Horizon Europe Guidance Document for the University of Twente (UT)

Aim of this document:

- This document serves as a guide with examples of how you could address sections falling under Part B -section 1.2. Methodology, namely; “Open Science” and “Research data management and management of other research outputs” of the **Horizon Europe Programme Standard Application Form** (Part B pages 7-9 af_he-ria-ia_en.pdf (europa.eu).
- Please use examples in this guidance document for inspiration and adapt according to your own research project.
- The templates for a specific call may slightly differ, please check your application form for the specifical call you are applying to (pg.2 af_he-ria-ia_en.pdf (europa.eu))

Please contact your faculty data steward and/or open access team for any questions and faculty-specific support & advise.

GUIDANCE FOR PART B (*Part B pg. 8 af_he-ria-ia_en.pdf (europa.eu)*)

1. Excellence

1.2. Methodology

- ***Describe how appropriate open science practices are implemented as an integral part of the proposed methodology. Show how the choice of practices and their implementation are adapted to the nature of your work, in a way that will increase the chances of the project delivering on its objectives [e.g. 1 page]. If you believe that none of these practices are appropriate for your project, please provide a justification here.***

Early and open sharing (Recommended)

Also check Horizon Europe Programme Guide Pages 42-43

(If applicable) Pre-registration: "We will pre-register and publish our research plan on the Open Science Framework with specific details such as study information, sampling plan, variables, design plan, and analysis plan. In this document, we will provide our research plan regarding the research hypotheses with further details on the sections listed above such as existing data, data collection procedures, sample size, measured variables, statistical analysis, scripts."

(if applicable) Registered reports: "We will submit [a registered report/our pre-registration document] to [an appropriate journal] to allow peer-reviewing of our study designs and proposed methods, prior to data collection. "

Preprints: "We will submit the publications to [preprint server, more info on preprints] prior to journal submission for formal peer-review, to ensure that project output will be made accessible as widely and as early as possible."

Reproducibility of research output (Mandatory)

(if doing pre-registration/registered reports): "By pre-registering research design and methodologies prior to data collection, we ensure that feedback can be provided to our research methods and questions early (e.g. potential duplication, suggestions for improvement) and that they are accountable."

Sharing negative results: "We will publish all negative results and orphan data generated from the project in the appropriate venues, e.g. 4TU.ResearchData for datasets, [preprint servers] and/or journal papers, so that others can also learn from these results."

(if applicable) Open protocols: "To increase reproducibility and reusability of our research, we will share our experimental protocols [and details of reagents] openly on [protocols.io (or other platforms)] as

soon as it has been tested, to allow the wider community to not only more easily use, but also help validate and offer feedback on our experimental protocols."

(if applicable) Electronic lab notebooks: "We will use [RSpace/eLAB Journal/other ELN systems] to ensure good record keeping throughout the project, and to allow the notes to be more easily shared, accessed and archived."

Statistical analysis: "When determining sample sizes, variables to measure and statistical models to apply, we will take extra care to determine the appropriate sample sizes to give robust results, check model assumptions and rigorously document the statistical inference process, including clear documentation about concepts, variables, and indicators in the statistical analysis [using Jupyter notebook/RMarkdown/other] to improve reproducibility."

(if applicable) Open-source software development: "We will develop our [research software/data analysis workflow] openly on [Gitlab/Github/Gitea] to make these assets available and transparent during research, and to collaboratively develop our codebase with the community. Careful attention will be paid to ensure that sufficient documentation and testing is in place to allow others to use, build and test the software, and increase reproducibility and reusability of the work. We will also provide example parameters and input/output data for testing purposes."

(if applicable) Open-source hardware: "We will develop our [open-source hardware] openly on [Gitlab/GitHub] to make the design available and transparent during and after research, and to collaboratively develop the hardware design with our community. Detailed documentation will be developed to ensure that others can easily study, modify, distribute and make the design/hardware based on the design."

FAIR data: "As also explained in detail in the "Research data management" section, [the datasets, (if applicable) including software] underlying the published papers resulting from the project will be made openly available through the [4TU.ResearchData/DANS repository] together with necessary documentation for others to validate, reproduce or reuse the research outputs, in line with the FAIR principles."

Open access to research outputs (Mandatory)

Open access research article: "We will share the all peer-reviewed scientific publications resulting from this proposal open access under CC-BY license by the time of publication to allow [researchers, citizens, industry, governmental organisations, NGOs, etc.] worldwide to benefit from and build upon this work. Where needed, please check Implementing rules on intellectual property (utwente.nl)"

(If applicable) Open access monographs: "All monographs resulting from this proposal will be published open access under [the appropriate Creative Commons license, e.g. CC BY-NC, CC BY-ND, CC BY-NC-ND or equivalent], to allow [researchers, citizens, industry, governmental organisations, NGOs, etc.] worldwide to benefit from and build upon this work. Where needed, the UT library will assist us in retaining sufficient intellectual property rights during manuscript submission to achieve the aforementioned.

Open data, protocols, software, models, algorithms, hardware, and other types of research outputs: **"**Research outputs [Data, software, models, algorithms, hardware, etc.] produced in this project will be made publicly available at [4TU.ResearchData/DANS repository], at [the latest by the end of the project *(having a specific time period (e.g. at the end of each WP) would strengthen the application)*]. [4TU.ResearchData/DANS repository] is a trusted and certified research data repository. Research artefacts will be published with the appropriate open license to allow reuse and adaptations from others to widen accessibility and impact of the proposed work. Protocols used/produced will be shared on [protocols.io].

(In case of commercially-sensitive data) Some of the [data/software/hardware/other research output] has a commercially-sensitive nature as [it belongs to a third party/it is related to a patent application/etc.] and therefore is not suitable for public sharing. Please see the data management plan for the handling, sharing and storing of data, software or other artefacts not suitable for public sharing. For all of the research output, the DOI link to the research outputs will be shared under the Data Availability or Supplementary Data section (depending on the journal) of the corresponding academic papers."

(In case of personal data) Some of the [data/other research outputs] is sensitive because it contains personal data and therefore is not suitable for public sharing and should be handled accordingly.

Open Peer review (Recommended)

Scientific publications (preprints, journal articles, book chapters, etc):
Where [publishers *(give specific names if you are thinking of publishing in venue(s) that offer open peer review)*] offer the choice between open and close peer review for our submissions, we will opt for open peer review to increase transparency of and trust in our research and the quality of the peer review process. We will submit our preprints and academic papers to [a platform] that supports open peer-review.

Conference papers: We will submit our conference papers to conferences that support open peer-review[1].

(if applicable) Software: We will submit our software to a venue that enables open software review, such as the Journal of Open Source Software.

Participate as reviewers in open peer review processes: With insight and knowledge generated from this project, our team members strive to participate in open peer review of other works in the field, to support increasing trust in the peer review and scientific publishing system.

[1] For a (non-exhaustive) list of computer science, machine learning, etc conferences that have an open peer review system, please see https://openreview.net

Citizen, civil society and end-user engagement (Recommended)

Also check Horizon Europe Programme Guide Pages 52-54
For further support, please contact Citizen Science Hub Twente utwente.nl/nl/designlab/citizen-science/

GUIDANCE FOR PART B (*Part B pg.9 af_he-ria-ia_en.pdf (europa.eu)*)

1. Excellence

1.2. Methodology

- ***Research data management and management of other research outputs: Applicants generating/collecting data and/or other research outputs (except for publications) during the project must provide maximum 1 page on how the data/ research outputs will be managed in line with the FAIR principles (Findable, Accessible, Interoperable, Reusable), addressing the following (the description should be specific to your project): [1 page]***

- Please contact your faculty data steward for any questions and faculty-specific support & advise

MANDATORY

Types of data/research outputs

Add:
Reuse of existing data (if applicable, please explain the source and terms of use):
We will re-use existing data from […]. [As we are the creators of the data, there are no restrictions for us regarding reuse., The data is published with a CC-BY license. Accordingly, we will give proper credit to the original creators., Data is available under a data processing agreement. Accordingly, we will …, …].

Data collection method:
[Experimental/ observational/ images/ text/ numerical …] data will be collected.

The estimated size:
The estimated size is [less than 5GB, more than 5 TB, …]

Also add:
The research data gathered in this project will be described in a Data Management Plan (DMP) in the first six months of the project by consulting the Faculty Data Steward. The DMP will be regularly updated during the course of the project.

[F]indability of data/research outputs

The datasets underlying the published papers resulting from the project will be made openly available through the [4TU.ResearchData, DANS, Zenodo] repository, accompanied by rich metadata (adhering to DataCite metadata standard) to ensure findability and be assigned a Digital Object Identifier (DOI) to make them citable and persistently available.

[A]ccessibility of data/research outputs

Choose the relevant answer(s):

Situations where all data can be shared publicly (no confidential information):
The datasets underlying the published papers resulting from the project will be made openly available through the [4TU.ResearchData, DANS Zenodo, …] repository at the time of the publication of corresponding research papers resulting from this study.

Situations where not all data is suitable for public sharing because they contain personal data:
All raw data will be retained for […] on AREDA (ARchive REsearch Data at the UT) for the purposes of validation, with consent from the participants.

All anonymised or aggregated data, and/or all other non-personal data will be uploaded to [4TU.ResearchData, DANS Zenodo, …] with public access, with consent from the participants.

All pseudonymised data will be uploaded to 4TU.ResearchData with restricted access, with consent from the participants.

Situations where not all data is suitable for public sharing because they contain commercially-sensitive information:
Some of the data has a commercially-sensitive nature as ["it belongs to a third party'', "it is related to a patent application", …] and therefore is not suitable for public sharing.

The non-commercially sensitive datasets underlying the published papers resulting from the project will be made openly available through the [4TU.ResearchData, Zenodo, …] repository at the time of the publication of corresponding research papers resulting from this study.

If applicable add:

Data related to patents:
The data is related to a patent application and will be made publicly available once the patent application is complete.

[I]nteroperability of data/research outputs

Whenever possible, open/common file formats will be used and proprietary formats will be converted into open/common formats to make the data interoperable and reusable. A dictionary of all terms used in our datasets which will be publicly shared together with all the datasets at the end of the project in [4TU.ResearchData, DANS, Zenodo, …]. [4TU.ResearchData, DANS, Zenodo] uses Dublin Core Metadata Initiative (DCMI) and DataCite Metadata Schema as the standard for metadata which are used worldwide and increase the interoperability.

[R]eusability of data/research outputs

All datasets will be licensed under a [CC 0, CC BY] license which requires attribution/credit for the original creation, while at the same time ensures broadest possible re-use. Additionally, any info that is necessary for others to validate, reproduce or reuse the research outputs will be shared along in a README file.

If applicable add:

Code on GitHub/GitLab shared via 4TU.ResearchData

The developed software and codes presented in academic papers will be shared on [GitHub/GitLab] and those [GitHub/GitLab] repositories will be published via 4TU.ResearchData, be made publicly available to anyone for re-use under an open license and assigned a Digital Object Identifier (DOI) to make them citable and persistently available.

Code on GitHub shared via Zenodo

The developed software and codes presented in academic papers will be shared on [GitHub] and those [GitHub] repositories will be published via Zenodo, be made publicly available to anyone for re-use under an open license and assigned a Digital Object Identifier (DOI) to make them citable and persistently available.

Curation and storage/preservation costs

Data and code shared via 4TU.ResearchData

4TU.ResearchData is able to archive 1TB of data per researcher per year free of charge for all UT researchers and that data is kept for 15 years. We do not expect to exceed this and therefore there are no additional costs of long term preservation. Each work package leader will be responsible for ensuring data management activities and quality assurance.

DANS is able to archive 50 GB of data per individual researcher, free of charge and there is no size limit on communities. We do not expect to exceed this and therefore, there are no additional costs of long-term preservation. Each work package leader will be responsible for ensuring data management activities and quality assurance.

Data and code shared via Zenodo

Zenodo is able to archive 50GB of data per submission free of charge for all users and there is no size limit on communities. We do not expect to exceed this and therefore there are no additional costs of long term preservation. Each work package leader will be responsible for ensuring data management activities and quality assurance.

Data and code shared archived at Areda

All data, including data that contains personal and/or commercially-sensitive information, will be archived in the data archive of the University of Twente, AREDA, which is free of charge for its employees, and data is kept for at least 10 years.

If applicable add: The dedicated data manager hired in the project will be responsible for data management in the project. ###