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	**Feili Lai, ORCID: 0000-0002-4945-0737**
Contributor name(s) (+ ORCID) & roles	**Johan Hofkens, ORCID: 0000-0002-9101-0567, role: supervisor**
Project number[1] & title	1298323N, Autonomous light tracking system containing well-developed photocatalysts for enhanced nitrate reduction to ammonia generation
Funder(s) GrantID[2]	D-2023-1846
Affiliation(s)	☑ 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	Nitrate reduction to ammonia (NRA) is a very recent yet effective strategy to convert NO_3^- into NH_3 at ambient conditions. It is a promising alternative for the Haber-Bosch process and as such an important step to realizing a fully carbon-neutral Europe. By employing density functional theory calculations, the first aim of this project is to screen different NRA catalysts (e.g., transition metal compounds), investigate the relationship between NRA properties and various catalyst modification strategies, and unveil NRA mechanisms on different activated surfaces. Based on these insights, highly efficient NRA photocatalysts will be developed. Due to the poor use of light energy in a traditional photocatalytic system, we will combine the new photocatalysts with an autonomous light tracking (ALT) system. Our ATL is based on a highly oriented hydrogel structure. To fabricate rapidly responsive ALT photocatalytic systems, which can fully use the light energy and maintain maximal photocatalytic NRA efficiency by keeping optimal alignment between its top surface and incident light, we will employ a combination of 3D printing, bi-directional freeze-drying, and UV polymerization. Furthermore, for exploring the potential of industrial application of the ALT photocatalytic system, scalable "artificial water plants" will be assembled with self-healing and transplantable properties that enable them to meet extreme environments and different photocatalytic reactions.

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
PP	Photocatalytic performance	☒ 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	
DFT	Theoretical models	☒ 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	☐ < 100 MB ☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	

					☐ .gml ☒ other: ☐ NA		

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.	No
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:

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

[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).	For the theoretical calculations, I will provide the raw data of the models that can be checked by the software of VASP. For the characterization, I will provide the raw data from the machines, such as XPS, XRD, TEM, SEM, and so on. For the photocatalytic performance, I will provide the txt data.
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: The data will be organized in folders with meaningful names, one per simulation run. Each folder will contain a README.txt explaining how to run the specific case and explaining what the related input/output files are. In addition, the header of the case file (.CFcase) will summarise the conditions and purpose of the testcase. All the output files will be put inside an internal folder RESULTS, also providing its own README.txt with additional information concerning the content of those files and visualization tips if any (for instance guidance on how to use additional scripts/macros for etracting some relevant quantities). 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?	Our larger volume data will be temporarily stored on the VSC cluster, then periodically selected and copied to our research unit central storage facility and, partly, to Box when needing to share data more efficiently within the team. Other data (e.g. source code, input simulation files) will be directly stored on the COOLFluiD GitHub repository.
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.*	The data will be stored on the university's central servers with automatic daily back-up procedures. In addition, all data uploaded to the COOLFluiD github repository will be automatically backed up under versioning control.
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: The storage and backup capacity is sufficient. The larger data in our case will be the mesh files and especially solution files in TECPLOT and/or CFmesh and/or Paraview formats.Those will be initially stored in the VSC cluster and then selected and backed up in our research unit central storage facility and/or Box. 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]	We won't be working with sensitive data so this issue is really not a concern for us. At the end of the day, we will be working on a open source code and try to release on the public repository as much as we can with no particular worries about who can use or access our data.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	We will make sure to limit the expenses for data storage to 2000 euros maximum for the whole duration of the project, which can be directly taken from the operational budget we allocated in the FWO proposal.

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...).	1) all new source code will be preserved and automatically backed up; 2) a selection of representative CFcase files and meshes will be stored; 3) a selection of input magnetograms will be stored; 4) a selection of the numerical output (~30% of the total, due to high storage requirements) will be stored; 5) all scripts will be stored; 6) a selection of images and videos will be stored. No particular contractual or legal restrictions shall apply on the usage of 1) 2) 3) 5). However, data of type 4) and 6) resulting from this project may be reused by third party only after negotiating, case by case, the conditions for their exploitation with the PI and co-PI of the FWO project.
Where will these data be archived (stored and curated for the long-term)?	A small part of our data will be archived directly on the COOLFluiD Github repository (source code, scripts, input files, some images and videos) while the larger part (big meshes, solution files, all images and videos) will be archived on the Large Volume Storage (LVS) system, after having deleted data corresponding to unsuccessful or useless simulation runs.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	The expected cost for storing data on the LVS system is 156,60/TB/Year. This expense will be covered by the allocated project budget.

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:
If access is restricted, please specify who will be able to access the data and under what conditions.	The source code, scripts, input files for representative numerical simulation, some images and videos of the results will be directly available through the public access COOLFluiD github repository, while the larger mesh/solution files, images and videos will be stored on LVS system and made temporarily available via Box upon explicit request via e-mail.
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.	No.

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'.*	The source code, scripts, input files for representative numerical simulations, some images and videos of the results will be directly available through the public access COOLFluiD github repository immediately after the end of the project (and partially even during the project), while the larger mesh/solution files, images and videos will be made available via Box upon explicit request via e-mail after the publication of the corresponding 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]	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?	We expect a maximum allocation of 1000 euros for data sharing to be charged directly from the allocated project budget. Considering that Box only costs 10 euro/year for 100 GB, this leaves room to more expensive solutions in case of need.

[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?	Johan Hofkens (co-PI)
Who will manage data storage and backup during the research project?	Johan Hofkens (co-PI)
Who will manage data preservation and sharing?	Johan Hofkens (co-PI)
Who will update and implement this DMP?	The PI and co-PI bear the end responsibility of updating & implementing this DMP.