Open Science

A Practical Guide for Early-Career Researchers

NWO

Universiteiten van Nederland

Data Archiving and Networked Services
DANS

Universiteitsbibliotheken & Koninklijke Bibliotheek

Contents

1. Welcome to Open Science

'In Africa there is a concept known as 'ubuntu' – the profound sense that we are human only through the humanity of others; that if we are to accomplish anything in this world it will in equal measure be due to the work and achievement of others.'

> Nelson Mandela

Reliable science is not the sole work of superhuman geniuses, but a collaborative process. Researchers rely and build upon each other's work. Together, we build theories, collect evidence and assess the research of colleagues. However, we can only build upon others' work if we know exactly what our predecessors have done. What were their methods, relevant materials, data and outputs? Therefore, sound science ideally equals Open Science, in which all phases of the research cycle are as transparent and accessible as possible (for examples, see Figure 1a + 1b).

Currently, the phases of the research cycle are often not as approachable as they could be. Many research articles and edited volumes are locked behind expensive paywalls, and more often than not, research data, methods, or procedures are not accessible to others at all.

Open Science represents the international aim to make scientific knowledge openly available, accessible, and reusable for everyone, and to promote scientific collaboration for the benefit of science and society (UNESCO Recommendation on Open Science, 2021). Different countries are moving towards openness at a different pace. Internationally, the Netherlands is regarded as a front-runner. Since 2017, the Dutch government has stated in its coalition agreements (2017-2021, 2021-2025) that Open Science should become the norm in science and has followed suit by investing in these democratising endeavours as much as possible.

Dutch universities are bringing forth Open Science Programmes to make Open Science a reality. Even though it is the responsibility of the institutions and policymakers to remove barriers and enable researchers to transform their habitual workflow, researchers (that's you!) are ultimately the ones who will have to make commitments and apply practices to open up science. As you begin to prepare for and work on your doctorate – not unimportantly the last stage of your formal education, and the first stage of your professional life – it is up to you to become an agent of change and to put these new open research principles into practice.

Figure 1a

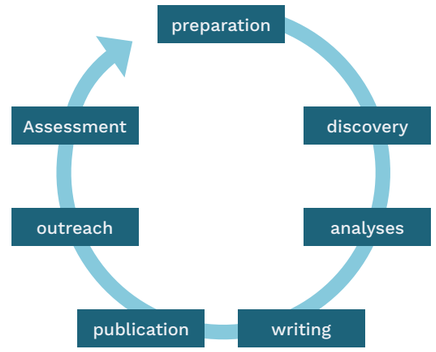
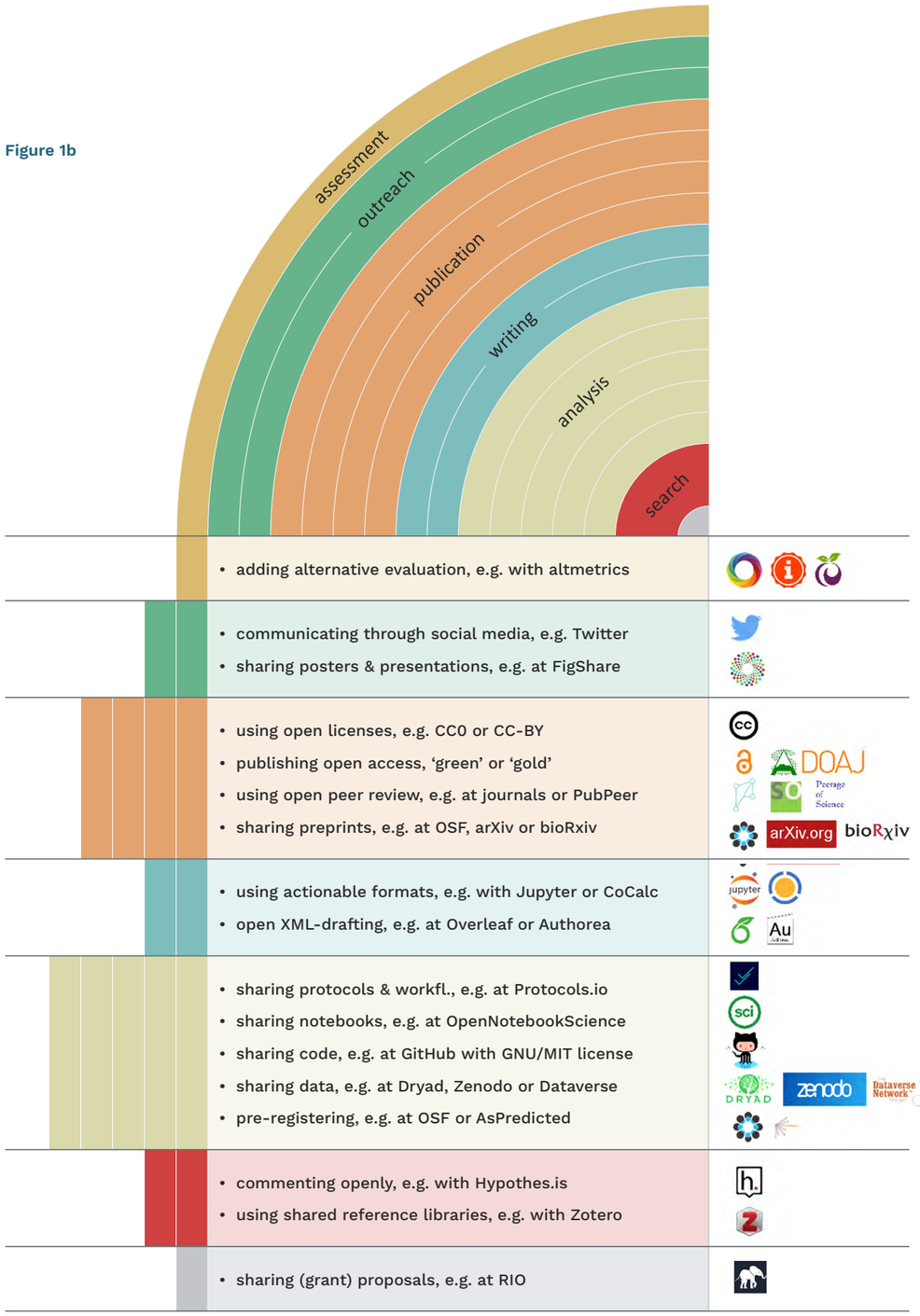


Figure 1b



DOI: 10.5281/zenodo.1147025

This guide is aimed at PhD candidates, Research Master Students, and early-career researchers from all disciplines at Dutch universities and research institutes. It is designed to accompany researchers in every step of their research, from the phase of preparing your research project and discovering relevant resources (chapter 2) to the phase of data collection and analysis (chapter 3), writing and publishing articles, data, and other research output (chapter 4), and outreach and assessment (chapter 5). Every chapter provides you with the best tools and practices that can be implemented immediately.

If the information in this guide feels overwhelming: Do not worry! Open Science is a journey, and you are not alone in this. There are always colleagues who are happy to help you out along the way. You can find fellow researchers or support staff at your local University Library and at your Open Science Community (see chapter 6).

We hope these instructions will motivate you and help you to practise Open Science, by sharing all aspects of your research with as many people as possible.

The Future is Open!

2. Prepare & Discover

'Diverse science leads to better science!'

> Carolyn Bertozzi

The preparatory phase of research is all about further shaping your first ideas and discovering relevant resources and perspectives. The sooner you are familiar with the principles and possibilities of Open Science, the more beneficial the move towards openness will be for your project. When you design your research in an open manner from the get-go, Open Science will not become an afterthought. Instead, it will frame the way in which you set up your research.

Finding Stakeholders

One thing that is often overlooked in the preparatory phase, is the possibility of early outreach to possible stakeholders. Indeed, opening up your project from the start can be very helpful in refining your research question, and to make it as (scientifically and socially) relevant as possible. As such, it may also result in more impact during and after the project or lead to additional resources that can help expand your research. So go ahead, and find as many experts, companies, non-profit organisations, politicians and policymakers as possible to follow your project closely!

Useful Tools

You can quite easily set up an interactive project page (for instance via PubPub) to inform others, share a newsletter or a blog, or communicate via social-media platforms such as Mastodon, Twitter or LinkedIn. Open Science platforms such as ResearchEquals also offer templates that can help you to immediately start sharing all sorts of information and output.

University librarians Bianca Kramer and Jeroen Bosman developed a handy tool for thinking about your dissemination strategy.

Defining stakeholders early on also helps you to think about a publication and dissemination strategy for your research project. By defining who you want to involve with your research and what you hope to achieve, you can anticipate what kind of outputs you want to generate and how you want to share them. Your specific topic or approach may beg for additional outputs, such as educational materials or translations for which time and funding need to be allocated, preferably at the start of the project.

Citizen Science

In some cases, it can be interesting to take your stakeholder management a step further, and to open up your research to citizens. We call this **citizen science**.

A citizen scientist is more than a respondent, subjected to a survey or an interview. Rather, a true citizen scientist is someone who complements the team of academic researchers in order to gather, analyse, and/or evaluate information.

An example of citizen scientists is the group of civilians who suspected the nearby petrochemical company Formosa of polluting the environment. They began to gather evidence and formulated hypotheses that would eventually be included in the official research report, produced by scientists and the police.

Within the Dutch National Programme Open Science (NPOS), a Citizen Science Network was established to connect initiatives and to discover the full potential of citizen science. More examples can be found here.

Discovering Information

Research always starts with evaluating what others have done before. More and more academic resources are openly available. Via freely accessible search engines, such as Lens.org or OpenAlex, you can easily find Open Access literature and data.

Here are some examples: The Directory of Open Access Journals (DOAJ) contains more than 18.000 trustworthy academic Open Access journals; the Directory of Open Access Books (DOAB) indexes almost 60.000 peer-reviewed Open Access books; OAPEN enlists monographs and books published Open Access; OpenDOAR is an international platform for finding open repositories; OpenAIRE gives access to millions of publications, data sets and research software, largely openly available, from many data sources; with BASE you can search through the content of many repositories.

Specialised search engines can help you find Open Access materials: Core is a search engine specialised in Open Access academic publications (books, articles, theses, etc.); Unpaywall can be installed as a browser extension and leads you to possible Open Access versions of all the articles you are searching on the web.

Useful Tools

Openly available literature can be found through search engines and repositories such as Lens.org, OpenAlex, BASE, Core, Dimensions, EuropePMC, Directory of Open Access Journals (DOAJ), Directory of Open Access Books (DOAB), OAPEN, OpenDOAR, OpenAIRE, and Unpaywall, which offers a handy browser extension to find open versions of closed-access literature.

You can also check platforms for papers in progress, so-called **preprints**: These articles are not yet peer reviewed, type-set or published in a traditional publishing venue, but have already been shared so that others can discuss the contents (see also Preprints). The availability of these in-progress articles also means that researchers can quickly discover the latest research conducted on a certain topic. Open Science Framework (OSF) assembles more than 30 preprint servers and many

other preprint repositories are available. Additionally, Lens org, Dimensions and EuropePMC (for life sciences) are search engines that can help you find preprints.

Discovering Data

Another source of information that you may want to get access to, is data. Commonly, researchers promise to make data sets available 'upon request', but in reality, this process can be cumbersome (e.g., because the email address is outdated, or researchers simply do not respond to requests). However, there are places where data is stored findable and accessible. More and more often, data is linked to journal articles. If so, the journal article will specify where you can access the data and under what conditions you are allowed to reuse the data.

Another way to search for data is through data warehouses; central repositories that combine data from a variety of sources. Data warehouses can either be multidisciplinary or specialised, and allow you to deposit, conserve, and share research data.

Useful Tools
FAIRsharing.org lists available data warehouses in the life sciences. DataverseNL is the Dutch data warehouse for online storage, sharing and registration of research data. EASY is the largest online archiving system for depositing and reusing research data in the Netherlands. 4TU.ResearchData guarantees the storage, reuse, and continued accessibility of science and engineering research data. Databases for Open Access theses – such as DART-Europe and Open Access Theses and Dissertations (OATD) – aggregate the digital thesis collections of universities and research centres. In addition to data, you can also find software in the repository of the Netherlands eScience Center and 4TU.ResearchData.

Lastly, it is also useful to mention the existence of open research software made possible by the Open Source Community, which is part of Open Science. The Netherlands eScience Center runs the Research Software Directory, where you can find all sorts of open software. In addition, you can find pieces of software (and data) you need on GitHub or GitLab (see also Reproducibility).

Open Science > A Practical Guide for Early-Career Researchers

2. Prepare & Discover

Preregistration

Another Open Science practice you can implement in the very first stages of your research is sharing and registering your initial ideas, theory, hypotheses and research design, open notebooks, and, if applicable, your Systematic Literature Review or Scoping Review of scholarly articles on your subject.

This way, you can tease out early feedback and discussion. It might be a somewhat frightening idea to share your 'early' work: Will others copy your approach? Another way to look at this, however, is that early sharing enables you to 'claim' and protect your view on the topic right away.

The most 'official' way to share hypotheses (and other elements of your research design) is called **preregistration**. This practice of making hypotheses explicit in a preliminary stage acts as a countermeasure against sloppy science, scientific fraud, and the replication crisis that haunts many scientific disciplines. The primary assumption is that preregistration will prevent researchers from working in reverse, such as by dredging data sets to find statistically significant patterns (also known as "p-hacking") or by postulating hypotheses a posteriori.

Preregistration must be executed before the actual research is conducted, but is often only made publicly available after completion of the research project. Some scholars go even further and make their preregistrations available at the onset of their studies, inviting colleagues and stakeholders to discuss their plans and to provide useful comments. A special case of this is the so-called **registered report,** in which a journal indicates – after peer review of the preregistration – that they will publish the results of that study in the future, regardless of the outcomes. This is a countermeasure of **publication bias**.

Useful Tools
You can share many different research elements digitally, for instance on the OSF, ResearchEquals, or GitHub. Systematic reviews can be shared via Prospero, or the more general platforms mentioned before. Guidelines for opening up your systematic review can be found here.

Templates for preregistration can be found via the Open Science Framework.

A preregistered hypothesis provides greater transparency of the scientific process, and underscores the quality of a project. Preregistration also makes it easier for other researchers to replicate your research approach, which reinforces the reproducibility of research (see also Reproducibility).

Although preregistration is more common and most applicable in fields of quantitative research and statistics, there are also voices that plea for the introduction of some form of preregistration in qualitative studies, e.g., fields within humanities or anthropology/ ethnography studies.

Licences!

Indispensable in the practice of Open Science is understanding and using licences. A licence tells you what you can and cannot do with public materials. There are specific sets of licences for software, data, and publications. Widely used and most well-known are Creative Commons licences.

Before using, adapting, or disseminating Open Source or Open Access materials, it is strongly recommended to check out the licence and what it means. When you publish materials yourself, always add a licence so there will be no doubt about the open status of your materials.
In Chapter 4, Write & Publish, you will find advice on what licence(s) to choose for your publications, your research data, and other research output.

Useful Tools

Creative Commons (CC) licences are often used for publications. You can usually find a CC licence in the form of an emblem in the colophon, somewhere on the first or last pages of the publication, or in the readme.txt that accompanies a file, piece of software, or a document. They look like this:

PUBLIC DOMAIN	0 – Available in the public domain without any restrictions (**Zero**)
BY	BY – Credit must be given to the creator. This is part of all the licences below. (**BY** whom was it made?)
BY SA	SA – Adaptations must be shared under the same terms (**S**hare **A**like)
BY NC	NC – Only non-commercial uses of the work are permitted (**N**on-**C**ommercial)
BY ND	ND – No derivatives or adaptations of the work are permitted (**N**o **D**erivatives)
	Combinations are also possible:
BY NC SA	Credit must be given to the creator, only non-commercial uses of the work are permitted, and adaptations must be shared under the same terms.
BY NC ND	This licence only allows reusers to copy and distribute the material in any medium or format in unadapted form and for non-commercial purposes.
	Creative Commons has developed a tool to help you choose a licence.

Open Science Prepare & Discover Checklist

- *Who are my stakeholders, and how can they be reached early in the research process?*
- *Would it be advantageous to let citizens partake in the research project?*
- *While discovering literature and data, were the openly available sources included in the searches?*
- *According to the licence, how are you allowed to use, adapt and disseminate the discovered literature and data?*
- *Is it possible to share early ideas, research designs, or even to preregister hypotheses?*

Photo: Shutterstock – Aksonsat Uanthoeng

3. Collect & Analyse Data

'If I only had an hour to chop down a tree, I would spend the first 45 minutes sharpening my axe.'

> Abraham Lincoln

From the very start of your research, you will collect, produce, and use data. Research data is defined as "factual records (such as numerical scores, textual records, images and sounds) [...] used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings" (source: OECD Council's Recommendation on Access to Research Data from Public Funding).

A research data set constitutes a systematic, partial representation of the subject under investigation. Examples of research data are text documents, spreadsheets, statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings (audiotapes, videotapes), images, database contents, models, algorithms, and scripts.

In order to use and analyse this assortment of data in a reproducible and open fashion, it is essential to carry out **Research Data Management (RDM)**. A carefully constructed Research Data Management Plan covers all activities involved in collecting, describing, storing, processing, analysing, archiving, and accessing data, and makes your research more understandable and reproducible for peers.

Data Management Plan

Data management needs to be anticipated at the very beginning of a project by creating a **Data Management Plan (DMP)**. A DMP maps out the entire data life cycle. It helps you think about how the data will be created or gathered. Should you use new data or reuse existing data? How will you organise your data, files, and other supporting documents during and after the project? How will the research data be documented, and what metadata and documentation do you use? Where will the data be stored during and after the research? Will the data be shared (how, when, embargo period, appropriate data licences) and how will you protect sensitive data?

If applicable, the DMP also describes the hardware and software needed to use the data. In this case, you use the DMP to communicate with, for instance, your colleagues from the IT department. A DMP is an ongoing document that needs to be updated throughout your research project. Many research funders, including the Dutch Research Council NWO, ZonMw, and the European Commission (Horizon Europe programme), require you to provide a DMP.

Sensitive Data

When handling and dealing with sensitive data, be mindful that you should give specific attention in your DMP to collecting, processing, handling, and storing data throughout the research process.

Since the EU General Data Protection Regulation – GDPR concerning personal and sensitive non-personal data came into existence, sensitive data can only be handled under strict regulations. Examples of sensitive data are:

- personal data; this is any information that relates to an identified or identifiable individual (name, address, identification number, but also biometric data, genetic data);
- confidential data (trade secrets, investigations, data protected by intellectual-property rights);
- biological data (such as endangered plant or animal species, where their survival is dependent on the protection of their location data).

Photo by Fili Santillán on Unsplash

Personal Data and Data Minimisation

When dealing with personal information, you should only collect and use the data necessary for your specific purpose(s) [Article 5, EU GDPR]. For instance, some survey-collection software collects dates, IP addresses, and latitudinal and longitudinal coordinates by default settings. If this does not serve any purpose for your research or if you have no permission to do so, please do not collect such data and adjust the default settings.

Useful Tools

There are several online tools to help you create a DMP. Examples are: ARGOS (DMP tool developed by OpenAIRE) and DMPonline (includes templates from NWO, EC, Dutch universities). Many Dutch universities and University Medical Centres have their own DMP templates. Examples of public Horizon-2020 DMPs can be found at DMPonline or in the collection Horizon 2020 DMPs.

23 Things for Researchers and PhD Candidates gives an overview of practical resources and tools that you can begin using today to incorporate research-data management into your research workflows. The guide for researchers 'How to deal with sensitive data' gives a clear overview of how to handle sensitive data.

Reproducibility

'Eat, sleep, rave, repeat.'

> Fatboy Slim

A methodological premise of many scientific domains is that when the same analyses are performed on the same data, the same results should be obtained. This is also called **reproducibility** (or: computational reproducibility). Reproducibility differs from related concepts such as replicability, generalizability, and robustness, which entail different data and/or different forms of analyses (see table below). Reproducibility and Open Science are separate concepts, yet they are intertwined; Open Science strives to make knowledge openly available so it can be reproduced and reused.

		Data	
		Same	Different
Analysis	Same	Reproducible	Replicable
	Different	Robust	Generalisable

Table adopted from The Turing Way

Reproducibility of research is neither a goal in itself, nor an absolute guarantee of quality, as research may be one hundred percent reproducible, while of low inherent value. Nonetheless, there are many good reasons to make your research as reproducible as possible:

- When all steps of your research are properly documented and available to others, errors are more easily detected and corrected. It makes your work more credible and adds to your accountability as a researcher. That is why working in a reproducible manner is part of the Dutch Code of Conduct for Research Integrity.
- Reproducible workflows make it easier for others to build on your research. They may reuse your data or the way you collected or analysed your data. It increases the impact of your work.

Working in a reproducible way is beneficial for science and society, but that does not mean that it is not beneficial for yourself as well. In 2015, Florian Markowetz sketched five selfish reasons to make your research more reproducible:

- Reproducibility helps to avoid disasters
- Reproducibility makes it easier to write papers
- Reproducibility helps reviewers see it your way
- Reproducibility enables continuity of your work
- Reproducibility helps to build your reputation

Working in a reproducible manner makes it much easier to revisit your data or analysis later on, which saves you precious time. Invest in reproducible workflows now, and your 'future self' will be most grateful!

Open Science > A Practical Guide for Early-Career Researchers

3. Collect & Analyse Data

Reproducibility in Practice

Reproducibility can be operationalised in different ways, depending on your discipline and the methods used, but always hinges on the availability of the research data and the methods by which the data was processed and analysed. Examples of reproducible research are experimental protocols that can be rerun identically, the reproduction of statistical processing of quantitative data, the reconstruction of the stages of analysis of a corpus of images or texts, and so forth. The key thing is to think ahead. Act *before* and *during* the process of your research to make sure your work is reproducible *afterwards*.

Reproducibility starts with proper data management. In your **Data Management Plan** (DMP), you describe how you are going to make your research reproducible. Think about how much time and effort another person will need to reproduce your results, and how you can minimise that effort. A detailed description of your data (metadata) as well as a description of your procedures and analyses accompanying your data is a good first step. When making your research reproducible, it is imperative to consider the following 10 things:

Useful Tools

Recently, an international working group of the Research Data Alliance compiled a list of 10 Things for Curating Reproducible and FAIR Research for you to consider when making your work reproducible. It focuses on the **research compendium** that "accompanies, enhances, or is a scientific publication providing data, code, and documentation for reproducing a scientific workflow" (https://research-compendium.science/).

10 Things for Curating Reproducible and FAIR Research

1 Completeness
The research compendium contains all of the objects needed to reproduce a predefined outcome.

2 Organization
It is easy to understand and keep track of the various objects in the research compendium and their relationship over time.

3 Economy
Fewer extraneous objects in the compendium mean fewer things that can break and require less maintenance over time.

4 Transparency
The research compendium provides full disclosure of the research process that produced the scientific claim.

5 Documentation
Information describing compendium objects is provided in enough detail to enable independent understanding and use of the compendium.

6 Access
It is clear who can use what, how, and under what conditions, with open access preferred.

7 Provenance
The origin of the components of the research compendium and how each has changed over time is evident.

8 Metadata
Information about the research compendium and its components is embedded in a standardized, machine-readable code.

9 Automation
As much as possible, the computational workflow is script- or workflow-based so that the workflow can be re-executed using minimal actions.

10 Review
A series of managed activities needed to ensure continued access to and functionality of the research compendium and its components for as long as necessary.

Sharing the scripts that you used for your analyses and statistical queries, along with the data, in a trustworthy repository, enhances reproducibility, provided that it is documented in such a way that others understand what is computed where, and why certain decisions were made in the analyses.

If analysis code plays a central role in research, it is advisable to use version-control software to track changes to the code over time. From an Open Science perspective, Git is particularly appealing for version control because it retains a complete historical backlog of all changes (commits). Users can also go back to a previous version of the code (for example, after making a mistake). The functionality of Git is enhanced by services such as GitHub or GitLab. GitHub is best understood as cloud storage with social networking functionality. You can use GitHub as a 'living research archive' or as a cloud backup.

The social-network aspect of GitHub comes into play when a repository is made 'public': This allows other researchers to see how the work was done, or to copy it to build on your work. GitHub excells as a collaborative tool, because people can submit changes to the text or code for your consideration, or simultaneously work on different parts of a project, and their changes can be compared and merged.

Useful Tools
The article 'A Workflow for Open Reproducible Code in Science' is a good starting point for getting acquainted with reproducibility in your daily practice. The usage of Git and GitHub fits seamlessly in an Open Science workflow.

Check out the Executable Research Articles (ERA) in the journal eLife. On their website, you will find a 'launch' button. Pressing the button will reanalyse the data in real time and include the results of all computations in the document, including all figures and statistical analyses. The format allows you to review and adapt the code to rerun your own analyses as you see fit. A list of currently available ERAs can be found here.

Open Science > A Practical Guide for Early-Career Researchers

3. Collect & Analyse Data

Research Software

The next things to consider are **software**, **algorithms**, and **code**. What software is required to reproduce your work? Is that software openly available? Often, analysis depends on particular versions of software, and there may be dependencies between particular versions of software packages that are used in analyses. Will these analyses still work years from now? This is not automatically the case. It is, therefore, important to describe the version numbers of all software used in your analyses.

So, let's say you put in the effort to make your work reproducible. How do you know if you did a good job? The proof of the pudding is in the eating! Give your research package (your data, analysis scripts, software, and instructions) to a colleague and see if they can reproduce your work or reach out to the Data Steward.

Photo by Gery Wibowo on Unsplash

Data Stewards

The information above may seem daunting, especially if you have just started your own research. The good news is that you are not on your own. Universities have experts on these topics who are there to help you (e.g., see this overview). Most often, these people are called Data Stewards, but they sometimes go by the name of Research Data Managers or Research Data Specialists. They may be located at your faculty or at the University Library, or they may be part of your local Digital Competence Centre. Don't hesitate to make use of their expertise!

Open Science Reproducibility Checklist

- *Is your data available to others?*
- *Is 'reproducibility' included in your Data Management Plan?*
- *Does your research folder include a 'readme' file, explaining the context of the research, the file structure, and procedures?*
- *Does your research folder include a codebook, explaining all variables in your research?*
- *Are your analysis steps or scripts well described and available to others?*
- *Does your project minimise the effort it takes to reproduce your research?*
- *Is your research software available to others by attribution of a licence?*
- *Will your results be reproducible in the future?*
- *Has the reproducibility of your work been verified by a colleague?*
- *Do you know whom to contact if you need help with reproducibility?*

4. Write & Publish

'Nothing has really happened until it has been recorded.'

> Virginia Woolf

From the perspective of Open Science, the preferred way to publish is Open Access. The Dutch government also strongly encourages researchers to publish **Open Access** (OA), and aims to reach the goal of one hundred percent OA of Dutch scholarly output in 2024. Dutch research funders, such as NWO and ZonMw, have already implemented the principles of the international cOAlition S in their grant requirements. This means that all future publications must be available in Open Access immediately upon publication.

The primary advantage of publishing OA is that anyone in the world can read your work for free and share and reuse it. This way, more readers will be able to find, read, and cite your work. You have many options to publish your work OA. Are you still working on a draft that needs to evolve and be discussed with peers? Then consider sharing a **preprint**. Or will your paper soon reach a state of completion? Then submit your manuscript to be published OA as a journal article, conference paper, or even a book.

Photo by Umberto on Unsplash

Preprints

It is often possible to openly share a preliminary version of your article online. These so-called preprints are "[academic] manuscripts that have not been peer reviewed or published in a traditional publishing venue" (Malički et al., 2020). By posting preprints, you can rapidly share your findings and start early and open academic discussions, which strengthens your work.

Useful Tools
Most journals allow you to submit your article when it is already posted as a preprint. The database Sherpa Romeo presents journal policies for preprints. Much more practical information on preprints can be found in the Practical Guide to Preprints.

Some might fear sharing preprints, as it may give others the opportunity to 'scoop' your topic or idea. This fear is unjustified. What happens instead, is that posting your preprint records your findings immediately!

Open Publications

Open-Access Journal Articles

When a journal publishes your work Open Access (OA), it is immediately and permanently open with a licence for reuse. This is known as **Gold Open Access**. As an author, you may get charged for OA publishing, in the form of an **Article Processing Charge** (APC). You may be able to pay APCs from your research-funding budget, or get a reimbursement for your department from your faculty or University Library.

Useful Tools
From the perspective of Open Science, the recommended licence for OA publications is CC BY, the most liberal Creative Commons licence. For practical guidance, see this Guide to Creative Commons for Scholarly Publications and Educational Resources.

Thanks to agreements between the universities in the Netherlands (UNL) and publishers, corresponding authors who are affiliated with a Dutch university can publish OA at no cost to them. The costs are already covered by the agreement. As a result, over 10.000 journals from publishers such as Elsevier, Taylor & Francis, and Springer offer you a full discount on their APC.

Useful Tools
The Open Access Journal Browser gives information about academic journals and their options for Open Access. Over 10.000 journals offer full discounts on OA publishing for authors of Dutch universities. When you submit an article to such a journal, always use your university email address to make sure you are eligible for the discount.

The drawback of APC-based publishing is that it gave rise to **predatory publishers** that are more interested in making money than in offering a sound peer-review process. When you consider publishing in an OA journal, always check if it is an authentic outlet for scholarly publishing.

Useful Tools
The Directory of Open Access Journals (DOAJ) indexes quality OA journals. Many of these journals do not charge APCs, but some do. If this happens to be the case, make sure your funder or department can fund the APCs before you submit your work.

To check the trustworthiness of a journal, use the tool Think. Check. Submit. Of course, you can always contact your University Library to help you out.

Many OA journals have a non-APC business model, also known as **Diamond Open Access**. Authors do not pay APCs for publishing in these journals. Instead, the costs are covered, for instance by institutions. Out of the almost 18.000 journals enlisted by DOAJ, over 12.000 do not charge APCs.

The Dutch platform openjournals.nl is an initiative to stimulate diamond OA. Openjournals hosts diamond journals and helps scholars who want to start a new diamond journal or turn an existing journal into one.

Making Closed Publications Openly Available

When you cannot find a suitable outlet to publish your work OA and you are forced to publish your article closed, the publisher will ask you to sign a **Copyright Transfer Agreement** (CTA). As a result, you are not allowed to immediately share and reuse your article as you please. Nonetheless, there are still ways to make your article openly available. You can share your work through your university's repository. You usually have to respect an embargo period and your work will probably not have a licence that allows others to reuse your article, but at least this way others can read your work and build upon it .

Making closed (traditional, non-OA) publications openly available in a repository is called **Green Open Access**. When you click on a journal's title in the Open Access Journal Browser and scroll down, you will see all your options for green OA, including **Taverne**, an amendment in the Dutch Copyright Act proposed by politician Joost Taverne. This amendment, Article 25fa of the Dutch Copyright Act, allows authors affiliated with a Dutch university to share the final published version of their articles, conference papers, and chapters in edited volumes in their university's repository. Authors only have to respect an embargo period of six months, after which they can share their work openly, regardless of having signed a CTA.

University Libraries have different workflows in place for supporting Taverne. Some will open up your work after six months by default; others will help you when you ask them to apply Taverne to your work. Contact your University Library for more information about conditions and implementation at your university.

Open Dissertations

All Dutch universities require PhD candidates to make their dissertation (PhD thesis) openly available by depositing a copy in their university's repository.

When you are finished writing your dissertation, you will likely be asked to submit a digital copy to the University Library several weeks before your defence. You will be asked to sign an agreement in which you can also request that your dissertation or certain chapters need to be placed under an embargo of six to twelve months, and sometimes even longer. During this period, your dissertation will only be available to the university community. This may be necessary when your dissertation is to be published as a monograph by a commercial academic publisher.

However, most academic publishers do allow you to openly share your dissertation, as it usually requires quite some reworking and editing to convert a dissertation into a publishable and marketable book. Almost every publisher will allow you to reuse published articles in your dissertation, or to submit new articles that are based on dissertation chapters.

Open Science > A Practical Guide for Early-Career Researchers

4. Write & Publish

Open Monographs

Monographs are a common way to publish research results in certain domains, such as Humanities and Law. Publishers often set high **Book Processing Charges** (BPCs) for monographs, in order to compensate for lost sales.

One can question whether this is justified; Research shows that a free and open online version has little effect on the sales of printed versions. There are ways, however, to publish OA at a lower cost.

It is often possible to only publish the digital version openly, while the printed version still needs to be paid for. Furthermore, dedicated academic OA book publishers deliver high-quality OA books and often ask for reasonable compensation. In some cases, they deploy a diamond model of publishing, which means they do not charge BPCs at all.

Useful Tools

The Directory of Open Access Books (DOAB) contains high-quality OA books and can also be used to find a trusted publisher. You can also use the Think. Check.Submit. tool for books and book chapters to review the trustworthiness of an OA book publisher.

Another option is to send your book proposal to one of the university presses based at University Libraries, such as Radboud University Press, University of Groningen Press, TU Delft OPEN Publishing or Open Press Tilburg University. These are mostly driven by scholars and library staff who aim to help you publish monographs and other academic content Open Access, without the need to make a profit.

If your project is funded by a research funder (such as NWO), you may be eligible for BPC funding. Some universities also offer (partial) BPC reimbursement. Contact your University Library to learn about your options.

Photo by Will H McMahan on Unsplash

Open & FAIR Research Data

'Data is the new gold.'

> Neelie Kroes

Apart from publishing Open Access, Open Science entails sharing your data sets as much as possible, and making your data FAIR: Findable, Accessible, Interoperable, and Reusable. Not all data - think of sensitive data - can be made openly available. A key principle is that data should be 'as open as possible and as closed as necessary'. In other words, data should be widely disseminated, but access restrictions may be justified in certain situations.

Sharing Your Research Data in a FAIR Way

The ultimate goal of the FAIR (Findable, Accessible, Interoperable, Reusable) principles is to optimise the reuse of data.

Findability facilitates the discovery of data by humans and computer systems. This requires the description and indexing of data and metadata. To do so, you should assign persistent identifiers (such as the Digital Object Identifier - DOI - and ORCID iD), provide rich metadata, use keywords, and use versioning of data sets.

Accessibility is about how to access the data, possibly including authentication and authorisation. Metadata should remain accessible even when the data is no longer accessible, so people, institutions or publications can be tracked down.

Interoperability means that the data can be integrated with other data. The data needs to interoperate with applications and workflows for analysis, storage, and processing. You should use formal, broadly applicable language, and use standard vocabularies and qualified references.

Reusability increases data reuse by specifying licences. You should use a licence to permit the widest reuse possible; specify a data embargo, if needed, and describe how long the data will remain reusable and describe data-quality-assurance processes.

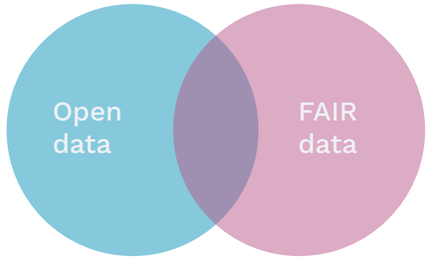
Many of the requirements to make your data FAIR will themselves be met by the chosen repository, for example a persistent identifier like the DOI.

Useful Tools
There is a tool to help you make your data FAIR: FAIR-Aware.

FAIR-Aware helps you assess your knowledge of the FAIR Principles. It also helps you understand how making your data (set) FAIR can increase the potential value and impact of your data (developed by DANS).

FAIR Data Is Not the Same as Open Data

Accessibility does not imply 'open for everyone'! Open data is data that is free to use, reuse, or redistribute. However, there are sound reasons why the openness of research data may need to be restricted, but any restrictions must be both justified and justifiable; in other words, open when possible, closed when necessary. Even when the raw data is restricted, the metadata of that data can still be made publicly available to flag that the data exists, and to enable others to contact the authors of that data.



Sharing Sensitive Data by Anonymisation

Your Data Management Plan should contain information about how you will handle sensitive data (see Data Management Plan).

It is possible to share sensitive data when you anonymise it. Anonymous or anonymised data cannot be linked back to individuals. The GDPR does not apply when all copies of the data set have been anonymised. Fully anonymising personal data is difficult, because i) a specific combination of even a few variables can increase identifiability, and ii) different data sets can be linked which increases identifiability. The processor of data must prove that identification will not be possible by current and even future technological methods.

It is also possible to create synthetic data, where the original data is replaced to prevent disclosure while the statistical features of the data are preserved. Synthetic data is fake data generated from real data and enables verifiability. To generate a synthetic data set, one takes a real dataset and models its distributions (e.g., shape and variance) and structure (e.g., correlations among the variables). One then uses this model to generate or synthesise the observations that make up the synthetic data set.

Useful Tools
Use tools to anonymise sensitive personal data (but keep the above caveat in mind), such as ARX – Data Anonymization Tool, or Amnesia that was developed by OpenAIRE.

Please consult the data supporter of your institute before you assume that your data is fully anonymised.

Disseminating Research Data

Why should you disseminate research data via a **trusted digital repository**? A few of the reasons are that making data open on the international level allows all researchers to reuse data sets produced by others, and the cost of creating, collecting, and processing data can be very high. If you put your data online, it helps increase the visibility of your work and enables you to be cited more often. Preserving and sharing research data

makes it possible to conduct comparative research, for example, if you want to study the socio-economic composition of a current population and compare it to the composition of that population 30 years ago.

Funders (e.g., Horizon Europe, the Dutch research funders NWO and ZonMw), Dutch universities, and University Medical Centres want to improve the openness of science, and increasingly require that data gets published and shared, when possible. These organisations ask for a Data Management Plan, where you describe if and how you (will) publish your research data. The national government has been providing subsidies for research into replacing, reducing, and refining animal testing for some time. One of the ways to do this is by having researchers share data from studies, so that fewer animals are needed for the testing.

Photo by ThisisEngineering on Unsplash

Keeping Data Available

A minimum retention period of 10 years is often stated in data policies and academic codes, but data can be valuable for ages, whether it be in climate research, sociology, human geography, health sciences, astronomy, or linguistics. Check the standard retention period in your domain; for instance, the NFU recommends that UMCs adhere to minimum periods of 15, 25, and 30 years for different kinds of data. Look beyond minimum retention periods where relevant. Select what data you will need and want to retain. Some selection criteria: Data underlying publications; data that cannot be recreated, like interviews or environmental recordings; what is potentially useful to others; and what has scientific, cultural, or historical value.

Long-Term Preservation

For long-term preservation, deposit your data in a data repository. You can use an external, disciplinary data archive or repository in your research domain to preserve the data according to the recognised standards in your discipline. In the Netherlands, you can use the data archives of DANS or 4TU.ResearchData. Alternatively, use your institutional repository, if you have one where the data will also be available in the long term, such as YODA (Utrecht University) or DataverseNL. You can also deposit your data sets in a catch-all repository such as Zenodo (CERN), Open Science Framework and Dryad, or you can search the global re3data.org portal, that provides several filtering options, for a fitting repository.

Choosing a Repository for Long-Term Preservation

If you want to make sure that your data will be kept safe, then check if the repository is certified as a trusted digital repository (for example, certified by CoreTrustSeal) with an explicit ambition to keep the data available in the long haul. Of course, the data repository should match your specific data needs (e.g., formats accepted, mixture of open and restricted access, licences). Look for a repository that provides guidance on how to cite the deposited data, and one that offers data depositing, data documentation, and support at a price that matches your budget. The repository should provide your submitted data set with a persistent and globally unique identifier for sustainable citations, and should link back to particular researchers and grants.

Useful Tools
See the OpenAIRE diagram about depositing your data in a data repository for long-term preservation: Where, Why, What, When, How, What to keep.

Choosing a Licence for Research Data

Each data repository has its own licensing options. Some repositories require you to use a certain licence if you want to deposit your data with them.

When you plan to reuse someone else's research data, remember that you can only use data when there is a licence or rights waiver attached, and that the Code of Conduct already requires you to cite the data properly, regardless of any licence.

The Creative Commons licences can also be used for research data. If your research data is a database or a data set (unstructured data that does not meet the database definition), the best option is usually CC0. This waives all your rights in the database.

For more information about licensing your research data, visit the website of OpenAIRE.

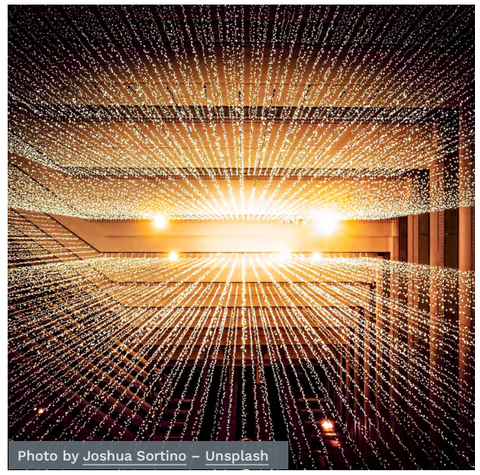


Photo by Joshua Sortino – Unsplash

Other Open Research Output

Open Education

Open education strives to open up academic education, and to take away barriers for those without easy access to educational materials in a variety of ways, such as by offering free, open courses online, or by publishing open textbooks and other educational materials.

Although education is usually seen as a different domain than research, many university scholars argue that it is impossible to separate education from science. Education provides input for scholarly activities, and vice versa, educational material is often the result of scholarly work. Both UNESCO and the Dutch government mention open education in one breath with open-science ambitions.

Many library-based OA publishers (mentioned above) can help you publish open handbooks **(open textbooks)**, and many universities nowadays also have facilities to create online courses (or even MOOCs: Massive Open Online Courses) that are released under an open licence. Eventually, this can prove to be cost-efficient: University Libraries need fewer funds to acquire expensive teaching materials and digital handbooks, and lecturers can reuse, repurpose and adapt material from others, so they will not have to (re)design a course from scratch.

Of course, open textbooks and MOOCs are the eye-catching end products of open education, but there are more open educational resources (OER) that you can share, such as slides, drawings, recordings, assignments, notes, and so on. They can all be shared on open educational platforms such as Edusources (owned by SURF/ Dutch institutions of Higher Education), OER Commons, the Open Textbook Library and Merlot.

Open Software

Whenever possible, it is good practice to share your software together with your data and code. There are tools and services available to store all current versions of your software (including your operating system and computing environment) for future use. A comprehensive overview of the tools and services that are currently available for these purposes can be found in a review article by Konkol, Nüst and Goulier (2020).

Note that the FAIR principles that were introduced earlier, in the context of research data, likewise apply to sharing software. Software can also be a research output in itself. You can check out these five recommendations to make your research software FAIR.

Choosing a Licence for Other Open Research Output

You can use Creative Commons licences for most of your output, such as presentations and educational materials. For guidance in adding a licence to your software, check the five recommendations to make your research software FAIR.

Open Science Write & Publish Checklist

- ***Does the journal of your choice allow preprints?***
  If so, are there suitable preprint servers for your work? If yes, post your work as a preprint.
- ***Does your funder require you to publish OA?***
  Check Sherpa Juliet and the grant requirements of your funder.
- ***Is it possible to publish OA in the journal of your choice?***
  Check the Open Access Journal Browser. If not and you do want to publish OA, select another journal.
- ***Can you cover the cost of OA publishing?***
  Does the journal charge APCs and if so, is it part of an agreement for a full discount? Check the Open Access Journal Browser. If not, will your funder, university, faculty or department cover the cost?
- ***In case of an OA journal, is it listed in the DOAJ?***
  If not, use the checklist on Think. Check. Submit. to find out if the journal is trustworthy.
- ***Have you uploaded your publications to your university's repository?***
  If not, do so. For help and more information, contact your University Library.
- ***Does your OA publication have a Creative Commons licence?***
  Make sure it does. Check the Guide to Creative Commons for Scholarly Publications and Educational Resources.
- ***Did you make a Data Management Plan?***
  If not, use an online tool like ARGOS or DMPonline.
- ***Have you made your research data FAIR?***
  Check this with the tool FAIR-Aware.
- ***Does your research data have a Creative Commons licence?***
  Make sure it does. The best option is usually CC0. For help and more information, contact your data supporter.
- ***Did you submit your research data in a trusted repository?***
  If not, do so. For help and more information, contact your data supporter.

5. Outreach & Assessment

'The public is the only critic whose judgement is worth anything at all.'

> Mark Twain

The ultimate goal of Open Science is to increase the scientific and societal impact of science. Opening up your data and output is pivotal, but you can do more to ensure that your research finds its way to others within and beyond academia.

Open Peer Review

Before you publish your final conclusions in a journal article or dissertation, you should allow others to engage with your preliminary works. Conventionally, journals employ a closed system of peer review, such as the single-blind and the double-blind model, which means that the identities of both the author(s) and reviewer(s) are anonymised. The reviews themselves are only made available to the editor of the journal and to the authors of the article.

The premise of single or double-blind review is that reviewers can be as autonomous and honest as possible when judging the quality of your article. Some open-science advocates envision a more open form of peer review where the discussion and interactions are public and, in some cases, the identity of the reviewers is disclosed. They argue that opening the black box of peer review would deepen academic discussion and result in more affable behaviour between reviewer and reviewee, and that peer-review reports gain value, because they can become a (micro) publication in their own right, findable and quotable as such.

There are a few progressive journals that disclose the identity of both the author(s) and the reviewers (open identities) or publish the peer review alongside the article (open reports); others use a platform for discussing your manuscript (open participation), or a combination of these practices.

Will open peer review improve the peer-review culture? Or will it instead cause bias (for instance because the reputation of the author will play a role in the reviewing process), and if so, will it be better to stick to a closed or blind mode of peer review? We still need to try out and build evidence in order to be conclusive.

Postpublication Peer Review

Open Science is also about stimulating discussion and engagement of work after publication. When colleagues openly comment on each other's publications, this is called postpublication peer review. More and more journals provide formats for postpublication peer review. Others, such as PubPeer and Hypothes.is, provide generic solutions for commenting or annotating any published work. These comments are visible to all who have activated the required browser plug-ins on their computers. Commenting on each other's work is considered a contribution to the field. Postpublication peer review also provides means to indicate whether others have been able to reproduce original findings, or to link to other related findings.

Science Communication

Your research matters. Not only to you, but also to colleagues in your field and others outside academia. The impact of your work does not only depend on its content; it also depends on how you communicate (about) your research. This holds true for both the scientific impact and the societal impact of your work. Publications in scientific journals and books are the default for communicating with your academic peers. But in many (or most) cases, research is also of interest to others, outside academia, such as policymakers, clinicians, teachers, industry, or citizens. In many cases, alternative outputs are more effective to reach and engage these stakeholders.

Open Science is therefore also about diversity in the formats of research output, such as blog posts, podcasts, videos, newsletter articles, and public speaking in front of non-academic audiences. After you have published your research findings, you want to make sure it reaches the people for whom it may be interesting. Traditionally, the journal in which you published your work largely determined who read your work. This, however, is changing. Social media provide ample opportunities to raise awareness and increase engagement with your work, both inside and outside academic bubbles. Indeed, articles featured on social media are more likely to be cited in the following years. Moreover, it is not just the citations that matter; other forms of impact are becoming increasingly important, as will be discussed next.

Rewards and Incentives for Open Science

The transition to Open Science is reflected in the way we reward and incentivise research. In order for it to become the norm, it is essential that researchers get rewarded when putting Open Science into practice.

By publishing the position paper Room for everyone's talent, all Dutch universities, University Medical Centres together with the Royal Academy, and funding councils NWO and ZonMw, launched a call to modernise how academics are being evaluated. Open Science plays a central role in this endeavour. A more balanced assessment is needed to do justice to the diversity of tasks that academics have. Less emphasis must be placed on the number of published papers, and more on societal impact and on education, public engagement, and (patient) care. This transition is an ongoing, international process. Recently, over 350 organisations from over 40 countries signed the CoARA agreement, which states that rewards and incentives in academia should reflect Open Science practices.

A first step towards this aim is to take a different look at the metrics through which we assess research and researchers. Traditionally, publishing in scientific journals of high prestige, with high **Impact Factors**, was considered to be the most accurate metric for research quality and impact. In Open Science, the focus no longer lies on the particular journal that you publish in, but on the research itself. Journal prestige is considered to be less important, especially since many excellent findings are published in outlets with less prestige

(and vice versa). That is why the Dutch government, universities, and research funders have signed the San Francisco Declaration of Research Assessment (DORA), stating that journal prestige should no longer be considered in evaluating research output. Instead, DORA states that a broad range of impact measures must be examined to evaluate research, including qualitative indicators of research impact. Moreover, the value and impact of all research output must be taken into account, including the reuse of your data, scripts and software.

Often, the impact of your work is not properly captured by metrics, and is better explained in words. This has been recognized by universities and research funders, which is why a narrative approach to research assessment, in addition to quantitative metrics, is becoming more popular, both for career advancement within universities and for assessment of grant applications (see the information page of NWO).

Research Funders' Policies for Open Science

NWO has a longstanding commitment to guarantee that the research it funds gets shared as widely and openly as possible. Funding requirements of NWO include that all publications arising from NWO funding must be made fully Open Access, without embargo, and with open licences allowing reuse. NWO also expects that research data resulting from NWO funding is shared as openly as possible and as FAIR as possible.

For these reasons, a Data Management Plan is requested. NWO also stimulates and rewards Open Science, for instance by making open-science practices visible in application procedures, by supporting open-science infrastructures, and by recognizing front-runners in Open Science through its Open Science Fund.

At the European level, the funding programme of the European Commission reinforces the open-science policy, introduced under the Horizon-2020 programme. This includes the obligation to ensure immediate Open Access to all publications and the sharing of data according to FAIR principles. Open Science is also one of the assessment criteria for projects under the current funding scheme Horizon Europe. Moreover, the European Commission is funding the European Open Science Cloud (EOSC), an initiative to create a giant web of FAIR linked data across Europe.

Open Science Outreach & Assessment Checklist

- *Who can potentially benefit from your research? How are you reaching them?*
- *How do you communicate your results to people outside academia?*
- *Are societal stakeholders involved in the process of your research?*
- *What are the criteria on which you will be evaluated? What are your criteria for success? To what degree are open-science practices part of these criteria?*
- *Is Open Science part of your research portfolio?*

6. Join the movement!

'I am going on an adventure!'

> Bilbo Baggins

Congratulations! You made it to the very end of this Open Science Guide. We assume that by now, you are all fired up and ready to put Open Science into practice! Here is a shortlist of open-science practices that you can implement, starting today:

- Consider **using existing data** in your (next) study; it may save you time and is very rewarding for the Open Science enthusiasts who shared this data.
- Specify your hypotheses and study design in a **preregistration**.
- Think of the **end users** of your research in society. How can you reach them? Can you engage with them early in the process of your research?
- Share your articles prior to publication as a **preprint**.
- Submit your articles to **Open Access** journals registered in the Directory of Open Access Journals (DOAJ).
- Deposit your publications in the institutional **repository** of your university.
- Write a **Data Management Plan (DMP)** to describe how you intend to handle your research data during and after your research project.
- Manage and share your research **data and software** according to the FAIR principles.
- Submit your research data in a **trustworthy repository**, but remember: open if possible, closed as necessary.
- Join your local Open Science Community. Open Science Communities are local hubs of people interested in Open Science. Here, you can learn from your colleagues how to put Open Science into practice and have some fun while you're at it!

Finally: Spread the Word!

You are now ready to put Open Science into practice and be part of the transition towards Open Science. But for Open Science to become the norm, we need everyone on board. That is why we invite you to spread the word about this guide on social media (Twitter, Mastodon, LinkedIn and so on), for example like this:

> **Twitter** @Twitter
>
> I just finished reading the Open Science Guide for early-career researchers (link). Ready to put #openscience into practice! #OSGuide

If you have any feedback on this guide, please leave your comments on PubPeer.

7. Attribution

This work is a derivative of "A Passport for Open Science. A practical guide for PhD students" written by Johann Berti, Marin Dacos, Gabriel Gallezot, Madeleine Géroudet, Sabrina Granger, Joanna Janik, Claire Josserand, Jean-François Lutz, Christine Okret-Manville, Sébastien Perrin, Noël Thiboud (Paris 2020) available through www.ouvrirlascience.fr and licensed under a CC BY-SA 4.0 International Licence. The original text has been cut, edited and mixed with original content focused on the Dutch research landscape.

This guide has been written by:

Loek Brinkman, DANS, OSC-NL.
loek.brinkman@dans.knaw.nl
orcid: 0000-0003-3997-1173

Elly Dijk, DANS.
elly.dijk@dans.knaw.nl
orcid: 0000-0002-5089-8774

Hans de Jonge, NWO.
h.dejonge@nwo.nl
orcid: 0000-0002-1189-9133

Nicole Loorbach, University of Twente.
n.r.loorbach@utwente.nl
orcid: 0000-0002-2766-8567

Daan Rutten, Tilburg University.
d.rutten@tilburguniversity.edu
orcid: 0000-0002-6645-9581

The authors would like to thank Marjan Grootveld (DANS), Vera Heininga (Groningen University), Melanie Imming (NPOS, Imming Impact), Bianca Kramer (Sesame Open Science), Caspar van Lissa (Tilburg University), Dan Rudmann (Leiden University) and Raúl Zurita-Milla (University of Twente) for their constructive feedback on earlier versions of this guide, and Meike Robaard for copy-editing the final version of the text.

April 2023

DOI
https://doi.org/10.5281/zenodo.7716153

Coordination
Daan Rutten, Tilburg University

Cover photo:
Shutterstock, Artemis Diana

Design:
Willem Dijkstra, www.inontwerp.nl