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	Name : Malek Ben Khalifa & ORCID :0000-0001-6944-5663
Contributor name(s) (+ ORCID) & roles	Name : Jérôme Loreau & ORCID : 0000-0002-6142-1509 Role : project supervisor
Project number[1] & title	Project number : 12E6623N Title : Ro-vibrational quenching of molecules of astrophysical, atmospheric and planetary interest.
Funder(s) GrantID[2]	FWO
Affiliation(s)	☐ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: Provide ROR[3] identifier when possible:
Please provide a short project description	For astronomers to understand their observations, chemists must answer a fundamental question: What is the excitation and (de)-excitation scheme of the molecule in interstellar regions? Is it excited via collisions with molecules, or with photons? In this project, we study the collisional processes that take place in astronomical environments, in fact, the inelastic collisions between molecules play an important role in the populating of the rovibrational quantum levels which are observed. In fact, the probability for a molecule to transit from one state to another via molecular collisions at a given temperature is related to the collisional rates. The range of physical conditions reigning in astrophysics requires the knowledge of collisional rates between 5 and a few thousand Kelvin, which is the objective of our project.

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.
2Funder(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.
3Research Organization Registry Community. https://ror.org/

2. Research Data

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
Potential energy surfaces (PES)	Compute the N-dimensional PESs presenting the interaction between He and/or H_2 and the molecule under study subjected to bending and stretching vibrational motion using Molpro package	☐ 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	
Cross sections	compute the ro-vibrational cross sections based on analytical development of the PESs computed before.	☐ Generate new data	☐ Digital	☐ Simulation data	☐ .tab	☐ < 100 GB	
Ro-vibrational Rates	Calculate rate coefficient using cross sections computed before	☐ Generate new data	☐ Digital	☐ Simulation data	☐ .tab	☐ < 100 MB	
Codes for the dynamic	write codes for the study of ro-vibrational excitation	☐ Generate new data	☐ Physical	☐ Software	☐ .txt	☐ < 100 MB	

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.	We don't use existing data
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:

4Add rows for each dataset you want to describe.
5These 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:

6See 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 potential energy surfaces, cross sections and rate coefficients, we will provide numbers, tabulated and easy to understand with comments in the files which describes each column. For the codes that will be written during the research, we will add documentations in order for the code to be easy to read and understand by others We will also add a 'readme' text file to provide details and explanations of the data. The data will be recorded in a data base archive of Professor Jerome Loreau's group with links to each item.
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: We will add a 'readme' text file in order to provide details and explanations of the data.

4. Data Storage & Back-up during the Research Project	
Where will the data be stored?	During the research activities, data will be stored in the external hard drives and the cloud system. Outputs and results will be kept in an archive space of Professor Jerome Loreau's group at the VSC (Flemish Supercomputer Center). This archive will be available for any member of the group, and can be made also available to other researchers interested in the results. The most important data will be included as supplementary information in research papers. Some data will also be included in online databases used by the community (EMAA, LAMDA, BASECOL).
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.*	During the research, I will collect all the important documents for writing the manuscripts in the external hard drives and the cloud system. After the research, all essential data will be kept in an archive space at the VSC (Flemish Supercomputer Center) as well as in the external hard drive at least five years after the end of the research and/or until they have been published in peer-reviewed international journals.
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: Currently, I save my data in the external hard drives and the cloud system If no, please specify:

7Source: 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]	Only me and professor Jérôme Loreau will have access to the data during the research. I will share data with professor Loreau via cloud system.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	The storage and back up costs 240 euros/year and these costs will be covered by me (FWO bench fee)

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 data that will be retained after the end of my project are quantum chemistry data (electronic energies for various molecules of interest in the project), as well as ro-vibrational rate coefficients of CO2, NH3 and CH4 by collision with He and/or H2 molecule. These data will always be available online for use by the astrophysical community in several database such as Basecol, Lamda and EMAA. These databases are devoted to collisional ro-vibrational excitation of molecules by colliders such as atoms, ions, molecules or electrons.
Where will these data be archived (stored and curated for the long-term)?	all data (potential energy surfaces and cross sections) will be kept in an archive space at the VSC. Ro-vibrational rate coefficients will be made online in Lamda, Basecol and EMAA websites.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	On the Lamda, Basecol and EMAA websites, data preservation is free. I will also save the data in cloud, for a costof 240 euros/year. These costs will be covered by me for 3 years at least.

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: Codes developed during this project will be reused by me as well as by our group at KU Leuven for the study of the more detected molecules after the project. Ro-vibrational rate coefficients will be made available online for the scientific community interested by these results. Cross sections will be kept in our archive and shared upon request.
If access is restricted, please specify who will be able to access the data and under what conditions.	N.A
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.	all data (potential energy surfaces and cross sections) will be kept in an archive space at the VSC, Ro-vibrational rate coefficients will be available online in Lamda, Basecol and EMAA websites.
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 data will be available to the community once the results based on these data are published in scientific journals, or upon request for collaborators. The final data i.e ro-vibrational rates, will be always available online.
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]*	I will not provide licenses, once our results are published in peer-reviewed international journals. They will be in open access for the scientific community,
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:

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

What are the expected costs for data sharing? How will these costs be covered?	No additional resources are needed.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	Malek Ben Khalifa
Who will manage data storage and backup during the research project?	Malek Ben Khalifa
Who will manage data preservation and sharing?	Malek Ben Khalifa and jérôme Loreau
Who will update and implement this DMP?	Malek Ben Khalifa