FWO DMP Template - Flemish Standard Data Management Plan

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	**Dimitrios Konstantinidis 0000-0002-2134-6823**
Contributor name(s) (+ ORCID) & roles	**Kevin Verstrepen 0000-0002-3077-6219, supervisor**
Project number[1] & title	ProteoYeast 1283423N
Funder(s) GrantID[2]	ProteoYeast 1283423N
Affiliation(s)	x KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: Provide ROR[3] identifier when possible:

[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/

Please provide a short project description	Protein synthesis is a complex, crucial cellular process, involving many quality control, sorting and trafficking steps. This complexity hampers a full mechanistic understanding and drastically impedes pathway improvements, required for several biotech and medical applications. Prior studies on protein synthesis yielded only limited results, mainly because they focus on single gene perturbation and secretion maximization—which are only some steps of the sophisticated process. Most studies solely use 1 genetic background, thus missing how different (natural) genetic variations could affect protein levels. The yeast Saccharomyces cerevisiae combines many unique traits, making it the perfect model to study protein synthesis. It is also commonly used for heterologous protein expression, facilitating direct valorization of results. Here, I will identify genetic factors underlying efficient protein production and accumulation by exploring S. cerevisiae natural biodiversity. I will use an innovative and multidisciplinary approach, combining my omics skills with Round-Robin QTL mapping and state-of-the-art high-throughput proteomics. Ultimately, this will lead to (i) genome-wide mapping of the S. cerevisiae protein accumulation biodiversity, (ii) mechanistic understanding of genetic factors underpinning this phenotype, and (iii) development of new protein over-accumulating strains. Today's demand for a sustainable bio- economy make this project timely and fitting to societal needs.

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
High-throughput screen	A collection of *Saccharomyces cerevisiae* strains (500 in total) from diverse genetic backgrounds able to secrete selected proteins.	☒ 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: ☐ NA	☐ < 100 MB ☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	In total 10 96-well plates.
Next Generation Sequencing	Whole Genome sequences of selected strains (parental strains) originating from the above mentioned screen, as well	☒ Generate new data	☒ Digital ☒ Physical	☒ Experimental	☒ .csv ☒ .txt ☒ other: .bam, .vcf	☒ < 100 GB	In total 60 96-well plates of single isolated *S. cerevisiae* spores.

	as sequencing of pooled F1 yeast spores						

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.	

[4] Add rows for each dataset you want to describe.
[5] These data are generated by combining multiple existing datasets.

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:
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:

[6] See Glossary Flemish Standard Data Management Plan

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: Yes We do not exclude that the proposed work could result in research data with potential for tech transfer and valorization. Ownership of the data generated belongs to KU Leuven and VIB in accordance with the framework agreement of both institutes. VIB has a policy to actively monitor research data for such potential. If there is substantial potential, the invention will be thoroughly assessed, and in a number of cases the invention will be IP protected (mostly patent protection or copyright protection). As such the IP protection does not withhold the research data from being made public. In the case a decision is taken to file a patent application it will be planned so that publications need not be delayed. The use of strains/constructs will be subjected to the terms described in their respective MTAs. Moreover, The host lab identifies in an early phase the valorization potential of research lines and has a vast network of industrial contacts to efficiently start the route to commercialization. The lab is supported in this matter by Dr. Stijn Spaepen, IOF innovation manager responsible for research valorization. For research with valorization potential, the host lab actively protects its IP by filling patent applications. Type of data with potential for tech transfer and valorization: yeast strains isolated and generated during the timeframe of this project, sequencing information generated during the timeframe of this project and (analysis) data and models hereof derived. Valorization potential includes strain licensing, linking a specific sequence variant to a phenotype.
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:

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).	Data will be generated following standardized protocols. Metadata will be documented by the research and technical staff at the time of data collection and analysis, by taking careful notes in the electronic laboratory notebook (E-notebook) and/or in hard copy lab notebooks that refer to specific datasets. Cryotubes of biological samples (bacterial and yeast strains) stored at -80°C will be labelled with a reference number that links to an entry in or strain database. All datasets will be accompanied by a README.txt file containing all the associated metadata (see more details below). The data will be generated following standardized protocols. Clear and detailed descriptions of these protocols will be stored in our lab protocol database, and published along with the results.

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: While specific data types might require particular metadata, as a general rule the metadata will be based on a generalized metadata schema such as Dublin Core or DataCite. We will closely monitor MIBBI (Minimum Information for Biological and Biomedical Investigations) for metadata standards that are more specific to our data. Metadata will include the following elements: • Title: free text • Creator: Last name, first name, organization • Date and time reference • Subject: Choice of keywords and classifications • Description: Text explaining the content of the data set and other contextual information needed for the correct interpretation of the data, the software(s) (including version number) used to produce and to read the data, the purpose of the experiment, etc. • Format: Details of the file format, • Resource Type: data set, image, audio, etc. • Identifier: DOI (when applicable) • Access rights: closed access, embargoed access, restricted access, open access. Additionally, we will closely monitor MIBBI (Minimum Information for Biological and Biomedical Investigations) for metadata standards more specific to our data type. For specific datasets, additional metadata will be associated with the data file as appropriate. Give details as needed for the project. Specific examples: - SOPs for biological data generation are kept on a dedicated KU Leuven shared drive. A central excel file is stored on that same drive, detailing for examples: (1) sample ID; (2) SOP with which data generation was performed; (3) abnormalities or deviations from SOP in data generation; (4) experimental QC values (e.g. DNA concentrations); (5) location of the source sample in the freezer. - For bioinformatics processing, a data analysis log will be kept that details: (1) sequencing run ID; (2) the bioinformatics SOPs/scripts that were applied; (3) location of source files; (4) abnormalities or deviations. The final dataset will be accompanied by this information under the form of a README.txt document. This file will be located in the top level directory of the dataset and will also list the contents of the other files

	and outline the file-naming convention used. This will allow the data to be understood by other members of the laboratory and add contextual value to the dataset for future reuse. If no, please specify (where appropriate per dataset or data type) which metadata will be created:

4. Data Storage & Back-up during the Research Project	
Where will the data be stored?	secure server (digital vault). - Omics data: omics data generated during the project will either be stored on KU Leuven servers or on The Flemish Supercomputer Centre (VSC), initially in the staging area and later in the archive area. - Vectors: As a general rule at least two independently obtained clones will be preserved for each vector, both under the form of purified DNA (in -20°C freezer) and as a bacteria glycerol stock (-80°C). All published vectors and the associated sequences will be sent to the non-profit plasmid repository Addgene, which will take care of vector storage and shipping upon request. - Bacterial and yeast strains will be stored in a -80°C freezer in the lab of Kevin Verstrepen. Costs are covered by general lab expenses. - Nucleic acid and protein sequences: All nucleic acid and protein sequences generated during the project will be stored on KU Leuven servers. Upon publication, all sequences supporting a manuscript will be made publicly available via repositories such as the GenBank database or the European Nucleotide Archive (nucleotide sequences from primers / new genes / new genomes), NCBI Gene Expression Omnibus (microarray data / RNA-seq data / CHIPseq data), the Protein Database (for protein sequences), the EBI European Genome-phenome Archive (EGA) for personally identifiable (epi)genome and transcriptome sequences.

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.*	KU Leuven drives are backed-up according to the following scheme: - data stored on the "L-drive" is backed up daily using snapshot technology, where all incremental changes in respect of the previous version are kept online; the last 14 backups are kept. - data stored on the "J-drive" is backed up hourly, daily (every day at midnight) and weekly (at midnight between Saturday and Sunday); in each case the last 6 backups are kept. - data stored on the digital vault is backed up using snapshot technology, where all incremental changes in respect of the previous version are kept online. As standard, 10% of the requested storage is reserved for backups using the following backup regime: an hourly backup (at 8 a.m., 12 p.m., 4 p.m. and 8 p.m.), the last 6 of which are kept; a daily backup (every day) at midnight, the last 6 of which are kept; and a weekly backup (every week) at midnight between Saturday and Sunday, the last 2 of which are kept. - All omics data stored on the Flemish Supercomputer Centre (VSC) will be transferred on a weekly basis to the archive area which is backed up. Incremental backups are done daily from one 20 TB QNAP NAS to a second 20 TB QNAP NAS. - For bacterial and yeast strains: A backup of selected strains will be stored in -80°C freezers in a physically different location (CMPG main building in the Arenbergpark) than the initial stock in a more compressed form (96-well plates instead of cryotubes).
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: There is sufficient storage and back-up capacity on all KU Leuven servers: - the "L-drive" is an easily scalable system, built from General Parallel File System (GPFS) cluster with NetApp eseries storage systems, and a CTDB samba cluster in the front-end. - the "J-drive" is based on a cluster of NetApp FAS8040 controlers with an Ontap 9.1P9 operating system. If no, please specify:

[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]	Both the "L-drive" and "J-drive" servers are accessible only by laboratory members, and are mirrored in the second ICTS datacenter for business continuity and disaster recovery so that a copy of the data can be recovered within an hour. Access to the digital vault is possible only through using a KU Leuven user-id and password, and user rights only grant access to the data in their own vault. Sensitive data transfer will be performed according to the best practices for "Copying data to the secure environment" defined by KU Leuven. The operating system of the vault is maintained on a monthly basis, including the application of upgrades and security patches. The server in the vault is managed by ICTS, and only ICTS personnel (bound by the ICT code of conduct for staff) have administrator/root rights. A security service monitors the technical installations continuously, even outside working hours. All private data will be rendered anonymous before processing outside the digital vault. Only the PI will be granted access to the server to deposit private data. The PI will be the only responsible for linking patient information, survey data and/or tissue samples, and will strictly respect confidentiality. All de-identified data will be exported from the database by the PI, and stored on KU Leuven servers from where it can be accessed by the research and technical staff from the laboratory. For bacterial and yeast strains: unauthorized people do not have access to the strain collections.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?	The total estimated cost of data storage during the project is 5000EUR which can be covered by the host lab. This estimation is based on the following costs: -The costs of digital data storage are as follows: 173,78€/TB/Year for the “L-drive” and 519€/TB/Year for the “J-drive”. Electricity costs for the -80° freezers present in the labs are included in general lab costs. Data storage and backup costs are included in general lab costs. Yeast/bacteria strains are easily kept alive for several weeks. This costs on average <5 euro for sets of hundreds of strains. When no experiments are planned with a specific strain, and in compliance with the 3R's rule (https://www.nc3rs.org.uk), cryopreservation will thus be used to safeguard the line, prevent genetic drift, loss of transgene and potential infections or breeding problems. -80°C freezers are present in the lab of Kevin Verstrepen and costs are included in general lab costs.

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...).	The minimum preservation term of 5 years after the end of the project will be applied to all datasets. All datasets will be stored on the university's central servers with automatic back-up procedures for at least 5 years, conform the KU Leuven RDM policy. The costs (€156 per TB per year for “Large volume-storage”) will be covered by the host lab.

Where will these data be archived (stored and curated for the long-term)?	As a general rule, datasets will be made openly accessible, whenever possible via existing platforms that support FAIR data sharing (www.fairsharing.org), at the latest at the time of publication. For all other datasets, long term storage will be ensured as follows: -Digital datasets: files will be stored on the "L-drive". -Omics data: datasets will be stored on the "L-drive" or, for larger datasets, on the Vlaams Supercomputer Centrum. -Vectors: As a general rule at least two independently obtained clones will be preserved for each vector, both under the form of purified DNA (in -20°C freezer) and as a bacteria glycerol stock (-80°C). -Genetically modified organisms: Drosophila lines will be housed locally. All other lines that are no actively used for experiments will be cryopreserved. - Following publication, the results associated with each study will also be deposited in the Dryad repository, where they will be preserved indefinitely. What are the expected costs for data preservation during these 5 years? How will the costs be covered? *Although FWO has no earmarked budget at its disposal to support correct research data management, FWO allows for part of* ***the allocated project budget*** *to be used to cover the cost incurred.* The total estimated cost of data storage during 5 years
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	The total estimated cost of data storage during 5 years after the end of the project is <3000EUR. This estimation is based on the following costs: -The costs of digital data storage are as follows: 173,78€/TB/Year for the "L-drive" and 519€/TB/Year for the "J-drive". Electricity costs for the -80° freezers present in the labs are included in general lab costs. Data storage and backup costs are included in general lab costs.

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: Participants to the present project are committed to publish research results to communicate them to peers and to a wide audience. All research outputs supporting publications will be made openly accessible. Depending on their nature, some data may be made available prior to publication, either on an individual basis to interested researchers and/or potential new collaborators, or publicly via repositories (e.g. negative data). We aim at communicating our results in top journals that require full disclosure upon publication of all included data, either in the main text, in supplementary material or in a data repository if requested by the journal and following deposit advice given by the journal. Depending on the journal, accessibility restrictions may apply. Biological material will be distributed to other parties if requested (unless we first want to file for IP)
If access is restricted, please specify who will be able to access the data and under what conditions.	

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: Yeast strains will be shared upon simple request following publication, unless we identify valuable IP. In this case, we will first protect commercial exploitation, either through patenting or via an MTA that restricts the material from commercial use. We aim at communicating our results in top journals that require full disclosure of all included data. Biological material will be shared upon simple request following publication, unless we identify valuable IP, in which case we will first protect commercial exploitation, either through patenting or via an MTA that restricts the material from commercial use
Where will the data be made available? If already known, please provide a repository per dataset or data type.	In an Open Access repository, Upon request by mail
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, After an embargo period.

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]	Whenever possible, datasets and the appropriate metadata will be made publicly available through repositories that support FAIR data sharing. As detailed above, metadata will contain sufficient information to support data interpretation and reuse, and will be conform to community norms. These repositories clearly describe their conditions of use (typically under a Creative Commons CC0 1.0 Universal (CC0 1.0) Public Domain Dedication, a Creative Commons Attribution (CC-BY) or an ODC Public Domain Dedication and Licence, with a material transfer agreement when applicable). Interested parties will thereby be allowed to access data directly, and they will give credit to the authors for the data used by citing the corresponding DOI. For data shared directly by the PI, a material transfer agreement (and a nondisclosure agreement if applicable) will be concluded with the beneficiaries in order to clearly describe the types of reuse that are permitted.
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: Accession or Doi numbers are not yet available.
What are the expected costs for data sharing? How will these costs be covered?	It is the intention to minimize data management costs by implementing standard procedures e.g. for metadata collection and file storage and organization from the start of the project, and by using free-touse data repositories and dissemination facilities whenever possible. Data management costs will be covered by the laboratory budget.

7. Responsibilities

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

Who will manage data documentation and metadata during the research project?	Metadata will be documented by the research and technical staff at the time of data collection and analysis, by taking careful notes in the physical or electronic laboratory notebook (E-notebook) that refer to specific datasets.
Who will manage data storage and backup during the research project?	The research and technical staff will ensure data storage and back up, with support from René Custers and Alexander Botzki for the electronic laboratory notebook (ELN) and from Raf De Coster for the KU Leuven drives.
Who will manage data preservation and sharing?	The PI (Kevin Verstrepen) is responsible for data preservation and sharing, with support from the research and technical staff involved in the project, from René Custers and Alexander Botzki for the electronic laboratory notebook (ELN) and from Raf De Coster for the KU Leuven drives.
Who will update and implement this DMP?	The PI (Kevin Verstrepen) is ultimately responsible for all data management during and after data collection, including implementing and updating the DMP.