

## **Innovative study of Solar Energetic Particle events using state-of-the-art coronal and CME flux-rope models**

### **Application DMP**

---

#### **Questionnaire**

**The questions in this section should only be answered if you are currently applying for FWO funding.  
Are you preparing an application for funding?**

- No

## Innovative study of Solar Energetic Particle events using state-of-the-art coronal and CME flux-rope models

### DPIA

---

DPIA

Have you performed a DPIA for the personal data processing activities for this project?

- Not applicable

**Innovative study of Solar Energetic Particle events using state-of-the-art coronal and CME flux-rope models**  
**GDPR**

---

GDPR

Have you registered personal data processing activities for this project?

- Not applicable

# Innovative study of Solar Energetic Particle events using state-of-the-art coronal and CME flux-rope models

## FWO DMP (Flemish Standard DMP)

### 1. 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.

				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
		<i>Please choose from the following options:</i> <ul style="list-style-type: none"> <li>• Generate new data</li> <li>• Reuse existing data</li> </ul>	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> <li>• Digital</li> <li>• Physical</li> </ul>	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> <li>• Observational</li> <li>• Experimental</li> <li>• Compiled/aggregated data</li> <li>• Simulation data</li> <li>• Software</li> <li>• Other</li> <li>• NA</li> </ul>	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> <li>• .por,</li> <li>• .xml,</li> <li>• .tab,</li> <li>• .csv,.pdf,</li> <li>• .txt, .rtf,</li> <li>• .dwg,</li> <li>• .gml, ...</li> <li>• NA</li> </ul>	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> <li>• &lt;100MB</li> <li>• &lt;1GB</li> <li>• &lt;100GB</li> <li>• &lt;1TB</li> <li>• &lt;5TB</li> <li>• &lt;10TB</li> <li>• &lt;50TB</li> <li>• &gt;50TB</li> <li>• NA</li> </ul>	
PARADISE	software for modeling energetic particles	reuse and update the software	Digital	Software	C++ and python scripts	< 100MB	
EUHFORIA	software to model the solar wind	reuse and update the software	Digital	Software	C++ and python scripts	< 100MB	
ICARUS	software to model the solar wind	reuse and update the software	Digital	Software	Fortran scripts	< 100MB	
PARADISE configuration files	configuration files used to perform PARADISE simulations	generate new data	Digital	Simulation data	txt	< 1GB	
EUHFORIA and ICARUS configuration files	configuration files used to perform solar wind simulations with EUHFORIA and ICARUS	generate new data	Digital	Simulation data	txt	< 1GB	

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:

NA

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- No

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.

- No

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 in the comment section to what data they relate and what restrictions are in place.

- No

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 in the comment section to what data they relate and which restrictions will be asserted.

- No

## 2. 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).

### 1. Simulation Configuration Files

All configuration files for the simulations were stored, ensuring that each simulation can be reproduced exactly as it was originally executed. These files are well-documented and include details such as parameter settings, input conditions, and environmental variables.

### 2. Code Versioning

The version of the code used for each simulation was recorded by including the specific commit hash from the Git repositories of the models (stored on Bitbucket and GitHub). This ensures precise traceability of the codebase, allowing future users to retrieve and replicate the exact code version used.

**Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.**

- No

### **3. Data storage & back-up during the research project**

**Where will the data be stored?**

#### **1. Local Storage Facility**

Primary data storage is on the local storage facility of the research group, which is managed with appropriate access controls and security measures.

#### **2. Cloud Backup**

A backup of all data is maintained on OneDrive, ensuring an additional layer of redundancy and easy access for collaborators.

#### **3. Version-Controlled Repositories**

The models and associated configuration files are stored on Bitbucket and GitHub repositories.

These repositories facilitate version control, enabling precise tracking of changes and fostering collaborative development

**How will the data be backed up?**

Regular synchronization with OneDrive (as mentioned earlier) for cloud-based redundancy.

Scheduled backups to the local storage facility of the research group.

Version-controlled repositories on Bitbucket and GitHub ensure additional backups for models and configuration files.

**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

**How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?**

We will not be working with sensitive data so this issue is really not a concern for us. All types of data stored on the research unit central storage facility follow the KU Leuven and in-house security standards: all data are secured and password protected with access restrictions on file-system level. Access to the OneDrive backups is restricted to authorized accounts. The Bitbucket and GitHub repositories are private, with access granted only to approved collaborators. Commit histories will ensure accountability for any changes made.

**What are the expected costs for data storage and backup during the research project? How will these costs be covered?**

All (potential) data storage costs are covered by the research unit's working budget.

### **4. 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 PARADISE and EUHFORIA software

The configuration files of the various simulations performed with those models.

**Where will these data be archived (stored and curated for the long-term)?**

A small part of our data is archived directly on the PARADISE repository on Bitbucket while the larger part (configuration files of all the performed simulations) is archived on our research unit central storage facility.

**What are the expected costs for data preservation during the expected retention period? How will these costs be covered?**

The Bitbucket and Github repositories are free. The simulation configuration files are lightweight text files. All (potential) data storage costs are covered by the research unit's working budget.

## 5. Data sharing and reuse

**Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.**

- Yes, in a restricted access repository (after approval, institutional access only, ...)

**If access is restricted, please specify who will be able to access the data and under what conditions.**

1. Research Team Members:

Members of the research group can have access to the source code of the software. Access is granted through user-specific permissions.

2. Collaborators:

Approved external collaborators will be granted access to specific datasets via the private Bitbucket or GitHub repositories. Access is limited to the scope of their involvement in the project.

**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 in the comment section per dataset or data type where appropriate.**

- No

**Where will the data be made available? If already known, please provide a repository per dataset or data type.**

A description of the full scientific analysis, tools and deliverables will be published through peer-reviewed journals, which are available through the respective publisher websites, the open access arXiv journal-repository website (<https://arxiv.org>), and the KU Leuven Lirias.

The EUHFORIA and PARADISE repositories on Bitbucket and GitHub are private, with access granted only to approved collaborators. Commit histories will ensure accountability for any changes made.

**When will the data be made available?**

After publication of research results

**Which data usage licenses are you going to provide? If none, please explain why.**

NA

**Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.**

- No

**What are the expected costs for data sharing? How will these costs be covered?**

No costs

## **6. Responsibilities**

**Who will manage data documentation and metadata during the research project?**

Stefaan Poedts

**Who will manage data storage and backup during the research project?**

Stefaan Poedts

**Who will manage data preservation and sharing?**

Stefaan Poedts

**Who will update and implement this DMP?**

Stefaan Poedts