Structure Inversions for Massive Stars from Gravity-mode Asteroseismology
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
		Please choose from the following options: • Generate new data • Reuse existing data	*Please choose from the following options:* • Digital • Physical	*Please choose from the following options:* • Observational • Experimental • Compiled/aggregated data • Simulation data • Software • Other • NA	*Please choose from the following options:* • .por, .xml, .tab, .cvs,.pdf, .txt, .rtf, .dwg, .gml, … • NA	*Please choose from the following options:* • <100MB • <1GB • <100GB • <1TB • <5TB • <10TB • <50TB • >50TB • NA	
Forward stellar models	Stellar models generated with MESA (stellar evolution code), including mode properties computed using GYRE.	Generate new data	Digital	Simulated data	HDF5	<1TB	
Observed mode frequencies	Detected mode frequencies in stars based on light curve analysis.	Reuse existing data	Digital	Compiled/Aggregated data	CSV, TXT	<100MB	
Inferred stellar structure	Inferred structures of stars, including structure profiles, rotation profiles and magnetic field strengths	Generate new data	Digital	Compiled/Aggregated data	HDF5, TXT	<1GB	
Inversion codes	Software developed during the project for performing inversions. Generates the "Inferred stellar structure" dataset	Generate new data	Digital	Software	TAR or Git repository	<100MB	
Simulation code	Software used to simulate stars (MESA, GYRE).	Reuse existing data	Digital	Software	TAR or Git repository	<100GB	

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:

Observed mode frequencies: obtained from multiple sources, often tables in peer reviewed papers.
Simulation code: the MESA and GYRE codes are available on GitHub (https://github.com/MESAHub/mesa, https://github.com/rhdtownsend/gyre) and the MESA code is also available on Zenodo (https://zenodo.org/record/7319739)

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).

- For the data that will be reused from previously published works (be it software or other types of dataset), I will provide links (DOI when available) to those works and include information on which parts of those works were used. If transformations of storage formats are used (for example from tables in published works to CSV files), those transformations will be clearly documented in accompanied TXT files.
- For output from reused software, I will refer to documentation of those software tools. Should such documentation not exist, I will include TXT files describing the outputs.
- Outputs from software produced during the project as well as the software itself will be documented via TXT files (for example README.txt) and inline code comments.

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.

- Yes

The specific standard metadata used will be based on the repository where the data is made available.

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

Where will the data be stored?

The Institute of Astronomy has a well-developed in-house storage and back- up system which will allow for continuous data preservation. This will be used during the project both for in progress data as well as finished data products.

How will the data be backed up?

Data stored on the Institute of Astronomy's storage system and data from laptops or desktops of institute members is backed up regularly (daily) on servers inside the Institute of Astronomy and one of the ICTS data centers.

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

The size of the resulting data products is negligible (<1TB) compared to the overall size of the storage cluster in use by the institute of astronomy.

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

The data can only be accessed from inside the Institute of Astronomy's network or via secure Wireguard tunnel. Physical access to the devices storing the data is limited to specific staff members. Data stored on laptops or other mobile devices is encrypted at rest. These datasets require no additional security measures as no personal information or otherwise confidential information is processed.

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

The costs for the data stored on the Institute of Astronomy's in-house storage system will be covered by shared funds of the institute.

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...).

All data products will be preserved.

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

All data will be stored long term on the Institute of Astronomy in house storage system. Additionally, datasets that are made publicly available will be stored indefinitely on archiving services like Zenodo.

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

Given the small total size of the datasets, the storage of all the publicly available data are expected to be free of charge.
The costs for the data stored on the Institute of Astronomy's in-house storage system will be covered by shared funds of the institute.

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 an Open Access repository

All developed software will be made publicly available on GitHub or Zenodo or similar services. Other data will be made available on a case by case basis. For example

- Data that cannot be reproduced from other data sources will be made available. Observations are a type of dataset for which this would apply (although currently such datasets are not planned for this project).
- Data which is valuable to a wide range of researchers, for which not making this data available would be an unnecessary barrier to building further upon my research. This would be the case for the main data product of this project: "Inferred stellar structure"

Other datasets will be made available upon request. For example, this includes the simulated stellar models.

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

Not applicable.

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.

Software will be made available on repositories like GitHub or Zenodo. Other datasets will be archived in repositories like Zenodo or the CDS database (cds.u-strasbg.fr).

When will the data be made available?

Upon publication of research results.

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

Software datasets produced during the project will be published under the GNU General Public License version 3.0.
Other types of datasets will be published under the Creative Commons Attribution 4.0 International license.

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.

- Yes

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

Given the small total size of the datasets, the storage of all the publicly available data are expected to be free of charge.

6. Responsibilities

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

The researcher

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

The researcher, assisted by the IT team of the Institute of Astronomy.

Who will manage data preservation and sharing?

The researcher during the project, the supervisor assisted by the IT team of the Institute of Astronomy after the project has completed.

Who will update and implement this DMP?

The researcher during the project, the supervisor assisted by the IT team of the Institute of Astronomy after the project has completed.