13

1 What Open Science means

OS is a summary of various practices and principles that allow free, immediate, and remote access to the results of publicly funded research.

1.1 Basic elements of open science:

- **open access (OA)** – open electronic access to the results of publicly funded research. The goal is to achieve immediate, free, permanent, and unrestricted access to publications for the end user
- **open data** – open research data
- **other OS tools** – open source (SW), public licenses for works, open peer review, open research notebooks, open educational materials, or citizen science

For the needs of the employees of the Institute of Macromolecular Chemistry of the CAS (hereinafter referred to as IMC), this methodology will focus primarily on the elements of OA and open data.

1.2 Advantages and disadvantages of OS

- **Advantages**:
    - immediate and free access to professional publications
    - greater availability and visibility of scientific work
    - making research visible and increasing its credibility
    - higher citation impact
    - faster exchange of knowledge
    - higher transparency and quality of scientific work
    - efficient use of results, work, and finances
    - an informed public
    - retention of property rights to publications (compared to the publishing regime in journals with a subscription model)
- **Disadvantages:**
    - high financial costs for fees associated with publication
    - existence of predatory journals
    - the necessity of handling copyright and publishing rights for data
    - potential demands on technical infrastructure
    - long-term sustainability of financing

- **Possible solutions:**
    - use of transformative agreements reducing publication costs
    - rigorous choice of journals and licenses
    - national and institutional support

1.3 Licenses and copyright

For publications or data to fulfil the principle of openness, they must be freely accessible, without legal barriers. Therefore, for an open work, the author indicates a so-called public license, which stipulates under what conditions the work can be made publicly available and used. The author retains the copyright to the content of the work.

Most commonly applied, and also directly required/recommended by some project providers, are the public Creative Commons (CC) licenses. Using these licenses, it is possible to grant others the right to download, reuse, print, modify, share, and/or copy the work with citation of the author and source. CC licenses are irrevocable and can be terminated by a breach of the licensing conditions by the user. There are several types of CC licenses, which differ in the levels of rights granted and reserved by the author to the work. Their more detailed breakdown is shown in Figure 1.

Some providers state in the project documentation the so-called CC0 as one of the required/recommended licenses. This license has no basis in the legal system of the Czech Republic, and therefore it cannot be granted in good faith to any copyrighted work. The providers' requirement to publish metadata for datasets in the CC0 regime is made possible by a slight balancing act in the Czech legal system, as the filled-in "form" through which metadata is entered is not considered a copyrighted work, and metadata therefore do not enjoy copyright protection, which is de facto the providers' requirement aiming to enable the extraction of metadata without the obstacles of citation.

A separate chapter is the legal classification of research data, which, especially in the case of chemistry and materials sciences, is very often of an instrumental nature. Such data cannot automatically be considered a copyrighted work according to the Copyright Act (Act No. 121/2000 Coll.), however, under certain circumstances, this data can be protected as a so-called sui generis database. A decision tree can help determine whether a work is protected by the Copyright Act in the Czech Republic.

Regardless of the current legal classification, providing a license for a dataset is an internationally recognized practice that is also required by most financial support providers.

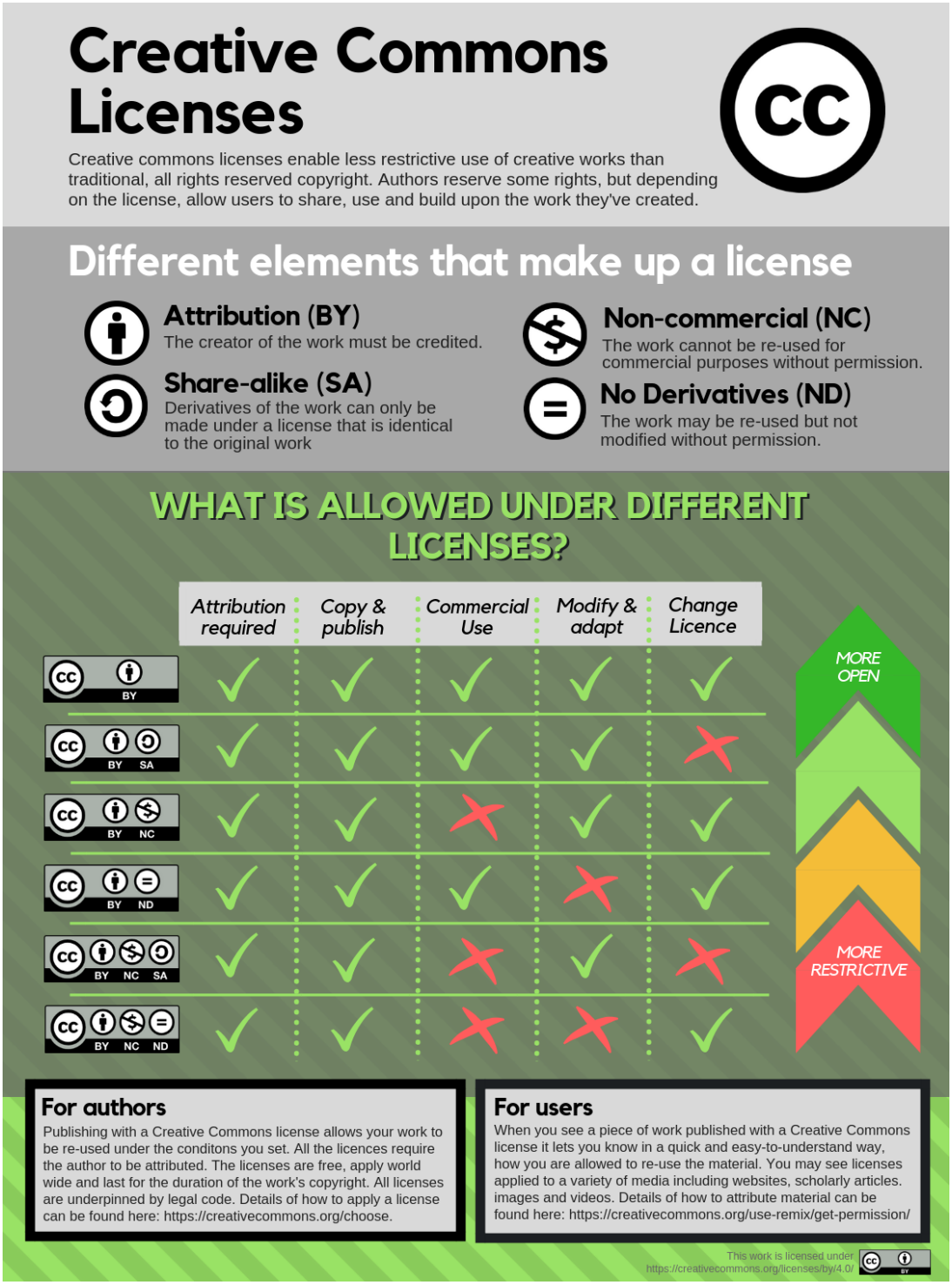

Figure 1 Overview of Creative Commons licenses. Adapted from: McTaggart, David. CC License infographic. In: City, University of London. City Blogs. [online] October 24, 2018. [cit. 2024-06-14].

1.4 OS requirements in research projects

Research funding providers in the EU and the Czech Republic mostly require open access to publications and source data. In a number of countries, legislative measures are gradually being introduced for the implementation of OA principles in science (e.g., in the Czech Republic, Act No. 130/2002 Coll. and No. 328/2025 Coll.). For example, part of the evaluation and outputs of some projects is the openness of publications and data in the project.

Generally, the requirements of project finance providers can be summarized as conditions for making the publication accessible under a public license, storing it in a trusted open-access repository, providing open access to research data, and creating and updating a data management plan. **Specific conditions of individual providers may vary.**

A useful starting point for information regarding OS and provider requirements can be the National Platform for Open Science (Providers – National Platform for Open Science). Detailed information on conditions can be found on the websites of individual grant agencies:

- **Horizon Europe**
- **OP JAC**
- **Program Exceles**
- **MEYS Research and innovation support**
- **Technology Agency of the Czech Republic (TAČR)**
- **Czech Science Foundation (GAČR)**
- **Agency of the Czech Ministry of Health (AZV ČR**)
- **CAS Research support**

2 Publishing via OA

The publication process in the OA regime and the subscription regime basically does not differ for authors – the manuscript is assessed by an editor, subjected to a peer-review process, and accepted for publication. The only fundamental difference is whether, at the end of this process, property rights to the manuscript are transferred to the publisher (the conventional path of publication in the subscription regime) or whether the authors retain property rights alongside copyrights (OA).

2.1 Stages of manuscripts

From the moment of writing until publication, an author's work typically goes through the following stages:

- **pre-print/author original manuscript (AOM)**
  = the original sent to the editorial office without journal typesetting and without peer review revisions
- **post-print/author accepted manuscript (AAM)**
  = the version after the peer review process, but without journal typesetting
- **publisher's version (version of record, VoR)**
  = the version after the peer review process and after the journal's graphic adjustments

2.1.1 Forms of OA

An author can ensure open access to their publications mainly through these paths:

- green OA,
- gold OA or
- platinum/diamond OA.

2.2 Modes of open access

2.2.1 Green OA

In Green OA, the author publishes the full text of the article in a classic subscription-based journal, but simultaneously deposits the full text, either the AOM or AAM, in an open repository.

During this so-called self-archiving, **the author must comply with the terms and conditions of the journal publisher**. However, for it to be recognized by the support provider, the **provider's requirements must also be met**. Attention must be paid to verify whether the article can be made accessible, in what version, in which repository, and with what embargo. A suitable tool for a quick survey of publisher conditions is the **Open policy finder**, which gathers exactly this information. However, authors are bound by the specific rules of the given journal, and therefore it is advisable to always check the rules of the given journal on the publisher's website before sending the manuscript to the editorial office. In the case of pre-prints, it is more common to publish them on so-called pre-print servers, such as arXiv, or ChemRxiv.

The advantage of Green OA is the fact that it is an OA option without costs for authors/the institute.

If the rules of the journal and the provider allow it, IMC authors are recommended to deposit self-archived publications either in the **IMC CAS Zenodo Community**, or in **ASEP**.

2.2.2 Gold OA

In Gold OA, the author publishes in a so-called open journal, where the submitted article undergoes a peer-review process and is immediately freely accessible upon acceptance. The costs of publishing in this case are covered by the author, or their employer, through a publication fee, the so-called **article processing charge (APC).**

The Gold OA model also includes publications in so-called hybrid journals, which contain both articles available via classic subscription and freely accessible articles for which authors pay to publish in the OA regime. Information on funding methods is in section 2.3.

In a number of grants, the self-archiving of VoR published in the Gold OA regime is also required. However, the rules here are easier, and it is recommended that authors self-archive VoR either in the **IMC CAS Zenodo Community**, or in **ASEP**.

2.2.3 Platinum/Diamond OA

In **platinum/diamond OA**, publication is free for the author, and readers are provided open access just like in gold OA. The costs of running the journal are typically covered by a scientific institution or, for example, a learned society.

There are also other paths of OA, e.g., the **bronze path** – publishing in journals without an attached license, WARNING – This is not true OA as proper licensing (e.g., CC BY) is absent. Project finance providers do not recognize this form due to legal uncertainty.

2.3 Options for financing the costs of publishing OA articles

- **The author can pay fees from the providers' subsidy** – for most grant agencies, APCs are eligible project costs, but these costs often need to be kept in mind at the time of the grant application.
- **The fee is covered by a so-called transformative agreement with the publish** – IMC authors can publish with selected publishers without the need to pay APCs, thanks to transformative agreements (so called "Read & Publish" or “tokens”) that we have concluded with publishers through the national consortium CzechELib. These agreements ensure classic full-text subscriptions for the institute while providing scientists with the opportunity to publish for free in the OA regime.

Current information on publishers and conditions for use of tokens by IMC authors is available on the IMC Library web or on the intranet.

2.4 How to select an open journal

A number of tools are available for selecting an open journal, the most common are:

- **Open policy finder**
  provides information on the open access rules of publishers from around the world, even by individual journal titles = whether publishers allow the author to self-archive
- **Directory of Open Access Journals**
  an overview of trusted OA journals incl. diamond OA

How to publish an article in the OA regime:

1. Find out the specific requirements of the project finance provider.
2. Choose an open journal (use DOAJ) and learn in advance about the APC.
3. Clarify options for potential APC financing – paid from the project or covered by a transformative agreement with the publisher (contact the library).
4. Deposit the article in a repository – proceed according to Directive No. 12.
5. If the provider's conditions require it and the publisher's policy allows it, you can deposit a preprint of the article on a preprint server (e.g., arXiv or ChemRxiv), where it is freely available before official publication.

3 Research Data Management

Research Data Management (RDM) is a supporting discipline focused on the organization, documentation, security, preservation, or publication of data created in the research process.

Research data is any data created during research, but not the information extracted from this data. Thus, this includes, for example, tables, images, audio, photographs, software, database data, outputs from measuring instruments, laboratory notebooks, questionnaires, 3D models, scripts, simulations, etc. This does not refer to data visualizations, which are sometimes published as so-called electronic supplementary materials, but to separate files that can be further worked with. Furthermore, this does not apply to reports, budgets, data management plans (DMPs), or other similar operational documents.

All collected data (collections) in various formats relating to a specific research/project are designated as a dataset, which is described, stored, and if deemed appropriate, can be shared through a trusted repository as a whole.

3.1 FAIR principles

FAIR principles (Figure 2) are mandatory by Act No. 328/2025 Coll. for all data originating from research supported by public funds. Fulfilment of these principles is generally required to a greater or lesser extent by all support providers incl. GAČR, AZV, TAČR, or MEYS. Emphasis on FAIR principles is placed especially by projects provided by the European Commission.

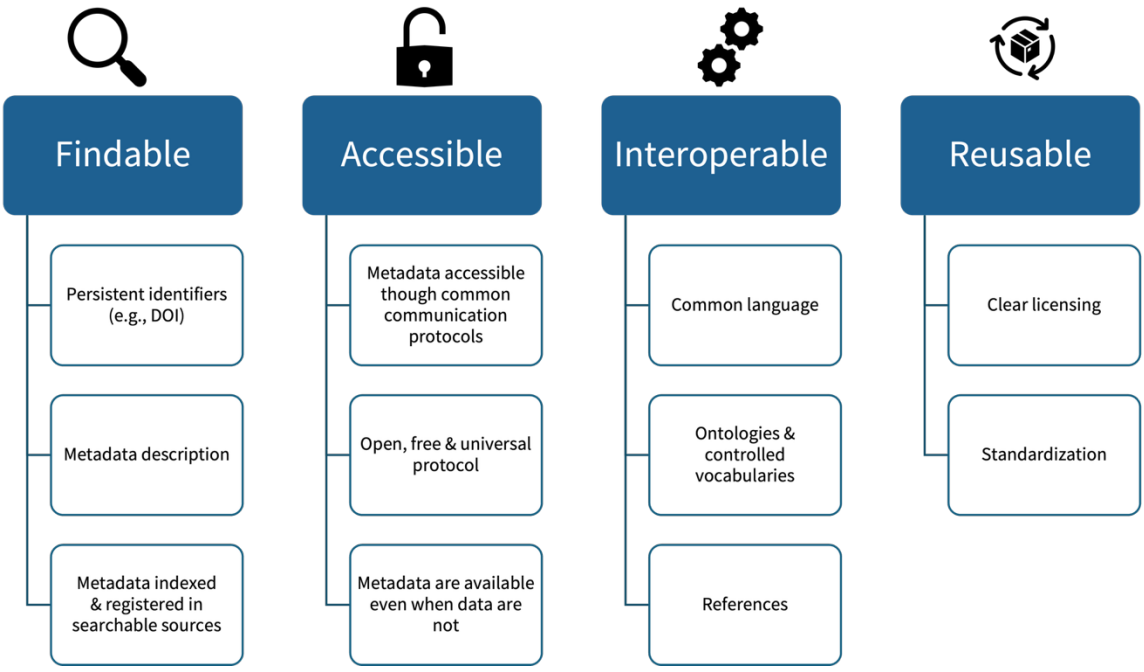

Figure 2 FAIR Principles, adapted from Jan Vališ 2026 [online].

3.2 Open data

The application of FAIR principles does not mean that data is necessarily open. Data sharing is always governed by the principle As open as possible, as closed as necessary. When deciding on the (non)publication of data, it is necessary to consider:

- legal and ethical aspects of their publication (personal or sensitive data)
- security impact (sensitive data, dual-use/military potential data)
- commercial interests (trade secrets, protection of intellectual property incl. patents)

These reasons for non-disclosure are usually accepted by support providers after proper justification in the Data Management Plan (DMP).

Some data may be published with a time delay, a so-called embargo – once the reasons for closing research data have passed, they can be made accessible. Some data may be sensitive in their original form but can be published openly as anonymized, or aggregated (the goal is to conceal the identity of all persons or organizations in the research outputs). Examples of anonymization tools: Amnesia, NML Scrubber, ARX Data Anonymization Tool. If the data cannot be anonymized, or if anonymization/aggregation would lead to an irreversible loss of data quality, it is

necessary to set up user authentication or authorization for access. This is enabled by repositories with so-called controlled access.

3.2.1 Storing open data

For support and assistance with research data management, do not hesitate to contact the institutional data steward of the IMC (datasteward@imc.cas.cz).

Data should be stored in so-called preferred formats (standard, open, independent of specific software, suitable for archiving) and provided with metadata, a persistent identifier, a license, and documentation:

- **formats** – standard text and spreadsheet (pdf, txt, rtf, csv, html, xml), video (mpeg, avi, mkv...), image (preferably lossless tiff, or jpeg or png), archive (zip), audio (mp3, flac, ...), measured data (in ASCII), software outputs;
- **metadata** – information about datasets enabling their machine traceability (at minimum the DataCite metadata schema – ASEP and Zenodo provide this);
- **persistent identifiers (PID)** – serve for the permanent and unambiguous identification of an entity (person, document, datasets, etc.) including their metadata, regardless of location:
- DOI/Handle for digital objects
- ORCID for authors
- ROR for organizations (IMC Persistent identifiers: ROR: 0143w7709; ISN:0000000106676325)
- **licenses** – serve to clearly define the rights of data users and protect the intellectual property of the IMC. It is necessary to balance the interests of the authors/IMC and the support provider. CC licence are typically recommended for datasets; established alternatives exist for SW, such as MITor GNU GPL;
- **documentation** – serves data users to understand the organization of the dataset and secure the necessary SW for opening and reusing data. Every dataset should contain at least a ReadMe file (e.g., Cornel University), or additionally instructions, protocols, etc.

Datasets are deposited separately into data repositories. Data publications through journal publication supplements are not accepted by providers because they do not have

their own persistent identifier, metadata, and license, and thus do not meet the FAIR principles.

3.2.2 Data repositories

Repositories are used for storing, publishing, protecting, and preserving final data after the conclusion of research. Grant providers usually impose the condition of depositing datasets into a trusted repository. Generally, it is recommended to find out who operates the repository, what functions it offers, whether it provides open access, provides data with a license, whether it assigns persistent identifiers for data, allows updates, or even holds a certificate (CoreTrustSeal, Nestor Seal,ISO16363). All this information can be found, for example, in the directories of OA repositories **OpenDOAR** or **re3data.org**. Repositories can be divided into disciplinary (e.g., **Chemotion**), general (e.g., **Zenodo**), or institutional (e.g., **ASEP**).

For data archiving, you can use a repository relevant to your field, or specific community. If none is available, a general repository can be chosen. If no suitable one exists, an institutional one can also be chosen. The choice of repository is advisable to consult with the IMC data steward (datasteward@imc.cas.cz).

For IMC authors who decide to publish their research data in the general **Zenodo** repository, we recommend publishing their datasets within the **IMC CAS Zenodo Community**. Publication in the community enables the institute's data steward to check the record before it is published.

IMC authors can also use the institutional repository ASEP, which is managed by the CAS Library and which we also use to record the publication activities of institute employees. In case of interest, please contact the Library (Eva Čechová, cechova@imc.cas.cz, the authorized person for data & publications entry in ASEP).

3.3 Data Management Plan

What data and how it will be created, managed, and preserved during research, how it will be accessible and further utilized, all this is specified by a Data Management Plan (DMP). Most project providers mandate its creation at the beginning of the project or with the first interim report and its regular updating during the project implementation in their conditions.

Various tools can be used to create a DMP. IMC investigators are **strongly encouraged to use the FAIR Wizard** tool (Figure 3). After answering the necessary questions, the tool will create a DMP in the template required by the providers (HE, MEYS, GACR, TACR (except for contract R&D) or, for example, AZV). For investigators outside the CAS who do not have access to the FAIR Wizard, there is an alternative in the form of the national instance Data Stewardship Wizard, or perhaps DMP Online, or Argos.

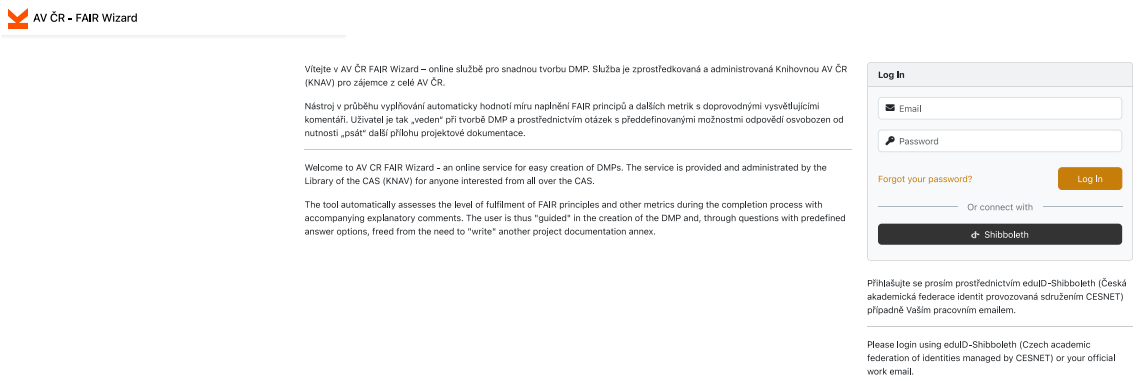

Figure 3 Fair Wizard: login screen, where it is necessary to select the black Shibboleth icon, which leads to login via eduID (UMCH login with number)

Detailed instructions for the Fair Wizard are available on the CAS Library website.

3.4 How to publish open research data

0. **Create a DMP:** Ideally at the beginning of the project with subsequent updates. Early DMP creation will allow you to set the basis for data management procedures (e.g., how to name files, in what formats, and where to store them), which will facilitate data publication.
1. **Find out the specific conditions of the project provider** – the acceptability of an embargo, freedom of repository choice, required licenses, reporting datasets, etc.
2. **Consider which data will be suitable for publication**, or under what conditions it can be published (embargo, controlled access, anonymization?). Specify which data can be shared and which cannot (describe in the DMP). All co-authors must agree to the sharing.
3. **Prepare the dataset.** Ensure that the data is well organized and in a suitable format. If possible, use commonly recognized standards and formats for your field. Attach documentation to the data that explains the data structure, how the data was

collected, the tools used, etc. The documentation can be, for example, in the form of a ReadMe file.

4. **Choose a suitable repository** – either according to the conditions of the project provider or according to disciplinary customs. If using the institutional repository ASEP please contact Ms. E. Čechová (extension 358), who is the authorized person for storing data in ASEP. In the case of publication on Zenodo, use the IMC CAS Zenodo community managed by the institutional data steward (extension 308).
5. **Fill in the required metadata** (creators, dataset title, abstract, keywords, creation date, and licensing conditions, related works such as journal publications). The author cannot assign a persistent identifier to the data themselves – the repository must ensure this.
6. **If the dataset relates to other works** (other datasets, SW, articles, etc.), it is advisable to indicate this as a so-called qualified reference. Similarly, it is good to state in the article that the data for the article is available through dataset DOI. The dataset should therefore be uploaded to the repository no later than the day the manuscript is sent to the editorial office. However, uploading does not mean publication – to protect publication priority, it is possible to close the dataset under a so-called embargo, thanks to which the content of the dataset will not be publicly accessible for a predetermined period and publication will only occur after the article is accepted for publication.
7. **License the data:**
    - Decide what rights and restrictions you want to apply and choose an appropriate license (e.g., CC BY 4.0), see. 1.3.
    - State the license:
        - in the dataset metadata in the repository and in
        - the attached documentation (e.g., in the form of a LICENSE file, or at least in the ReadMe file).
    - Clearly formulate the conditions of data use for their users. If you have specific requirements, e.g., restrictions on sensitive data, or you don't know exactly how to formulate the conditions of use, contact the legal IMC dept.
8. **You can indicate a recommended citation format for the data** – some repositories already offer this based on the entered metadata.
9. **Update data as needed**, upload new versions, update documentation. Some repositories also provide usage and citation statistics (Zenodo, Figshare).