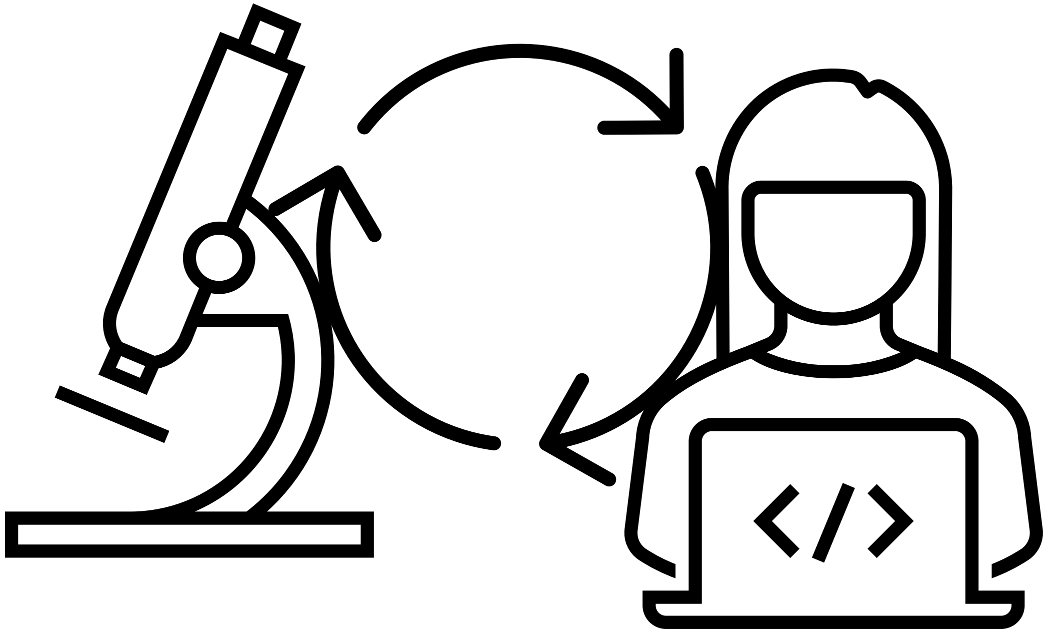
EUTOPIA TRAIN
Transforming Research and Innovation

Open and reproducible research

An introduction to practical measures (Version 1)



This programme has received funding fromthe European's Union's Horizon 2020 Framework programme under grant agreement N° 101017419 EUTOPIA-TRAIN.

VUB VRIJE UNIVERSITEIT BRUSSEL

CY CERGY PARIS UNIVERSITÉ

UNIVERSITY OF GOTHENBURG

University of Ljubljana

upf. Universitat Pompeu Fabra Barcelona

WARWICK

Contents

Background

- **Open Science has been in part a reaction to the 'reproducibility crisis' of science**. According to a study by Nature (2016), many researchers are indeed concerned that studies are too often deemed irreproducible or irreplicable. Later in 2018, the magazine dedicated an entire special issue to the theme "Challenges in irreproducible research".

- **Making data, code, software and other artifacts of the research processes accessible and re-usable is one way to combat these problems**, by achieving what can be called 'pre-reproducibility'.

- Open Science practices, which aim to make research processes more transparent and its outputs widely accessible, can provide some answers to the 'reproducibility crisis'. They offer an array of solutions, from tools and infrastructure to processes and workflows which seek to minimise the causes of irreproducible research. It can also mean tackling issues such as publication bias, lack of access to resources and results, changes to evaluation processes or reporting and dissemination actions (see e.g. Munafo et al . 2017).

- This guide presents an overview of **definitions of reproducibility** and **practical ways of ensuring reproducibility** of research results through Open Science. This includes **pre-registration and registered reports as mechanisms to reduce biases and improve study quality** as well as **different means of sharing workflows and protocols, software and code,** and **using computational tools and platforms**.

Reproducibility and replicability

Definitions of reproducibility and replicability can vary by domain and have been changing over time. Two widely used definitions are those brought forward by the *National Academies of Sciences, Engineering, and Medicine* (*NASEM 2019*) :

> ***Reproducibility*** *is obtaining consistent results using the same input data; computational steps, methods, and code; and conditions of analysis. This definition is synonymous with "computational reproducibility".*

> ***Replicability*** *is obtaining consistent results across studies aimed at answering the same scientific question, each of which has obtained its own data. Two studies may be considered to have replicated if they obtain consistent results given the level of uncertainty inherent in the system under study.*

Types of reproducibility

As noted by NASEM, their definition is about “computational reproducibility”. This is only one of **several types of reproducibility.**

Leonelli (2018) defines five different types: “Computational Reproducibility"; "Direct Experimental Reproducibility"; "Scoping/Indirect/Hypothetical Reproducibility"; "Reproducible Expertise"; "Reproducible Observation”.

Leonelli notes that reproducibility is **dependent on the research field, question, methods** and qualifies the five types based on dimensions such as the “assumed degree of control over research conditions”, the “dependence on statistics and computations”, the “precision of the research goals”, and “dependence on researchers’ judgement”.

Leonelli **warns against overly narrow or broad definitions of reproducibility** that do not capture the diversity of types and assumptions as shown in her typology (next slide).

Types of reproducibility

Source: Leonelli, S. (2018), "Rethinking Reproducibility as a Criterion for Research Quality", *Including a Symposium on Mary Morgan: Curiosity, Imagination, and Surprise* (*Research in the History of Economic Thought and Methodology, Vol. 36B*), Emerald Publishing Limited, Bingley, pp. 129-146. https://doi.org/10.1108/S0743-41542018000036B009

Type of Reproducibility	Assumed control	Dependence on statistics	Precision of goals	Dependence on judgement
Computational Reproducibility	total	high	high	none
Direct Experimental Reproducibility	high	high	high	low
Scoping/Indirect/Hypothetical Reprod.	limited	variable	limited	variable
Reproducible Expertise	variable	variable	variable	high
Reproducible Observation	low	low	low	high
Irreproducible Research	none	low	low	total

Types of reproducibility

These are not the only definitions of reproducibility. Others have proposed terms such as

methods reproducibility,

results reproducibility,

and *inferential reproducibility*.

The first would be close to the definition of reproducibility used by the National Academies of Sciences, Engineering, and Medicine, the second closer to replicability, and the latter describing the power of a study to "[make] knowledge claims of similar strength from a study replication or reanalysis" (Goodman et al. 2016).

Note: Reproduction and replication are two possible quality standards of good research among others – they are not the only ones, depending on the type of research or the specific field. These questions concern deeper debates about the philosophy of science or methodological rigour (see Penders et al 2019). Plus, a study not being reproduced or replicated might not automatically falsify its results. The German research funder DFG has published a position on replicability and reproducibility which emphasises such limitations.

There can be other problems related to reproducibility, such as lack of methodological or statistical training. Systemic flaws of research, such as the pressure to publish positive results, play a further role. And whether there is in fact a crisis of reproducibility and what it implies is a contested question in itself (see e.g., Fanelli 2018).

Resources and initiatives

Many initiatives, communities and networks are working to foster open and reproducible research:

- ReproducibiliTea is an international, bottom-up network of journal and discussion clubs. They maintain an active Zenodo community with many interesting documents.
- Framework for Open and Reproducible Research Training (FORRT) is a network sharing resources and training material on reproducible research, with a focus on teaching students about reproducibility.
- Reproducibility for Everyone (R4E) is an initiative organizing workshops and training on reproducibility.
- The UK Reproducibility Network (UKRN) is a national network on reproducibility in research with several sister networks in Germany, Sweden, Italy, and many other countries. They have developed a comprehensive primer on pre-registering research.
- rOpenSci is a *community* working on ensuring reproducible research through the R software.
- The FOSTER Open Science project has developed a self-paced course on open and reproducible research.
- The Carpentries are a *network providing training* on research data, software and library systems.
- The Turing Way is an *interactive handbook* for reproducible data science.

The paper "Reproducibility Starts from You Today" gives a range of suggestions how to ensure reproducibility (and inspired many contents of this guide). If you want to know reasons to work reproducibly beyond the yardstick of research quality and how it can benefit you as a researcher, we recommend the paper "Five selfish reasons to work reproducibly" (Markowetz 2015).

Pre-registration

Pre-registration aims to **prevent selection or publication bias by explaining research questions, hypotheses and methods in advance**.

Pre-registration

EUTOPIA TRAIN
Transforming Research and Innovation

- A pre-registration protocol is a detailed research plan (with information on the research questions, study design, hypotheses, methods and analysis) that is created before the data are collected or before analysis of already collected data is performed (Heers, 2020; Steward *et al.*, 2020; Banks *et al.*, 2020).
- A pre-registration protocol is given a timestamp and it cannot be edited anymore. When publishing the research results one may provide the pre-registration link to the reviewers so that they can check if the pre-registered plan was followed (Steward *et al.*, 2020).
- Usually, the pre-registered protocol is immediately made public or with an embargo period (Heers, 2020; Steward *et al.*, 2020; website).
- Some pre-registration platforms allow to keep the pre-registration private. Doing so might be interesting for sharing details with stakeholders (e.g. funders, government) without releasing the sensitive aspects. However, keeping the pre-registration private might negate some advantages of pre-registration.

Pre-registration of clinical trials conducted in the European Union or the European Economic Area is **mandatory** to increase the transparency and access to them for healthcare professionals and for the public (e.g. patients). This can be done in the European clinical trials register or respective national registers.

Pre-registration: advantages I

- Pre-registration **increases the reproducibility**[**and replicability** of findings, since others can find detailed information on the methodology and data collection to reproduce the study/analysis (Heers, 2020; website).
- With the timestamp, you can **take credit for your ideas**, **hypotheses and predictions,** even if somebody publishes a similar study, However, pre-registration **does not confer formal rights to exclusivity** (Steward *et al.*, 2020; website).
- You are building on a **positive reputation** that stands for openness and transparency (Steward *et al.* 2020). Readers of your final publication can check if you conducted the research as stated in your pre-registered protocol. Furthermore, they can judge for themselves if any deviation from the original plan is acceptable or not (website).

Pre-registration: advantages II

EUTOPIA TRAIN
Transforming Research and Innovation

- Pre-registering may help in **distinguishing between confirmatory and exploratory** research. Both types of research can be specified in the pre-registration protocol, although pre-registration of exploratory research is less well suited. Note that **it is still possible to deviate from the original pre-registration protocol**:
    1. You can perform additional analysis on the obtained data, which will be considered as exploratory research (see Nosek *et al.* 2018 for more on exploratory research).
    2. Deviations to the design and/or analysis might still be considered acceptable as confirmatory, as long as they are explicitly motivated in the final publication. A clear distinction between confirmatory and exploratory research needs to be made in the final manuscript. As such, researchers are protected against hindsight and confirmation bias, **increasing** their **credibility** (Heers 2020; Steward *et al.* 2020; Banks *et al.* 2020; Allen & Mehler 2019).
- While writing a pre-registered protocol can be **time-consuming** since the complete research needs to be planned upfront, it can benefit the research (Allen & Mehler 2019). In particular, writing this detailed protocol and/or answering the questions involved in a pre-registration encourages to reflect about the research plans and to **identify possible flaws**. Costly and hard to correct mistakes can be prevented (website).

Where can you pre-register research?

EUTOPIA TRAIN
Transforming Research and Innovation

Different online platforms can be used to pre-register a research plan, usually using a template to fill out, which can differ between the different platforms (Heers, 2020; Steward *et al.*, 2020). Popular preregistration platforms include:

- Open Science Framework (OSF): a general pre-registration platform managed by the Center for Open Science (COS). A description on how to create a preregistration in OSF can be found here. A collection by the Center for Open Science gives examples of pre-registration forms and templates. Filled out pre-registration forms, sorted by discipline and study type, can be found here.
- AsPredicted: a general pre-registration platform that allows a quick and easy registration. AsPredicted does not oblige you to make the pre-registration public after an embargo period. While keeping a pre-registration private might negate some of the advantages mentioned above, it might be useful for sensitive research topics (e.g., dual use), because there is a record of pre-registration which can be shared by the researcher with stakeholders (e.g. funders, government) without making the sensitive aspects public.
- PROSPERO: pre-registration platform for systematic reviews in health and social care, welfare, public health, education, crime, justice and international development, with a health related outcome.
- Evidence in Governance and Politics (EGAP): pre-registration platform for political science.
- Preclinicaltrials.eu: pre-registration platform for preclinical animal study protocols.
- The European clinical trials register or respective national registers.

Registered reports

A mechanism to **combat study biases and selective reporting**, which is also **addressing methodological questions before research is carried out** or the final publication is released.

Registered reports

Registered Reports, also known as reviewed pre-registration, is a model in which scientific publications goes through two stages of peer review (Chambers *et al.*, 2014; PhD on Track on Preregistration). A typical registered report will go through the following steps:

1. Before performing the study, a **stage 1 manuscript**, including an *Introduction* (background literature, results of preliminary work, theory and hypotheses) and a *Methods* (study procedures and analysis plan) section, is **peer-reviewed**.
2. Peer-reviewers check the research question(s), the hypotheses and the methodology to decide if they will accept the stage 1 manuscript, accept it pending revision to the study design/rationale or reject it (Chambers *et al.*, 2014; Chambers, 2019).
3. After peer-review of the stage 1 manuscript, the manuscript is offered **'In Principle Acceptance' (IPA)**, meaning that the *study will be published regardless of the outcome* as long as authors conduct the research according to the approved protocol (Chambers *et al.*, 2014; Chambers, 2019).
4. For a majority of journals, the stage 1 manuscript now needs **formal pre-registration** in a recognized repository, either publicly or under embargo (Chambers, 2019). Afterwards data can be collected and analysed.
5. This is followed by the submission of a **stage 2 manuscript**, including the *Introduction* and *Methods* section from the stage 1 manuscript plus *Results* and *Discussion* sections. The *Results* include the outcome of the analyses, as they were described in the pre-registration, together with possible additional unregistered analyses, clearly identified as exploratory analyses (Chambers *et al.*, 2014; Chambers, 2019).
6. The **original reviewers** check if authors followed the accepted pre-registration and after sharing the data on an (open access) repository, with exception in some cases (e.g. personal data, protection of intellectual property rights), the manuscript will be published (Chambers *et al.*, 2014).

Registered reports: advantages

EUTOPIA TRAIN
Transforming Research and Innovation

- Pre-registration **increases the reproducibility**[**and replicability** of findings, since others can find detailed information on the methodology and data collection to reproduce the study/analysis (Heers, 2020; website).
- With the timestamp, you can **take credit for your ideas**, **hypotheses and predictions,** even if somebody publishes a similar study, However, pre-registration **does not confer formal rights to exclusivity** (Steward *et al*., 2020; website).
- You are building on a **positive reputation** that stands for openness and transparency (Steward *et al*. 2020). Readers of your final publication can check if you conducted the research as stated in your pre-registered protocol. Furthermore, they can judge for themselves if any deviation from the original plan is acceptable or not (website).

Registered reports: advantages

- Pre-registering may help in **distinguishing between confirmatory and exploratory** research. Both types of research can be specified in the pre-registration protocol, although pre-registration of exploratory research is less well suited. Note that **it is still possible to deviate from the original pre-registration protocol**:
    1. You can perform additional analysis on the obtained data, which will be considered as exploratory research (see Nosek *et al.* 2018 for more on exploratory research).
    2. Deviations to the design and/or analysis might still be considered acceptable as confirmatory, as long as they are explicitly motivated in the final publication. A clear distinction between confirmatory and exploratory research needs to be made in the final manuscript. As such, researchers are protected against hindsight and confirmation bias, **increasing** their **credibility** (Heers 2020; Steward *et al.* 2020; Banks *et al.* 2020; Allen & Mehler 2019).
- While writing a pre-registered protocol can be **time-consuming**, it can benefit the research (Allen & Mehler 2019). In particular, writing this detailed protocol and/or answering the questions involved in a pre-registration encourages to reflect about the research plans and to **identify possible flaws**. Costly and hard to correct mistakes can be prevented (website).

Registered reports: disadvantages

EUTOPIA TRAIN
Transforming Research and Innovation

- A Registered Report can be a **lengthy process**, since you need to go through several review rounds (Chambers & Tzavella, 2020).
- A **minority of journals** that publish Registered Reports do not oblige to publish the stage 1 manuscript in a recognized pre-registration repository, leading to a **lack of protocol transparency** (Chambers & Tzavella, 2020).
- Registered Reports cannot be used for all fields of science or all sub-disciplines within the fields, as it is particularly suited for hypothesis-driven studies but **not** to improve the robustness or transparency of **purely exploratory studies** (Chambers *et al.*, 2014; Chambers, 2019; Chambers & Tzavella, 2020). Even more, sometimes it is also not suited for hypothesis-driven research itself. For example, studies that try to capture the effects of unpredicted events (e.g. such as solar flares or stroke-induced brain injury) must start collecting data as soon as it is feasible and they cannot wait for the lengthy reviewing process (Chambers, 2019).

Where to publish Registered Reports?

- By 2022, over 300 journals offer the Registered Reports publishing format as a new article type. These journals are not only spread across disciplines but also across subscription and Open Access journals and publishers. A regularly updated list can be found on the website of the Center for Open Science (tab ‘participating journals’).

- A template on what to include in a Registered Report, together with some tips, has been prepare by the Center for Open Science.

Source: https://osf.io/tvyxz/wiki/home/ (License: CC-BY)

Workflows and protocols

Reproducing a study requires a thorough understanding of the **steps taken during an experiment or an analysis**.

While usually described in the methods section of studies, the open sharing of protocols or workflows helps preserving and presenting this information.

Workflows and protocols

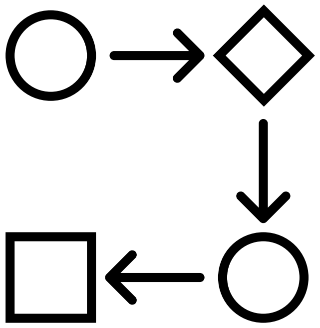
Some platforms for this purpose include:

- Protocol Exchange, an open repository for sharing scientific research protocols
- protocols.io, a platform for data management and protocol sharing
- myexperiment, a platform for sharing workflows

This may further include using open (lab) notebooks or shareable computational notebooks such as Jupyter Notebooks.

Other notable projects and platforms in an emerging stage are

- ResearchEquals, a platform allowing publishing different modules – or steps – of the research process in Open Access, and
- Octopus, an experimental platform supported by Jisc and the UK Reproducibility Network, which also aims to publish and connect problems, hypotheses, methods, data, analyses and interpretations in a modular way.



Sharing software, tools and code

The creation, management and analysis of research data is closely tied to the software and infrastructure.

Sharing software, tools and code

Initiatives supporting the sharing of research software include

- the Software Sustainability Institute, which lists a wide range of resources on the topic,
- the Software Carpentries, who organise training on research software management.
- The Research Software Alliance seeks to professionalise the field of research software engineers and works towards a better recognition of software as a research output.

Closely linked to the topic of software sharing are **efforts to apply the FAIR principles to research software** (see e.g. Hasselbring et al. 2020; Lamprecht et al. 2020; Katz et al. 2021).

> *"The core element of a sufficiently dynamic understanding of data quality is the precise documentation and disclosure of the measures, tools, the research software used and the procedural steps for generating, processing and making the data available."*
>
> The Data Quality Challenge (https://rfii.de/?p=4203, p. 72)

Software repositories, archives, papers and journals

Repositories and archives

Similar to publications and research data, repositories and services with an archive function can store and preserve research and other software for future use.

- Zenodo ('Issuing a persistent identifier for your repository with Zenodo') and Figshare ('How to connect Figshare with your GitHub account') can be used to archive GitHub repositories and make them citable.
- Savannah, SourceForge, and Launchpad are different hosting platforms or communities for software.
- Software Heritage operates an archive for software in source code form.

Software papers and journals

With software becoming a more recognized output of research, there are also **venues and journals where you can formally publish software or code**. The Software Sustainability Institute maintains a list of software journals across various disciplines.

In other cases, dedicated publishing platforms might offer an option to publish **software papers**. Open Research Europe and F1000 Research are examples for such services.

When working on a software paper, you can consider the "Ten simple rules for writing a paper about scientific software" by Romano & Moore (2020).

Software licences

Software sharing should be done with suitable licenses. These differ from the usual Open Access licenses which you would apply to a research paper or a dataset.

For research software you can follow the recommendations provided by Horizon Europe, meaning:

- Licenses listed as free by the Free Software Foundation
- Licenses listed as open source by the Open Source Initiative

The website choosealicense.com is also a helpful tool.

Note: Your Technology Transfer Office might offer support in licensing and sharing software.

Computational tools and platforms

Computational tools and platforms

- Some platforms integrate different tools, data sets, software and code in a single environment and ensure their proper documentation and preservation, going back to the idea of a reproducible 'research compendium' (see Gentleman & Lang 2020). This is mainly but not only relevant for research with a quantitative focus.

- To get started, you can consult the paper "Ten Simple Rules for Reproducible Computational Research" (Sandve et al., 2013).

- Again, there are many different solutions, platforms and resources, depending on the type of research and/or domain (see a comparison at "Publishing computational research - a review of infrastructures for reproducible and transparent scholarly communication" by Kokol et al., 2020). In many cases, these will require previous knowledge.

- On the next slides, we present a collection of different platforms and tools that broadly fall under this category.

Computational tools and platforms

EUTOPIA TRAIN
Transforming Research and Innovation

Codeocean (www.codeocean.com)	Codeocean lets you create and share 'capsules' of research data, software and code. These capsules are archived and given persistent identifiers, so that others can not only read but also copy and re-use the capsules content. Two examples are available at https://codeocean.com/capsule/8235972/tree and https://codeocean.com/capsule/9155944/tree/v1. A video summary can be watched at https://www.youtube.com/watch?v=aiyOa3n1Bwc.
Galaxy (https://galaxyproject.org/)	Galaxy is an analysis platform combining data management, workflows, and interactive research environments. The focus is on life sciences and biomedical data. Galaxy is an open-source tool. For introductions and training into Galaxy, go to https://training.galaxyproject.org/training-material/topics/introduction/
Reana (https://reanahub.io/)	Reana is an analysis platform developed by CERN researchers and therefore originally set up with high-energy physics research in mind. It allows you to create data analysis pipelines with remote computing clouds. For documentation, visit https://docs.reana.io/getting-started/.
The Turing Way (https://the-turing-way.netlify.app/welcome.html)	The Turing Way is an online learning resource for open data science, with an active community behind it. It contains tutorials and explanations how to make data science projects open and reproducible. It also features content on research communication, project organization, ethics and collaboration.
Project Jupyter (https://jupyter.org/)	Project Jupyter provides you with environments to work with and share data, code and notebooks. The main feature, Jupyter Notebook, is a widely used web application to create computational documents. JupyterLab gives you a more powerful environment to create notebooks and other components of your project.

Computational tools and platforms

Binder (https://mybinder.org/)	Binder is a tool that allows you to execute Jupyter Notebooks stored on a GitHub repository by building a Docker image. A tutorial is available at https://github.com/Build-a-binder/build-a-binder.github.io/blob/master/workshop/10-zero-to-binder.md and via the Turing Way.
Docker (https://www.docker.com/)	Docker is a tool based on containers, which create a "standard unit of software that packages up code and all its dependencies so the application runs quickly and reliably from one computing environment to another". In other words, Docker containers will always run in the same way, independent on the underlying infrastructure. Some advantages of using Docker are summarized in "An introduction to Docker for reproducible research" (Boettiger 2015) and tips given at "Ten simple rules for writing Dockerfiles for reproducible data science" (Nüst et al. 2020). Docker documentation and guides can be found at https://docs.docker.com/. The Turing Way also covers Docker containers (https://the-turing-way.netlify.app/reproducible-research/renv/renv-containers.html).
Wholetale (https://wholetale.org/)	Wholetale is an online platform for reproducible computational research projects. It has been funded by the US National Science Foundation and combines data, code, and a software environment into so-called executable research objects (or 'tales').
rOpenSci (https://ropensci.org/)	A community developing R packages specifically aimed to enable reproducible and open research.
ReproZip (https://www.reprozip.org/)	An open-source tool that helps packaging your data, code and software environment for sharing, reproduction and re-use. More information including some examples can be accessed via https://www.reprozip.org/about.html.

Annex: Disciplinary aspects

Discussions about reproducibility and/or replicability have emerged in many disciplines and it is impossible to summarise all of them, since they are deeply linked to methodological and epistemological debates and differences.

Often, such discussions arise within domains with quantitative approaches, but it is not limited to them. We have collected a few examples of papers discussing reproducibility from different perspectives a variety of domains.

Disciplinary aspects

Robotics	Bonsignorio, Fabio. "A New Kind of Article for Reproducible Research in Intelligent Robotics [From the Field]." IEEE Robotics & Automation Magazine 24, no. 3 (September 2017): 178–82. https://doi.org/10.1109/MRA.2017.2722918.
High-energy physics	Chen, Xiaoli, Sünje Dallmeier-Tiessen, Robin Dasler, Sebastian Feger, Pamfilos Fokianos, Jose Benito Gonzalez, Harri Hirvonsalo, et al. "Open Is Not Enough." Nature Physics 15, no. 2 (February 2019): 113–19. https://doi.org/10.1038/s41567-018-0342-2.
Analytical chemistry	Dryden, Michael D. M., Ryan Fobel, Christian Fobel, and Aaron R. Wheeler. "Upon the Shoulders of Giants: Open-Source Hardware and Software in Analytical Chemistry." Analytical Chemistry 89, no. 8 (April 18, 2017): 4330–38. https://doi.org/10.1021/acs.analchem.7b00485.
Preclinical cancer biology	Errington, Timothy M, Alexandria Denis, Nicole Perfito, Elizabeth Iorns, and Brian A Nosek. "Challenges for Assessing Replicability in Preclinical Cancer Biology." Edited by Peter Rodgers and Eduardo Franco. ELife 10 (December 7, 2021): e67995. https://doi.org/10.7554/eLife.67995.
Computer science/e-science	Freire, Juliana, Norbert Fuhr, and Andreas Rauber. "Reproducibility of Data-Oriented Experiments in e-Science (Dagstuhl Seminar 16041)." Edited by Juliana Freire, Norbert Fuhr, and Andreas Rauber. Dagstuhl Reports 6, no. 1 (2016): 108–59. https://doi.org/10.4230/DagRep.6.1.108. Peng, Roger D. "Reproducible Research in Computational Science." Science (New York, N.Y.) 334, no. 6060 (December 2, 2011): 1226–27. https://doi.org/10.1126/science.1213847.
Cognitive neuroscience	Gilmore, Rick O., Michele T. Diaz, Brad A. Wyble, and Tal Yarkoni. "Progress toward Openness, Transparency, and Reproducibility in Cognitive Neuroscience." Annals of the New York Academy of Sciences 1396, no. 1 (2017): 5–18. https://doi.org/10.1111/nyas.13325.
Geography and geosciences	Nüst, Daniel, and Edzer Pebesma. "Practical Reproducibility in Geography and Geosciences." Annals of the American Association of Geographers 111, no. 5 (July 29, 2021): 1300–1310. https://doi.org/10.1080/24694452.2020.1806028. Kedron, Peter, Wenwen Li, Stewart Fotheringham, and Michael Goodchild. "Reproducibility and Replicability: Opportunities and Challenges for Geospatial Research." International Journal of Geographical Information Science 35, no. 3 (March 4, 2021): 427–45. https://doi.org/10.1080/13658816.2020.1802032.

Disciplinary aspects

EUTOPIA TRAIN
Transforming Research and Innovation

Chemistry	McAlpine, James B., Shao-Nong Chen, Andrei Kutateladze, John B. MacMillan, Giovanni Appendino, Andersson Barison, Mehdi A. Beniddir, et al. "The Value of Universally Available Raw NMR Data for Transparency, Reproducibility, and Integrity in Natural Product Research." Natural Product Reports 36, no. 1 (January 25, 2019): 35–107. https://doi.org/10.1039/C7NP00064B.
Economics	McCullough, B. D. "Open Access Economics Journals and the Market for Reproducible Economic Research." Economic Analysis and Policy 39, no. 1 (March 1, 2009): 117–26. https://doi.org/10.1016/S0313-5926(09)50047-1. Vlaeminck, Sven, and Felix Podkrajac. "Journals in Economic Sciences: Paying Lip Service to Reproducible Research? Paying Lip Service to Reproducible Research?" IASSIST Quarterly 41, no. 1–4 (December 10, 2017): 16–16. https://doi.org/10.29173/iq6.
Health/machine learning	McDermott, Matthew B. A., Shirly Wang, Nikki Marinsek, Rajesh Ranganath, Luca Foschini, and Marzyeh Ghassemi. "Reproducibility in Machine Learning for Health Research: Still a Ways to Go." Science Translational Medicine 13, no. 586 (March 24, 2021): eabb1655. https://doi.org/10.1126/scitranslmed.abb1655.
Psychology	Nosek, Brian A., Tom E. Hardwicke, Hannah Moshontz, Aurélien Allard, Katherine S. Corker, Anna Dreber, Fiona Fidler, et al. "Replicability, Robustness, and Reproducibility in Psychological Science." Annual Review of Psychology 73 (January 4, 2022): 719–48. https://doi.org/10.1146/annurev-psych-020821-114157.
Evolutionary biology	O'Dea, Rose E., Timothy H. Parker, Yung En Chee, Antica Culina, Szymon M. Drobniak, David H. Duncan, Fiona Fidler, et al. "Towards Open, Reliable, and Transparent Ecology and Evolutionary Biology." BMC Biology 19, no. 1 (April 9, 2021): 68. https://doi.org/10.1186/s12915-021-01006-3
Behavioural research	Smaldino, Paul E., and Richard McElreath. "The Natural Selection of Bad Science." Royal Society Open Science 3, no. 9 (n.d.): 160384. https://doi.org/10.1098/rsos.160384.

Contributors

This document is based on earlier VUB material on registered reports and pre-registration prepared by Jone Paesmans (Vrije Universiteit Brussel).

Niek van Wettere (Vrije Universiteit Brussel), Jone Paesmans (Vrije Universeit Brussel), Lennart Stoy (Vrije Universiteit Brussel), Rangnar Nilsson (University of Gothenburg), Pieter De Bruyn (Vrije Universiteit Brussel) have contributed to various versions of the document.

License

Available on Zenodo at https://doi.org/10.5281/zenodo.7418817

This programme has received funding fromthe European's Union's Horizon 2020 Framework programme under grant agreement N° 101017419 EUTOPIA-TRAIN.