Characterization of evolved planetary systems around pulsating white dwarfs

A Data Management Plan created using DMPonline.be

Creator: Murat Uzundag

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: Murat Uzundag

Data Manager: Murat Uzundag

Project Administrator: Murat Uzundag

Grant number / URL: 1247624N

ID: 206609

Start date: 01-10-2023

End date: 30-09-2026

Project abstract:

White dwarf stars (WDs) are the end evolutionary state of low- and
intermediate-mass stars, which comprise more than 95% of all stars in our
galaxy. WDs play a unique and fundamental role in our understanding of
the formation and evolution of stars and planetary systems. An important
subset of WDs presents photometric variability caused by a range of
phenomena, including pulsations, binarity, and more recently transits by both
planets and planetary debris. Therefore, studying photometric variations in
WDs provides an invaluable source of information concerning the end stages of
stars and planets. This research statement focuses on identifying and
characterizing two types of variable WDs: those being transited by
planets/circumstellar planetary debris and those that pulsate.

Last modified: 22-04-2024

Characterization of evolved planetary systems around pulsating white dwarfs
Application DMP

Questionnaire

Describe the datatypes (surveys, sequences, manuscripts, objects …) the research will collect and/or generate and /or (re)use. (use up to 700 characters)

The research project encompasses a diverse range of data types sourced from both space and ground-based surveys. Predominantly, the data is stored in the standard FITS format. It includes photometric and spectroscopic data sourced from various surveys. Throughout the study, these datasets will undergo post-processing, resulting in associated output files saved as .png for images and spectra, or .csv for tables, tailored to the specific data type. The publications describing the complete scientific method will be available from the appropriate scientific journals and KU Leuven Lirias, and additional astronomical models and model output will be stored on CDS database (cds.u-strasbg.fr).

Specify in which way the following provisions are in place in order to preserve the data during and at least 5 years after the end of the research? Motivate your answer. (use up to 700 characters)

1. Designation of responsible person (If already designated, please fill in his/her name.): Dr. Murat Uzundag, as the FWO Junior Postdoctoral Fellow, will handle the day-to-day data management tasks. Over the long term, Prof. Conny Aerts, serving as the supervisor of the junior postdoctoral project at KU Leuven, will be the responsible person in the long-term.
2. Storage capacity/repository
   - during the research: The co-host institute, Institute of Astronomy, Department of Physics & Astronomy at KU Leuven, has provided ample storage capacity to meet the project's needs.
   - after the research: After the project, the datasets can be transferred to a KU Leuven repository (if needed).
     However, all the data outputs will already be kept in an online publicly available repository as explained in the previous subsection.

What's the reason why you wish to deviate from the principle of preservation of data and of the minimum preservation term of 5 years? (max. 700 characters)

N/A

Are there issues concerning research data indicated in the ethics questionnaire of this application form? Which specific security measures do those data require? (use up to 700 characters)

N/A

Which other issues related to the data management are relevant to mention? (use up to 700 characters)

N/A

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.

Dataset Name	Description	New or reused	Digital or Physical	Only for digital data Digital Data Type	Only for digital data Digital Data format	Only for digital data Digital data volume (MB/GB/TB)	Only for physical data 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, .csv,.pdf, .txt, .rtf, .dwg, .gml, … • NA	*Please choose from the following options:* • <100MB • <1GB • <100GB • <1TB • <5TB • <10TB • <50TB • >50TB • NA	
Searching for variability on polluted white dwarfs	The research focuses on identifying variability in 179 stars using data from the TESS satellite, spanning multiple sectors.	Generate new data	Digital	Observational	.csv	<100GB	N/A
A comprehensive search for hot subdwarf stars using Gaia and TESS	The analysis of TESS and Gaia datasets involved examining around 3000 light curves.	Generate new data	Digital	Observational	.csv, .pdf, .tab	<100GB	N/A
Asteroseismological analysis of pulsating compact stars	The discoveries of new H-deficient pulsating white dwarfs (WD) and pre-white dwarfs (pre-WD), alongside the theoretical instability strips and the asteroseismic analysis of these pulsators.	Generate new data	Digital	Observational Modelling	.csv	<100GB	N/A
Comparing spectroscopic, seismic, and astrometric masses of pulsating white dwarfs	The fundamental parameters of pulsating white dwarfs will be measured through the application of various approaches, including spectroscopy, asteroseismology, and astrometry, facilitating comprehensive testing.	Generate new data	Digital	Observational Modelling	.csv, .pdf, .tab	<100GB	N/A

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:

Datasets relevant to the published manuscript will be provided online. Data pertaining to hot subdwarfs and low-mass red giants will be archived at https://a15.astro.physik.uni-potsdam.de/w/projects/. For data concerning white dwarfs, upon submission, it will be accessible to the public via https://www.montrealwhitedwarfdatabase.org/. The models are available at: http://evolgroup.fcaglp.unlp.edu.ar/TRACKS/tracks.html

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 all the survey dataset will be given in cds.u-strasbg.fr with README.txt files. Such example can already be found in https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/684/A118 in the case of hotsubdwarfs.
Data pertaining to hot subdwarfs and low-mass red giants will also be archived at https://a15.astro.physik.uni-potsdam.de/w/projects/. For data concerning white dwarfs, upon submission, it will be accessible to the public via https://www.montrealwhitedwarfdatabase.org/.

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

A significant volume of data in the standard FITS format will be handled, encompassing photometric data sourced from missions such as Kepler/K2, TESS, and PLATO, alongside astrometric data from Gaia. Throughout the study, these datasets will be subjected to post-processing, resulting in the preservation of associated output files, which may manifest as .png images for visual representations and spectra, or .csv files for tabular data. All the dataset can be found in https://a15.astro.physik.uni-potsdam.de/w/projects/ and cds.u-strasbg.fr.

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

Where will the data be stored?

The Institute of Astronomy has an effective internal backup and storage mechanism that will enable ongoing data preservation. The publications describing the complete scientific method will be available from the appropriate scientific journals and KU Leuven Lirias, and additional astronomical models and model output will be stored on CDS database (cds.u-strasbg.fr) when applicable. In addition, all original astronomical observations and metadata are stored on public data repositories of the observatories. Any relevant codes will be made GitHub-accessible such as https://github.com/muratuzundag.

How will the data be backed up?

The IT staff will conduct routine backups. Nevertheless, we suggest a series of strategies for ensuring data reproducibility in observed dataset and model simulations.

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 primary storage & backup capacity is around 3 TB.

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

All data and model source codes will adhere to IT guidelines, ensuring that all computers and external drives are password protected. The IT staff consistently performs security upgrades. When sharing datasets, read-only permissions will be enforced.

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

We anticipate no supplementary expenses for data storage or backup, as the operational costs of the IT department are covered by the co-host institute, the KU Leuven.

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 output of the analyzed dataset and the models will be retained.

Processed Data: Processed datasets, including cleaned and analyzed data, as well as intermediate data generated during data processing pipelines, will be retained.

Documentation: Comprehensive documentation, including project plans, data dictionaries, codebooks, and metadata, will be explained on the manuscript that uploaded in Lirias.

Model Codes: Source code and algorithms used for data analysis and modeling will be preserved in mostly GitHub and Lirias to ensure reproducibility of results.

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

Major part of these data will be stored on the CDS database (cds.u-strasbg.fr) repository. If needed, non-published results can be stored at an internal
workstation as well.

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

The CDS repository is free of charge.

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

The published results will be made available in an open access repository (CDS).
The models will only be made available in http://evolgroup.fcaglp.unlp.edu.ar/TRACKS/tracks.html.

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

The sharing of the code will occur following mutual agreement, facilitated by a rules of agreement to be established through email correspondence.

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.

Most of the dataset can be found in .FITS format, which can be obtained from http://archive.eso.org/eso/eso_archive_main.html.

Also other big dataset will be stored in .csv format or can be download some other formats as well in https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/684/A118

When will the data be made available?

Publication datasets will immediately be made available upon the acceptance of manuscripts.

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

We will provide open access licenses for data usage.

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?

All the repositories such as CDS and Github are free of charge.

6. Responsibilities

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

Murat Uzundag

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

Murat Uzundag

Who will manage data preservation and sharing?

Murat Uzundag

Who will update and implement this DMP?

Murat Uzundag