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	Karl Meerbergen 0000-0002-1508-0248
Contributor name(s) (+ ORCID) & roles	Wim Michiels 0000-0002-0877-0080 (co-promotor)
Project number [1] & title	3E240025 - Global methods for nonlinear eigenvalue problems with eigenvector nonlinearities
Funder(s) GrantID [2]	G027624N
Affiliation(s)	X KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: ROR identifier KU Leuven: 05f950310

---

[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.

| Please provide a short project description | Nonlinear eigenvalue problems with eigenvector nonlinearities are algebraic eigenvalue problems whose matrix depends on the eigenvector. Applications range from electron structure computations to machine learning.

This class of eigenvalue problems is challenging as they are inherently nonlinear while the concept of linearization, in the sense of an equivalent, possibly infinite-dimensional linear representation, is not clear. As a consequence, the basis theory is incomplete, while the available algorithms almost exclusively rely on a fixed point iteration whose global convergence properties are only understood in specific cases. At the same time, the literature is limited and scattered over different domains.

The aim of the project is to develop structure exploiting algorithms for eigenvector nonlinearities with strong theoretical foundations. First, multivariable rational approximations will be constructed that induce spectrally equivalent linear problems. Second, global algorithms will be developed, able to reliably compute all eigenvalues in a prescribed region or interval. Third, using projection these algorithms will be scaled up to large-scale matrices, which is essential to bridge methods and applications.

The methodology integrates techniques from rational approximation, algorithms for multi-parameter eigenvalues and methods from algebraic geometry, and it exploits the complementary expertise of the applicants. |

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
Publications and presentation data	Publications, and presentation of data Publications generated using LaTeX, Microsoft Word, Powerpoint.	New	Digital	Textual	.docx/.pptx/.tex/.pdf/.png/.jpg	<5GB	
Research software	Software (Matlab,Python, Julia,C++,Fortran)/manual(text/html)/test cases.	Reused New	Digital Numerical	Textual	.m/.py/.jl/.cpp/.f/.txt/.dat/.html	<10GB	

---

[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.	• Problem description and software published with journal articles (source: journal website/author's repositories).
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: S67418 ☐ 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: S67418
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.*	1. Software is stored at gitlab.kuleuven.be and nextcloud.cs.kuleuven.be, using version control. README.txt files explain how to use the software and run the text cases. 2. For each paper, a description file including the exact command line is provided for reproducibility.

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: If no, please specify (where appropriate per dataset or data type) which metadata will be created: Data and version of the software for each publication will be stored with the publication in text files.

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: gitlab.kuleuven.be, Nextcloud.cs.kuleuven.be
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 Yes: NetApp storage is available at the department of computer science; gitlab.kuleuven.be has sufficiently large storage. In the case of excessive use of storage, we will be notified. In fact, the storage of software and data to run the software is not very large, since it is all text, and problem descriptions are very small. 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*	Research data are stored and managed by the KU Leuven ICTS and the CS department and are accessible only by the researchers working on the project.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	The Gitlab repository server is offered by KU Leuven. The Dept CS NetApp datastorage is covered by the budget of this project. After this project ends this cost is taken over by the research group(s)' reserves.

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 Dept CS NetApp storage solution (>10TB) will retain snapshots for every publication (source code, experiments, manuscript) for at least 10 years.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	The Dept CS NetApp longterm storage is covered by the research group's reserves. For this project, this is expected to be <100 euros / year (this does not include costs for computation which is only relevant during the project).

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: Latex files stay private, the publications (pdf) are available through the journals websites. Software and data to run the software are open, via gitlab.kuleuven.be.
If access is restricted, please specify who will be able to access the data and under what conditions.	The researchers of the project will be able to access the latex files.
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): gitlab.kuleuven.be ☒ Other (specify): nextcloud.cs.kuleuven.be

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?	We don't expect any costs regarding data sharing.

7. Responsibilities

Who will manage data documentation and metadata during the research project?	Prof Karl Meerbergen and Professor Wim Michiels.

Who will manage data storage and backup during the research project?	ICTS (KU Leuven) and CS (KU Leuven)
Who will manage data preservation and sharing?	Karl Meerbergen and Wim Michiels (PIs)
Who will update and implement this DMP?	Karl Meerbergen (PI) bears the end responsibility of updating & implementing this DMP.