Guidelines - Open Science workshop

The purpose of this document is to serve as a sort of checklist, or set of guidelines, to support personnel that aim to organize open science workshops within their academic institution. The intended target audience of this set of workshops are researchers varying stages of their career, though most notably early career researchers. The courses are mostly informative in nature, with a focus on practical tips for researchers to incorporate (certain elements of) open science within their daily research practice – either out of inherent interest or a need to adhere to (funder) guidelines.

Within the ENABLECARES project, the decision was made to discuss the broad open science movement in less detail due to the fact that only a limited number of open science components are of immediate interest to every participating institution: i.e. open access publishing and FAIR data management. While other relevant components are mentioned, the practical nature of the provided content did not allow for a detailed approach of every topic. Thus, after the summary introduction to open science, the course at large is further divided into two modules, with a focus on the practical implications and useful tools/tips related to open access (module 1) and data management (module 2).

These guidelines are meant to be used in conjunction with examples/templates available through the ENABLECARES training environment. These templates consist of a set of slide decks as well as some example quiz-assignments that address some of the more common questions/issues. The templates can be used by trainers as a basis to further build upon and adjust according to their own specific context. The questions in the guidelines below are intended to help local trainers further expand upon the templates by including slides on their own policies, requirements, tools, infrastructure, etc.

Please note that the current slide decks may contain images that are subject to certain licenses that prevent open access sharing. If you aim to share these slides (either in original or adjusted form), make sure you remove any images and/or replace them with open license images.

0.   Introduction to open science

This is intended to be an introduction to open science: definitions, main benefits, related topics, etc. Because of its summarizing nature, the slide deck covers only the basics of most of the open science components. The user is free to expand upon according to the level of knowledge already present amongst scholars, and the needs of the institution. Once a certain component is discussed in great detail, our advise is turn this into a module of its own.

1.1 Introduction to OA

The first sub-module focuses on the introduction of the topic of open access/open scholarly communication. While this may be well known amongst more senior researchers, be aware that for junior scholars a lot of this information (even basics, like academic publishing in general) may be entirely new. Hence the need to properly introduce the topic.

Themes:

- o What is OA exactly?
- o To what types of publications does OA refer?
- o Benefits and downsides of OA.

Checklist for local trainer:

- o If you feel this is can be of interest, inform yourself of local highlights/best practices (no. of repository full text downloads, for example). This allows you to either emphasize the need for more action, or show how prevalent the practice already is within your institution.

1.2 Open access in the publishing landscape

The goal of this sub-module is to fit the open access form of publishing into the wider context of academic publishing and its details. Open access does not stand on itself; it exists as part of a larger ecosystem that is varied and can be confusing, even to more experienced scholars. Use this module to highlight the differences between various forms of access, but also to explain how publishing actually works.

Themes:

- o Various forms of access to academic publications.
- o The publication process and paper versions.

1.3 The preprint server

Preprints are becoming increasingly important to both the research process in general, as well as academic (open access) publishing. This sub-module explains the principles, benefits and potential risks of preprints and how they fit within the publishing ecosystem. A guide is available for more information for both trainer as well as students.

Themes:

- o What is the preprint, and how does preprint sharing work.
- o Characteristics, benefits, risks of preprint publishing.
- o Preprint servers.

Checklist for local trainer:

- o Does your institution have a policy on preprint archiving/publishing?
  If so, make sure you explain this in detail during the course.
- o Are many scholars/departments within your institution particularly active in this area?
  You can use this to promote the practice within your institution.

1.4 Open access publishing

This sub-module goes into the details of the various forms of open access publishing, ranging from hybrid to diamond. Make sure you properly explain the specifics of each form, including benefits and (potential downsides). Explain Read&Publish/transformative deals, and inform student how to find these within your institution. Properly highlight and explain available tools and infrastructure. If there is funding available to publish in OA, inform students where to find this and how to apply. This sub-module should offer students most of the practical information to circumnavigate the world of open access publishing on their own – including making use of support available within your own institution.

Be aware that green open access can be confusing if not explained properly.

Checklist of local trainer:

- o Inform yourself about any latest developments (international, national, institutional). Open access is constantly in flux, with changing policies, strategies and academic cultures shaping behaviour of scholars, university management and support, and publishers alike.
- o Be aware of discipline-related specifics (if focused on a specific audience). What is the average/most common APC cost? What are the major full OA journals within your discipline? Any major platforms? What are the costs involved, for the researcher? Are things like diamond OA, preprint publishing common? Try to include all medium/large-sized platforms, and a selection of the most commonly used journals.
- o Are there any Read&Publish/transformative deals that your institution takes part in? Do these include full/hybrid journals? Where/how is this information shared? How does the workflow for these deals look like (both for the researcher and support personnel)?
- o Do you have any institutional tools/infrastructure that students can use (for selecting the right open access journals, for example)?
- o Is there an OA fund within your institution? And on a university/national level?
- o Is there expert support available within your institution? What sort, and where? How can students get into contact.
- o Familiarize yourself with databases and tools like the Directory of Open Access Journals (doaj.org), COPE, OASPA, Sherpa (Romeo & Juliet), the Journal Checker Tool, etc.
- o Do you have an institutional repository?
  - How does it work?
  - How is it connected to the local CRIS?

- Is the repository/CRIS linked to national or international databases (e.g. OpenAIRE)? And how?
- How is support organized? Who is responsible for what?
o Are there any other repositories in use within your institution (external; registered in OpenDOAR)?

1.5 Open access (funders) requirements

This sub-module covers the most important funders within your discipline, and how their open access requirements (greatly) impact the publication behaviour of grantees. Of considerable importance in discussing this topic is to emphasize what (types of) publication venues are allowed per funder, so that researchers are aware that some of their intended journals are off-limits.

You can highlight the effects of funder requirements on the personal context of individual researchers by having them check for themselves what journals they often consult/publish in are prohibited under certain funder requirements.

Checklist of local trainer:

o What is your institutional policy on open access?
- Is the type of venue important, and if so, how does this impact publishing behaviour?
- Do you have a way to monitor compliance? Are there penalties involved?
o What does your national copyright law state about OA publishing? Any additional jurisprudence of relevance to your institution?
o What funders are active within your field (national, discipline-focused, etc)? For European-based institutions, include at least Horizon2020/Europe/ERC (even if only as a comparison).
- What are the OA requirements?
  - Gold vs green vs hybrids
  - Archiving vs making OA available
  - Allowed embargo
  - Allowed licenses
o Does your national/commonly used funder participate in/adhere to the principles of cOAlition S?
- How do they implement the Plan S tenets?
- (If switch to Plan S is recent) How is this different from the previous situation?

1.6 Predatory/questionable Publishing

Predatory/questionable publishing is a real risk for especially junior researchers (though more senior scholars are equally susceptible), and so it is of great importance to explain the phenomenon, warn about the risks, and to offer tools, tips and guidelines to avoid these journals. In addition, this

sub-module can be used as a handle for the local trainer to cover quality in publishing in general, and debunk persistent misconceptions about quality and open access publishing.

Checklist of local trainer:

- o Familiarize yourself with major antagonists within your discipline.
    - Distinguish between truly fraudulent/predatory journals and those that seem to be more legitimate, but do conduct questionable practices
- o Familiarize yourself with online tools/databases:
    - Directory of Open Journals (DOAJ)
    - OASPA
    - COPE
    - Think, check, submit checklist
    - List of predatory Publishers (note the disadvantages of blacklists for which the criteria are not transparent)
- o Do you have supporting personnel with knowledge of predatory publishing that researchers can consult?
    - Any tools/databases/guides within your own institution/country?
- o Do you know of any best/worst practices within your discipline/institution?
    - Mentioned in local/national/institutional newsitems?

1.7 Copyright issues

This sub-module covers the principle of copyright, and how this relates to publishing in open access. From a practical perspective, the Creative Commons License is explained in detail, as well as how open access requirements from funders and institutions in combination with licenses allowed by funders may once again limit the choice of publication venue.

Checklist of local trainer:

- o Inform yourself of copyright
- o Do you have expert supporting personnel focusing on copyright?
    - Any websites/guides/tools to assist researchers?
- o Is there any policy on licensing within your institution/country?
- o Familiarize yourself with Creative Commons licenses
- o Identify most common possible conflicts between institutional/commonly used funder open access license restrictions and licenses allowed by commonly used publishers/journals.

1.8 Finding (open access) publications

While the above themes cover practical information on where and how to *publish* in open access, this sub-module focuses on where researchers can easily find and access (open access) publications.

The main aim of this module is to explain researchers how to best make use of the benefits of green and preprint open access publishing, though trainers can use this to highlight other (open access as well as non-exclusively open access) databases to scholars.

Checklist of local trainer:

o   Familiarize yourself with plug-ins/browser extensions.
  • Unpaywall
  • Google Scholar Button
  • Open access button
  • Kopernio
o   Familiarize yourself relevant databases that offer (legal) access to otherwise closed publications
  • OSF preprints*
  • CORE & Core Discovery*
  • Dimensions
  • Science Open
  • Directory of Open Journals (DOAJ)*
  • Zenodo*
  • PubMed, Scopus, etc.
(With asterisk only show open access publications)

o   Familiarize yourself with request PDF options
  • #icanhazpdf
  • Request from author
o   Does your institution offer services to access otherwise closed publications (via inter-library connections, copying of excerpts, etc)?

Please contact ENABLECARES to gain access to template slide decks and start tailoring your own Open Science and Research Data Management Workshop:

https://enablecares.eu/contact-us/

Guidelines – Research Data Management (RDM) workshop

2.1 Introduction to RDM

The research data management module starts with a general introduction, which is made up of an explanation of the most important definitions of research data, and main goals and requirements for proper research data management. Participants are made familiar with obligations in terms of applicable laws and regulations, and which elements are important for ensuring high quality research data. The introduction also explains where support can be found within (or outside) the participants' institute on RDM related topics.

Themes:

- Introduction to the topic research data management in general
    - Definition of research data and research data management (RDM)
    - Applicable laws and regulations, as well as institutional and (inter)national RDM policies
- Introduction to the research data lifecycle
- Introduction to the FAIR principles, as part of Open science and open data
- How is RDM support arranged within the institute?

Checklist for local trainer:

- To start with, it is advisable to first assess the knowledge and experience level of the participants, and their expectations of the workshop. After all, there are major differences between junior researchers and senior researchers and supporters. The type of research (human-related or not, big data or small-scale research, qualitative or quantitative data handling) is also relevant for the contents of the workshop.
- Does your institution, region, country have a RDM policy? If so, prepare an overview of relevant policies and also consider related laws and regulations, such as the GDPR.
- What services are available to researchers? Consider, for example, support in the areas of data stewardship, ICT, privacy, valorization, etc.

2.2 Data management planning

This sub-module emphasizes the importance of proper planning of data management activities, and how a data management plan (DMP) can be used for this purpose. The requirements of funders, journals and their institute on the DMP are further explained.

Themes:

- What are the purpose and contents of a DMP?
- Examples of DMP template(s) from funder(s), institute or research theme.

Checklist for local trainer:

- Does your institution have an institutional DMP tool and/or template? If so, it is advised to use this DMP as an example.
- Check which funder DMP templates are applicable to your institute, and familiarize yourself with the procedure(s) on how to deliver the DMP.
- Consider to use a data management plan as a practical basis for discussing all topics, let the participants fill in (or discuss) parts of the DMP throughout or after the course or by means of an assignment after the course.

2.3 Data creation and collection

This sub-module raises awareness of the possibility of reusing data, as a sustainable alternative to recreating pre-existing data. If reuse is not possible, it is important to know what aspects are involved in the creation and collection of research data.

Themes:

- Why and how to reuse existing data for research.
- How to create and capture new data for research.
  - o Examples of newly created data, along with the tooling that is used to create them
- Which data types and formats are usually reused / collected / generated within the institute.

Checklist for local trainer:

- Make an inventory of data sources, within and outside the institute, that are available to the researchers for reuse. What are the limitations on the use of these data sources?
- If possible, arrange or provide a demonstration of the tooling researchers can or should use for data capture. Think for instance about electronic lab journals, (clinical) data capture tools, etcetera.

2.4 Data storing and sharing during research

There are substantial differences between data storage and sharing during the research project, and longer-term storage (archiving) after the research. This submodule presents the storage options offered by the institute during the project, preferable by describing best practices.

Themes:

- Institutional and external storage facilities for research data.
- Which prerequisites exist for secure and sustainable storage of research data?
- Which facilities for data sharing and collaboration with other parties are available?

Checklist for local trainer:

- Prepare an overview of the available storage facilities for research data, and provide advice on how to use them, and for which type of data. Different requirements are imposed on the storage of human participant data than on non-sensitive data.
- Prerequisites for storing research data: backups, versioning, security.
- Be clear about bad examples and explain why not to use them, for instance private devices (laptops), portable data carriers and unsafe cloud solutions like for instance Dropbox.

2.5 Data documentation and organization

Although this topic is not subject to defining laws and regulations, properly structuring and documenting research data is of key importance for any research project to enable verification and replication of research findings, and to facilitate future reuse of data within the context of open science.

Themes:

- Documentation to describe the data and the analyses steps
- Explanation of different types of metadata
- Best practices in organizing datasets in terms of file structure and naming conventions

Checklist for local trainer:

- Know which metadata standards are applied in the field(s) of research of your institute. For instance, metadata catalogues, web portals, repositories require certain generic or discipline specific metadata about research datasets.
- Gather good examples of documentation, file structures and naming conventions from research groups within the institute, if present.

2.6 Archiving and (FAIR) sharing of research data

This sub-module addresses the long-term storage of data, data archiving, after the research project has ended and results are published. In addition, in the context of open science / open access, researchers are expected to share data publicly as openly as possible, restricted where necessary. The FAIR principles are explained, and practical examples of how these principles can put into practice.

Themes:

- Archiving of research data: which data must be stored for the long term in various types of research (human and non-human research).

- Valid reasons for applying limitations on data access.
- Data archiving facilities.
- FAIR principles to enable reuse of data.
- Tools for data and metadata sharing.

Checklist for local trainer:

- Based on the types of data used within the institute, an estimate can be made about what restrictions might be placed on sharing these types of data. Again, as in previous modules, sharing human participant data or other data of sensitive nature may be subject to restrictions. If applicable, prepare an overview to provide guidance on archiving and sharing of various data types like non-sensitive and sensitive data.
- What facilities exist within the institute for archiving and sharing data and metadata? For example, a data repository, a data archive, data catalogue(s), etcetera. Present these facilities alongside public resources for data sharing.
- Next to explaining the FAIR principles, provide practical examples on how the FAIR principles can be applied.
- FAIR is a subject that undergoes rapid developments, especially in terms of FAIR tooling: FAIR data points, FAIR implementation profiles, etcetera. The level of detail of this topic strongly depends on the maturity of the FAIR policies, expertise level and infrastructure of the institute.

2.7 Research Ethics: personal data and informed consent

Although not every research institute conducts research with human data, it is beneficial for the professional development of young researchers to also be aware of certain aspects that are specifically applicable in this type of research. After all, one can still go on to work in this research field later in their career, and then it is good to have some basic knowledge about handling personal data in research. Nevertheless, it is up to the local trainers to decide on whether to include this topic in their local training.

Ethical aspects around handling personal data are discussed in this submodule. Ensuring privacy of human research participants, the GDPR and informed consent are the topics that are covered.

Themes:

- Definitions of personal data
- The GDPR and its implications for research
- Pseudonymization and anonymization
- Informed consent

Checklist for the local trainer:

- Is personal data used for research within the institute?
- If yes, provide information on who's responsible (and/or provides support) for data protection and privacy issues.
- If no, decide on to what extent this topic will be included in the training.
- Is specific information or tooling available within the institute or open source on data or device encryption, anonymization and pseudonymization?
- Provide examples of best practices for wording in informed consent forms, and include not only using data for current research, but also for enabling reuse of data.

Please contact ENABLECARES to gain access to template slide decks and start tailoring your own Open Science and Research Data Management workshop:

https://enablecares.eu/contact-us/