

## 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	<b>Giovanni Lapeta, 0000-0002-3123-4024</b>
Contributor name(s) (+ ORCID) & roles	
Project number <sup>1</sup> & title	3E221369
Funder(s) GrantID <sup>2</sup>	G0B9923N
Affiliation(s)	<input checked="" type="checkbox"/> KU Leuven <input type="checkbox"/> Universiteit Antwerpen <input type="checkbox"/> Universiteit Gent <input type="checkbox"/> Universiteit Hasselt <input type="checkbox"/> Vrije Universiteit Brussel <input type="checkbox"/> Other: ROR identifier KU Leuven: 05f950310
Please provide a short project description	<p>One of the most important open questions in astrophysics is explaining why the solar wind is so hot. The solar surface is at about 6500K but quickly as one moves in the atmosphere of the Sun the temperature reaches millions of degrees and continues to increase more gently at larger distances before slowly starting to decrease. This goes against any intuition for an expanding gas that moves away from its source of heating.</p> <p>Two new missions have been launched to understand the physics behind this puzzle more in depth: Solar Orbiter (led by Europe) and Parker Solar Probe (led by the USA). These missions are going closer to the sun than any man made object before and will give us an unprecedented wealth of data.</p> <p>The objective of our project is to use the mission data, supplemented by our high-resolution simulations, to extricate the holy grail: what is the equation of state of the expanding solar wind? This equation will tell us how the temperature of the solar wind evolves given the other parameters measured. The novelty of our approach is that we derive this equation with a new machine learning tool that extracts physical laws from data. Potentially this approach can bring new discovery in the form of an equation that we can then interpret and use to explain the thermodynamics of the solar wind.</p>

<sup>1</sup> "Project number" refers to the institutional project number. This question is optional. Applicants can only provide one project number.

<sup>2</sup> 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 <sup>3</sup>.

Dataset Name	Description	New or Reused	Digital or Physical	ONLY FOR DIGITAL DATA	ONLY FOR DIGITAL DATA	ONLY FOR DIGITAL DATA	ONLY FOR PHYSICAL DATA
				Digital Data Type	Digital Data Format	Digital Data Volume (MB, GB, TB)	Physical Volume
iPic3D simulation output	Output of code iPic3D	<input checked="" type="checkbox"/> Generate new data <input type="checkbox"/> Reuse existing data	<input checked="" type="checkbox"/> Digital <input type="checkbox"/> Physical	<input type="checkbox"/> Audiovisual <input checked="" type="checkbox"/> Images <input type="checkbox"/> Sound <input checked="" type="checkbox"/> Numerical <input checked="" type="checkbox"/> Textual <input checked="" type="checkbox"/> Model <input checked="" type="checkbox"/> Software <input type="checkbox"/> Other:	HDF5	<input type="checkbox"/> < 1 GB <input type="checkbox"/> < 100 GB <input type="checkbox"/> < 1 TB <input type="checkbox"/> < 5 TB <input checked="" type="checkbox"/> > 5 TB <input type="checkbox"/> NA	50TB
Trained NN	Parameters of trained Neural Networks	<input type="checkbox"/> Generate new data <input type="checkbox"/> Reuse existing data	<input checked="" type="checkbox"/> Digital <input type="checkbox"/> Physical	<input type="checkbox"/> Audiovisual <input type="checkbox"/> Images <input type="checkbox"/> Sound <input checked="" type="checkbox"/> Numerical <input checked="" type="checkbox"/> Textual <input checked="" type="checkbox"/> Model <input checked="" type="checkbox"/> Software <input type="checkbox"/> Other:	Pytorch	<input type="checkbox"/> < 1 GB <input type="checkbox"/> < 100 GB <input checked="" type="checkbox"/> < 1 TB <input type="checkbox"/> < 5 TB <input type="checkbox"/> > 5 TB <input type="checkbox"/> NA	0.5TB
Training datasets	Data Sets used to train NN	<input checked="" type="checkbox"/> Generate new data <input type="checkbox"/> Reuse existing data	<input checked="" type="checkbox"/> Digital <input checked="" type="checkbox"/> Physical	<input type="checkbox"/> Audiovisual <input checked="" type="checkbox"/> Images <input type="checkbox"/> Sound <input checked="" type="checkbox"/> Numerical	FITS	<input type="checkbox"/> < 1 GB <input type="checkbox"/> < 100 GB <input type="checkbox"/> < 1 TB <input checked="" type="checkbox"/> < 5 TB	5TB

<sup>3</sup> Add rows for each dataset you want to describe.

				<input checked="" type="checkbox"/> Textual <input checked="" type="checkbox"/> Model <input checked="" type="checkbox"/> Software <input type="checkbox"/> Other:		<input type="checkbox"/> > 5 TB <input type="checkbox"/> NA	

***GUIDANCE:***  
 The data description forms the basis of your entire DMP, so make sure it is detailed and complete. It includes digital and physical data and encompasses the whole spectrum ranging from raw data to processed and analysed data including analysis scripts and code. Physical data are all materials that need proper management because they are valuable, difficult to replace and/or ethical issues are associated. Materials that are not considered data in an RDM context include your own manuscripts, theses and presentations; documentation is an integral part of your datasets and should be described under documentation/metadata.  
[RDM Guidance on data](#)

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 data reused
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.	<input type="checkbox"/> Yes, human subject data; provide SMEC or EC approval number: <input type="checkbox"/> Yes, animal data; provide ECD reference number: <input type="checkbox"/> Yes, dual use; provide approval number: <input checked="" type="checkbox"/> No Additional information:
Will you process personal data <sup>4</sup> ? 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).	<input type="checkbox"/> Yes (provide PRET G-number or EC S-number below) <input checked="" type="checkbox"/> No Additional information:

<sup>4</sup> See Glossary Flemish Standard Data Management Plan

<p>Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, ...)?</p> <p>If so, please comment per dataset or data type where appropriate.</p>	<p><input type="checkbox"/> Yes</p> <p><input checked="" type="checkbox"/> No</p> <p>If yes, please comment:</p>
<p>Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements, research collaboration agreements)?</p> <p>If so, please explain to what data they relate and what restrictions are in place.</p>	<p><input type="checkbox"/> Yes</p> <p><input checked="" type="checkbox"/> No</p> <p>If yes, please explain:</p>
<p>Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use?</p> <p>If so, please explain to what data they relate and which restrictions will be asserted.</p>	<p><input type="checkbox"/> Yes</p> <p><input checked="" type="checkbox"/> No</p> <p>If yes, please explain:</p>

### 3. Documentation and Metadata

<p>Clearly describe what approach will be followed to capture the accompanying information necessary to keep <b>data understandable and usable</b>, 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).</p> <p><i><a href="#">RDM guidance on documentation and metadata.</a></i></p>	<p><b>We will follow the existing procedure of the NASA Planetary Science Database (<a href="https://pds.nasa.gov">https://pds.nasa.gov</a>)</b></p>
<p>Will a metadata standard be used to make it easier to <b>find and reuse the data</b>?</p> <p>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.</p> <p><i>REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS.</i></p>	<p><input checked="" type="checkbox"/> Yes  <input type="checkbox"/> No</p> <p>If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used:  <b>We will adhere to the standards of the NASA Planetary Science Database (<a href="https://pds.nasa.gov">https://pds.nasa.gov</a>)</b></p> <p>If no, please specify (where appropriate per dataset or data type) which metadata will be created:</p>

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

<p>Where will the data be stored?</p> <p><i>Consult the <a href="#">interactive KU Leuven storage guide</a> to find the most suitable storage solution for your data.</i></p>	<p> <input type="checkbox"/> Shared network drive (J-drive)  <input type="checkbox"/> Personal network drive (I-drive)  <input checked="" type="checkbox"/> OneDrive (KU Leuven)  <input type="checkbox"/> Sharepoint online  <input type="checkbox"/> Sharepoint on-premis  <input checked="" type="checkbox"/> Large Volume Storage  <input type="checkbox"/> Digital Vault  <input checked="" type="checkbox"/> Other: Mango at KULEuven, Tier-1 Data component at VSC         </p>
<p>How will the data be backed up?</p> <p><i>WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS?</i></p>	<p> <input checked="" type="checkbox"/> Standard back-up provided by KU Leuven ICTS for my storage solution  <input type="checkbox"/> Personal back-ups I make (specify)  <input type="checkbox"/> Other (specify)         </p>
<p>Is there currently sufficient storage &amp; 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.</p>	<p> <input checked="" type="checkbox"/> Yes  <input type="checkbox"/> No             If no, please specify:         </p>
<p>How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?</p> <p><i>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.</i></p> <p><a href="#">Guidance on security for research data</a></p>	<p>We will follow the procedures of the VSC Tier-1 Data component</p>



What are the expected costs for data storage and backup during the research project? How will these costs be covered?	No cost expected, via Tier1 data allocation of VSC
---	--

5. Data Preservation after the end of the Research Project	
<p>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...).</p> <p><a href="#">Guidance on data preservation</a></p>	<p><input checked="" type="checkbox"/> All data will be preserved for 10 years according to KU Leuven RDM policy</p> <p><input type="checkbox"/> 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</p> <p><input type="checkbox"/> Certain data cannot be kept for 10 years (explain)</p>
<p>Where will these data be archived (stored and curated for the long-term)?</p> <p><i><a href="#">Dedicated data repositories</a> 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 <a href="#">interactive KU Leuven storage guide</a>.</i></p>	<p><input checked="" type="checkbox"/> KU Leuven RDR</p> <p><input type="checkbox"/> Large Volume Storage (longterm for large volumes)</p> <p><input type="checkbox"/> Shared network drive (J-drive)</p> <p><input type="checkbox"/> Other (specify):</p>
<p>What are the expected costs for data preservation during the expected retention period? How will these costs be covered?</p>	<p><b>No cost expected</b></p>

## 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](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)
- ☐ Other (specify)

When will the data be made available?	<input checked="" type="checkbox"/> Upon publication of research results <input type="checkbox"/> Specific date (specify) <input type="checkbox"/> Other (specify)
Which data usage licenses are you going to provide? If none, please explain why.  <i>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.</i> Check the <a href="#">RDR guidance on licences</a> for data and software sources code or consult the <a href="#">License selector tool</a> to help you choose.	<input type="checkbox"/> CC-BY 4.0 (data) <input type="checkbox"/> Data Transfer Agreement (restricted data) <input checked="" type="checkbox"/> MIT licence (code) <input type="checkbox"/> GNU GPL-3.0 (code) <input type="checkbox"/> Other (specify)
Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here.  <i>INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.</i>	<input checked="" type="checkbox"/> Yes, a PID will be added upon deposit in a data repository <input type="checkbox"/> My dataset already has a PID <input type="checkbox"/> No
What are the expected costs for data sharing? How will these costs be covered?	<b>No cost is expected.</b>

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	<b>Giovanni Lapenta</b>

Who will manage data storage and backup during the research project?	<b>Giovanni Lapenta</b>
Who will manage data preservation and sharing?	<b>Giovanni Lapenta</b>
Who will update and implement this DMP?	<b>Giovanni Lapenta</b>