FWO DMP Template - Flemish Standard Data Management Plan **stoelenstoe**

Project supervisors (from application round 2018 onwards) and fellows (from application round 2020 onwards) will, upon being awarded their project or fellowship, be invited to develop their answers to the data management related questions into a DMP. The FWO expects a **completed DMP no later than 6 months after the official start date** of the project or fellowship. The DMP should not be submitted to FWO but to the research co-ordination office of the host institute; FWO may request the DMP in a random check.

At the end of the project, the **final version of the DMP** has to be added to the final report of the project; this should be submitted to FWO by the supervisor-spokesperson through FWO's e-portal. This DMP may of course have been updated since its first version. The DMP is an element in the final evaluation of the project by the relevant expert panel. Both the DMP submitted within the first 6 months after the start date and the final DMP may use this template.

The DMP template used by the Research Foundation Flanders (FWO) corresponds with the Flemish Standard Data Management Plan. This Flemish Standard DMP was developed by the Flemish Research Data Network (FRDN) Task Force DMP which comprises representatives of all Flemish funders and research institutions. This is a standardized DMP template based on the previous FWO template that contains the core requirements for data management planning. To increase understanding and facilitate completion of the DMP, a standardized **glossary** of definitions and abbreviations is available via the following link.

1. General Project Information	
Name Grant Holder & ORCID	**Markus Wöhr (0000-0001-6986-5684)**
Contributor name(s) (+ ORCID) & roles	
Project number[1] & title	C1 - 3H220294 – Linking genetic alteration to brain development and behavior in a mouse model for Autism Spectrum Disorder
Funder(s) GrantID[2]	3H220294
Affiliation(s)	☒ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel v Other: Provide ROR[3] identifier when possible:
Please provide a short project description	Autism Spectrum Disorder (ASD) is among the most heritable neuropsychiatric dysfunctions and increasing evidence indicates that abnormal development of cortical inhibitory circuits and the disruption of excitation and inhibition balance are central to disease pathogenesis. While available evidence points to a complex set of genetic factors, Reelin (RELN) has emerged as a key candidate gene. Yet the exact relationship of how genetic mutations relate to circuit formation and behavioral alterations relevant to ASD remain completely unknown. Due to the multifaceted functions of Reelin in different cell types, we still do not know which cell type is most affected and drives changes in circuits and behaviors that are relevant to ASD. The lack of a clear (endo)phenotype significantly hinders the development of disease-modifying therapeutics. This project aims at bridging this gap by interrogating the function of a prominent ASD-associated gene, reelin, in a cell type specific manner, in order to map inhibitory cortical circuit development to behavioral changes.

[1] "Project number" refers to the institutional project number. This question is optional since not every institution has an internal project number different from the GrantID. Applicants can only provide one project number.

[2] Funder(s) GrantID refers to the number of the DMP at the funder(s), here one can specify multiple GrantIDs if multiple funding sources were used.

[3] Research Organization Registry Community. https://ror.org/

2. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data[4].

				Only for digital data	*Only for digital data*	*Only for digital data*	*Only for physical data*
Dataset Name	Description	New or Reused	Digital or Physical	Digital Data Type	Digital Data Format	Digital Data Volume (MB, GB, TB)	Physical Volume
Video files	Video files	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☒ Observational ☒ Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: avi or mp4 ☐ NA	☐ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☒ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Audio files	Video files	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☒ Observational ☒ Experimental ☐ Compiled/ aggregated data ☐ Simulation	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf	☐ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☒ < 5 TB	

[4] Add rows for each dataset you want to describe.

				data ☐ Software ☐ Other ☐ NA	☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: wav ☐ NA	☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Automated behavioral measurements	Automated behavioral measurements	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☒ Observational ☒ Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: xlsx ☐ NA	☒ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Biomaterial	Biomaterial	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☒ Observational ☒ Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt	☐ < 100 MB ☐ < 1 GB ☐ < 100 GB ☒ < 1 TB ☐ < 5 TB ☐ < 10 TB	

					☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: tiff ☐ NA	☐ < 50 TB ☐ > 50 TB ☐ NA	

Research activities in 3H220294 will result in a variety of data types. Data types include primarily: (1) video files (mouse behavior, avi-files or mp4-files, circa 4-5 TB, recorded with video cameras connected to a computer) and (2) audio files (mouse ultrasonic vocalizations, wav-files, circa 4-5 TB, recorded with ultrasound microphones connected to a computer). Additionally, (3) automated behavioral measurements (e.g. locomotor activity, social contacts) will be recorded in Excel (xlsx-files, circa 2-3 MB, recorded with infrared light sensors connected to a computer). Of note, recordings of mouse ultrasonic vocalizations are large in size, with one hour of recording resulting in about 2 GB. Video files will likewise result in a significant amount of data. Other datatypes will include biomaterials (e.g. DNA, RNA, protein, tissue samples, such as brain and tail samples), typically later converted into images (not expected to exceed 400 GB over the course of the project, details still need to be determined as ordering of relevant equipment is still ongoing). All data will be stored in digital form (e.g. AVI or MP4 files for video recordings, WAV files for audio recordings, TIFF files for gel images). Measurements derived from video and audio files, automated behavioral measurements, and biomaterial will be recorded in Excel (for long-term preservation, converted into CSV files), SharePoint Online and SPSS for statistical analyses. Manuscripts will be written in Word.

Guidance: *Data can be digital or physical (for example biobank, biological samples, …). Data type: Data are often grouped by type (observational, experimental etc.), format and/or collection/generation method.* *Examples of data types: observational (e.g. survey results, sensor readings, sensory observations); experimental (e.g. microscopy, spectroscopy, chromatograms, gene sequences); compiled/aggregated data[5] (e.g. text & data mining, derived variables, 3D modelling); simulation data (e.g. climate models); software, etc.* *Examples of data formats: tabular data (.por,. spss, structured text or mark-up file XML, .tab, .csv), textual data (.rtf, .xml, .txt), geospatial data (.dwg,. GML, ..), image data, audio data, video data, documentation & computational script.* *digital data volume: Please estimate the upper limit of the volume of the data per dataset or data type.* *physical volume: Please estimate the physical volume of the research materials (for example the number of relevant biological samples that need to be stored and preserved during the project and/or after).*	
If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type.	NA
Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, please describe these issues further and refer to specific datasets or data types when appropriate.	☐ Yes, human subject data ☒ Yes, animal data ☐ Yes, dual use ☐ No If yes, please describe: Animal experiments – ECD approved project n. 009/2023.

[5] These data are generated by combining multiple existing datasets.

Will you process personal data[6]? If so, briefly describe the kind of personal data you will use. Please refer to specific datasets or data types when appropriate. If available, add the reference to your file in your host institution's privacy register.	☐ Yes ☒ No If yes: - Short description of the kind of personal data that will be used: - Privacy Registry Reference:
Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.	☐ Yes ☒ No If yes, please comment:
Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements, research collaboration agreements)? If so, please explain to what data they relate and what restrictions are in place.	☐ Yes ☒ No If yes, please explain:
Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain to what data they relate and which restrictions will be asserted.	☐ Yes ☒ No If yes, please explain:

[6] See Glossary Flemish Standard Data Management Plan

3. Documentation and Metadata	
Clearly describe what approach will be followed to capture the accompanying information necessary to keep **data understandable and usable**, for yourself and others, now and in the future (e.g. in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).	Standardized protocols (SOPs, including all relevant details of experimental setup and procedures) will be applied and enforced by the head lab manager/ lab technician. For video files (mouse behavior) the following information will be noted: mouse ID(s), date, time, protocol (i.e. SOP), and experimenter. The methodology and protocol will be described in detail in the lab book. For audio files (mouse ultrasonic vocalizations) the following information will be noted: mouse ID(s), date, time, protocol (i.e. SOP), and experimenter. The methodology and protocol will be described in detail in the lab book. For automated behavioral measurements (e.g. locomotor activity, social contacts) the following information will be noted: mouse ID(s), date, time, protocol (i.e. SOP), and experimenter. The methodology and protocol will be described in detail in the lab book. For other datatypes the following information will be noted: mouse ID(s), date, time, protocol (i.e. SOP), and experimenter. The methodology and protocol will be described in detail in the lab book.
Will a metadata standard be used to make it easier to **find and reuse the data**? If so, please specify which metadata standard will be used. If not, please specify which metadata will be created to make the data easier to find and reuse. *Repositories could ask to deliver metadata in a certain format, with specified ontologies and vocabularies, i.e. standard lists with unique identifiers.*	☐ Yes ☒ No If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used: If no, please specify (where appropriate per dataset or data type) which metadata will be created: Metadata include a unique mouse ID, together with its birth date, genotype, sex, and experimental condition. All other data are linked to the individual animal through the unique mouse ID only (but not birth date, genotype, sex, and experimental condition) to avoid a bias during data acquisition and analysis. After completing relevant parts of the data acquisition process, data will be merged in SPSS linking all relevant data through the unique mouse ID. Metadata will be also stored in SharePoint Online.

4. Data Storage & Back-up during the Research Project

Where will the data be stored?	There are several provisions in place in order to preserve the data during and after the end of the research. Data will be stored on hard drives during experiments. After the experiments, data will be transferred to two external hard drives (two copies; one working copy and one backup copy) and metadata will additionally be transferred to the large, safe, and automatically backed up central network device of KU Leuven, OneDrive. In addition, relevant files will be also saved and stored on SharePoint Online. Moreover, it is planned to make exemplary audio files, i.e. recordings of ultrasonic vocalizations, available to the scientific community through the online platform mouseTube (or similar platforms). Furthermore, it is planned to upload metadata of key confirmatory studies to a general repository (e.g. Open Science Framework). For optimal storage of biomaterials, fridges, -20°C freezers, and -80°C freezers will be used.
How will the data be backed up? *What storage and backup procedures will be in place to prevent data loss? Describe the locations, storage media and procedures that will be used for storing and backing up digital and non-digital data during research.*[7] *Refer to institution-specific policies regarding backup procedures when appropriate.*	For backup, the data will be stored on external hard drives (two copies; one working copy and one backup copy) and metadata will be additionally transferred to the university's central servers with automatic daily backup procedures, for at least 10 years, conform the KU Leuven RDM policy. External hard drives (with a capacity of several TB) will be ordered before the start of data acquisition.
Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.	☐ Yes ☒ No If yes, please specify concisely: If no, please specify: External hard drives (with a capacity of several TB) will be ordered before the start of data acquisition.

[7] Source: Ghent University Generic DMP Evaluation Rubric: https://osf.io/2z5g3/

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons? *Clearly describe the measures (in terms of physical security, network security, and security of computer systems and files) that will be taken to ensure that stored and transferred data are safe.* [7]	Data acquisition is performed in an animal laboratory with limited access. External hard drives and computers are bitlocker/ password protected and PhD and postdoctoral researchers will keep the external hard drives in a cabinet that can be locked. The data stored on SharePoint Online will be accessible only by PhD and postdoctoral researchers and PI.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	Circa 2000 € for external hard drives, acquired during the project period and covered through FWO funding.

5. Data Preservation after the end of the Research Project	
Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).	Data storage during and after the end of the research project does not differ. There are several provisions in place in order to preserve the data during and after the end of the research. According to the Research Data Management policy at KU Leuven, all relevant research data will be kept for at least 10 years after the end of the research. Data will be stored on hard drives during experiments. After the experiments, data will be transferred to two external hard drives (two copies; one working copy and one backup copy) and metadata will additionally be transferred to the large, safe, and automatically backed up central network device of KU Leuven, OneDrive. In addition, relevant files will be also saved and stored on SharePoint Online. Moreover, it is planned to make exemplary audio files, i.e. recordings of ultrasonic vocalizations, available to the scientific community through the online platform mouseTube (or similar platforms). Furthermore, it is planned to upload metadata of key confirmatory studies to a general repository (e.g. Open Science Framework).
Where will these data be archived (stored and curated for the long-term)?	For long-term storage, the data will be stored on external hard drives (two copies; one working copy and one backup copy) and metadata will be additionally transferred to SharePoint Online and to the university's central servers with automatic daily backup procedures, for at least 10 years, conform the KU Leuven RDM policy.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	Circa 2000 € for external hard drives, acquired during the project period and covered through FWO funding.

6. Data Sharing and Reuse	
Will the data (or part of the data) be made available for reuse after/during the project? Please explain per dataset or data type which data will be made available. *Note that 'available' does not necessarily mean that the data set becomes openly available, conditions for access and use may apply. Availability in this question thus entails both open & restricted access. For more information: https://wiki.surfnet.nl/display/standards/info-eu-repo/#infoeurepo-AccessRights*	☒ Yes, in an Open Access repository ☒ Yes, in a restricted access repository (after approval, institutional access only, …) ☐ No (closed access) ☒ Other, please specify: It is planned to make exemplary audio files, i.e. recordings of ultrasonic vocalizations, available to the scientific community through the online platform mouseTube (or similar platforms). Furthermore, it is planned to upload metadata of key confirmatory studies to a general repository (e.g. Open Science Framework).
If access is restricted, please specify who will be able to access the data and under what conditions.	The exemplary audio files, i.e. recordings of ultrasonic vocalizations, will be available to the scientific community through the online platform mouseTube (or similar platforms; with restricted access, i.e. the user needs to register). Metadata of key confirmatory studies will be available through a general repository (e.g. Open Science Framework).
Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain per dataset or data type where appropriate.	☐ Yes, privacy aspects ☐ Yes, intellectual property rights ☐ Yes, ethical aspects ☐ Yes, aspects of dual use ☐ Yes, other ☒ No If yes, please specify:
Where will the data be made available? If already known, please provide a repository per dataset or data type.	The exemplary audio files, i.e. recordings of ultrasonic vocalizations, will be available to the scientific community through the online platform mouseTube (or similar platforms). Metadata of key confirmatory studies will be available through a general repository (e.g. Open Science Framework).

When will the data be made available? *This could be a specific date (dd/mm/yyyy) or an indication such as 'upon publication of research results'.*	Upon publication of the research results.
Which data usage licenses are you going to provide? If none, please explain why. *A data usage license indicates whether the data can be reused or not and under what conditions. If no licence is granted, the data are in a grey zone and cannot be legally reused. Do note that you may only release data under a licence chosen by yourself if it does not already fall under another licence that might prohibit that.* *Example Answer: E.g. "Data from the project that can be shared will be made available under a Creative Commons Attribution license (CC-BY 4.0), so that users have to give credit to the original data creators."* [8]	Still needs to be determined. The example listed here sounds like an interesting possibility: "DATA FROM THE PROJECT THAT CAN BE SHARED WILL BE MADE AVAILABLE UNDER A CREATIVE COMMONS ATTRIBUTION LICENSE (CC-BY 4.0), SO THAT USERS HAVE TO GIVE CREDIT TO THE ORIGINAL DATA CREATORS."
Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here. *Indicate whether you intend to add a persistent and unique identifier in order to identify and retrieve the data.*	☐ Yes ☒ No If yes:
What are the expected costs for data sharing? How will these costs be covered?	Still needs to be determined.

[8] Source: Ghent University Generic DMP Evaluation Rubric: https://osf.io/2z5g3/

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	PhD and postdoctoral researchers are responsible for day-to-day data management, including data documentation and metadata during the research project. Data management is enforced by the head lab manager/ lab technician.
Who will manage data storage and backup during the research project?	PhD and postdoctoral researchers are responsible for day-to-day data management, including data storage and backup during the research project. Data management is enforced by the head lab manager/ lab technician.
Who will manage data preservation and sharing?	PhD and postdoctoral researchers are responsible for day-to-day data management, including data preservation and sharing during the research project. Data management is enforced by the head lab manager/ lab technician.
Who will update and implement this DMP?	PhD and postdoctoral researchers are responsible for day-to-day data management, including update and implementation during the research project. Data management is enforced by the head lab manager/ lab technician. The PI bears the end responsibility of updating & implementing this DMP.