FWO DMP Template - Flemish Standard Data Management Plan

Version KU Leuven

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	**Romain Laugier - https://orcid.org/0000-0002-2215-9413**
Contributor name(s) (+ ORCID) & roles	**Denis Defrère - https://orcid.org/0000-0003-3499-2506**
Project number [1] & title	Advanced Infrared Interferometry Techniques for Exoplanets Characterization (AIITEC)
Funder(s) GrantID [2]	1234224N
Affiliation(s)	✅ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: ROR identifier KU Leuven: 05f950310
Please provide a short project description	The study of exoplanets is one of the big drivers of astrophysics and space missions today, as these worlds orbiting distant stars can reveal much about the origin of our own solar system, and the presence of life in the universe. Over 5000 exoplanets have been detected so far by using indirect methods. Only a handful have been detected directly to be studied in spectroscopy because current instruments only allow us to detect the very young, massive and very wide orbit planets. Interferometry offers a solution. By optically combining the light collected by widely separated telescopes, it can improve the angular resolution limit by a factor twenty (for existing interferometric facilities), opening the door to detecting planets in the inner orbits where most of the planet formation is expected to occur. In this AIITEC FWO project, I will on three generations of infrared interferometric instruments to enable their planet observation capabilities. I will provide novel high-precision data-reduction methods for MATISSE, at the Very Large Telescope Interferometer (VLTI). I will deploy the innovative atmospheric dispersion corrector of Asgard/NOTT in construction at KU Leuven for VLTI. Finally I will guide the architecture design decisions for LIFE, an ambitious space interferometry mission to detect life on rocky exoplanets.

---

[1] "Project number" refers to the institutional project number. This question is optional. 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.

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 [3].

				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
		☐ Generate new data ☐ Reuse existing data	☐ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:		☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
MATISSE_target	Raw MATISSE interferometry data on a science target	Reuse existing data	Digital	Interferometric observation data	Flexible Image Transport System (FITS)	< 100GB	
MATISSE_cal	Raw MATISSE interferometry data on calibrator stars	Reuse existing data	Digital	Interferometric observation data	Flexible Image Transport System (FITS)	< 100GB	
MATISSE_internal	Raw MATISSE calibration data	Generate new data	Digital	Low-level detector data	Flexible Image Transport System (FITS)	< 100 GB	

---

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

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.	Matisse raw data run/porgram id: 105.20HD.001 (aka 0105.C-0407(A)), accessible through the ESO archive http://archive.eso.org/eso/eso_archive_main.html
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, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.	☐ Yes, human subject data; provide SMEC or EC approval number: ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ✅ No Additional information:
Will you process personal data[4]? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).	☐ Yes (provide PRET G-number or EC S-number below) ✅ No Additional information:
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:

---

[4] See Glossary Flemish Standard Data Management Plan

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). _RDM guidance on documentation and metadata_.	**A significant part of the project is to create a new data reduction method to produce processed and analyzed data, and the details of these reduction algorithm will be described in a peer-reviewed paper. The output of this reduction will saved in Optical Interferometry FITS (OIFITS) format, the standard optical interferometry reduced data format to make it understandable.** **If we can demonstrate an improvement with our methods, it will be integrated to the standard data processing pipeline of the MATISSE instrument and documented as such.**

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 OIFITS data standard contains all the relevant metadata. Any other metadata required by the JMMC oidb database will be added. The intermediate data products have low scientific value and will not be kept.** 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? *Consult the [interactive KU Leuven storage guide](#) to find the most suitable storage solution for your data.*	✅ Shared network drive (J-drive) ☐ Personal network drive (I-drive) ☐ OneDrive (KU Leuven) ☐ Sharepoint online ☐ Sharepoint on-premis ☐ Large Volume Storage ☐ Digital Vault ☐ Other:
How will the data be backed up? *What storage and backup procedures will be in place to prevent data loss?*	✅ Standard back-up provided by KU Leuven ICTS for my storage solution ☐ Personal back-ups I make (specify) ☐ Other (specify)

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 no, please specify:
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.* *Guidance on security for research data*	Laptop has full disk encryption, network storage is managed by local IT with the security standards from KULeuven.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	The cost is 300€/TB, but the volume for this project is low enough that it will not be billed.

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...). *Guidance on data preservation*	✅ All data will be preserved for 10 years according to KU Leuven RDM policy ☐ All data will be preserved for 25 years according to CTC recommendations for clinical trials with medicinal products for human use and for clinical experiments on humans ☐ Certain data cannot be kept for 10 years (explain)

Where will these data be archived (stored and curated for the long-term)? *Dedicated data repositories are often the best place to preserve your data. Data not suitable for preservation in a repository can be stored using a KU Leuven storage solution, consult the interactive KU Leuven storage guide.*	✅ KU Leuven RDR ☐ Large Volume Storage (longterm for large volumes) ☐ Shared network drive (J-drive) ✅ Other (specifiy): **The reduced data will be stored on the relevant astronomical community archives: Strasbourg Astronomical Data center (CDS) for the astrophysical information, and Jean-Marie Mariotti Center (JMMC OIDB http://oidb.jmmc.fr/index.html) for the reduced observational data.** **The working copy will be kept by KU Leuven for a period of 10 years.**
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	**For the low volume that this project represents, the costs are forfeited as this will piggyback on larger volume projects.**

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, as open data ☐ Yes, as embargoed data (temporary restriction) ☐ Yes, as restricted data (upon approval, or 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.	

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.	☐ KU Leuven RDR ✅ Other data repository (specify) **Following the practice in use with users of this instrument, reduced data will be stored in the relevant databases (Jean-Marie Mariotti Center JMMC OIDB http://oidb.jmmc.fr/index.html) for open access upon the publication of the article using them.** ☐ Other (specify)
When will the data be made available?	✅ Upon publication of research results ☐ Specific date (specify) ☐ Other (specify)

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.* *Check the RDR guidance on licences for data and software sources code or consult the License selector tool to help you choose.*	☑ CC-BY 4.0 (data) ☐ Data Transfer Agreement (restricted data) ☐ MIT licence (code) ☑ GNU GPL-3.0 (code) ☐ Other (specify)
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, a PID will be added upon deposit in a data repository ☐ My dataset already has a PID ☐ No
What are the expected costs for data sharing? How will these costs be covered?	

7. Responsibilities

Who will manage data documentation and metadata during the research project?	**Romain Laugier - For the documentation of new data created for this project.** **Philippe Berio and Alexis Matter (LAGRANGE) - to help in exporting the reduced data to OIFITS format.**
Who will manage data storage and backup during the research project?	**Maarten Dirickx - Along with the KU Leuven IvS local IT department**
Who will manage data preservation and sharing?	**G. Duchêne - Responsible for OIDB at the JMMC**
Who will update and implement this DMP?	**Romain Laugier**