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	**Daans Nicolas, 0000-0003-2217-7758**
Contributor name(s) (+ ORCID) & roles	**Raf Cluckers (project supervisor), 0000-0002-3056-3280**
Project number [1] & title	**Decidability of polynomial equations over function fields**
Funder(s) GrantID [2]	**1208225N**
Affiliation(s)	☐ **KU Leuven** ROR identifier KU Leuven: 05f950310
Please provide a short project description	The aim of this project is to study decidability of the problem of solving polynomial equations over function fields, also known as "Hilbert's 10th Problem over function fields". While several specific cases have been studied over the last decades, a general framework has so far not been developed. Inspired by a recent breakthrough stemming from the fellow's PhD research, the current project seeks to push the frontier of our understanding to include function fields over (almost) arbitrary base fields.

[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.

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
Unpublished notes	Partially or fully developed mathematical ideas, not (yet) suitable for publication. Including in particular developed theorems and their proofs.	☐ **Generate new data**	☐ **Digital** ☐ (occasionally) **Physical**	☐ **Textual** (including source code) ☐ (occasionally) **Images**	*.tex, *.bib, *.bbl, *.pdf, *.jpg	☐ **< 100 MB**	Negligible amounts of paper notes.
Source files of manuscripts	Source files used to compile manuscripts (papers) to be published and shared with the mathematical community.	☐ **Generate new data**	☐ **Digital**	☐ **Textual** (including source code)	*.tex, *.bib, *.bbl	☐ **< 10 MB**	NA

[3] Add rows for each dataset you want to describe.

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 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.	NA
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.	☐ **No**
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).	☐ **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**

[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.	☐ **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 to what data they relate and which restrictions will be asserted.	☐ **No**

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*.	The research is fundamental and theoretical in nature without empirical component, and the value of the output lies entirely in the content of the produced manuscripts, to be written in plain English so as to be readable by any researcher with the appropriate general background. The manuscripts should as such document themselves and need no additional documentation to be correctly and fully understood. Two minor aspects to keep in mind before the research has been fully written up and published in manuscripts: • *Documentation of version history:* As mathematical ideas may be rewritten in different forms (e.g. to improve the presentation), it may be important to keep track not just of different versions which have existed, but also what motivated the choice in favour of one of them in the final presentation. For larger subprojects and subprojects which involve collaboration with other researchers, I will make use of version history software (namely Git, or Overleaf) which allows to include metadata documenting the changes made in each version. • *Annotation of source code:* Mathematical manuscripts will be written using LaTeX (rather than, e.g., Microsoft Word). Since there can be slight variations between LaTeX compilers, I will add sufficient annotations to the source code documenting the required dependencies and macros, to ensure that one can easily make small modifications to the code as required by different compilers. Both of these aspects are only relevant to my own management of the research and editorial process up until the moment of publication, after which point the unpublished notes are of no more additional value to the research community.
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.*	☐ **No** As explained above, the metadata produced during the writing stage of manuscripts (annotation of version history and source code) is only of value to myself before publication, not to other researchers.

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.*	☐ **OneDrive (KU Leuven)**, synchronized with desktop PC at the office and private laptop.
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)** For work on larger manuscripts and/or manuscripts with co-authors, regular manual back-ups will be made to a reputable platform with full version history (like GitHub or another platform supporting Git) to ensure no data is lost in the editing process. When working from home, data will also be automatically backed up on a private server.
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** (volume of digital data is negligible, see section 2)
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*	All digital data is regularly backed up both automatically and manually, as described above, with full version history where relevant, and in multiple ways. The chance of all such systems simultaneously failing or being the target of a dedicated malicious actor is very low. Physical data (handwritten notes) is temporarily kept in my (locked) office, until promptly digitized, after which the handwritten notes may be discarded. Only the project supervisor and myself will have access to the data; parts of the data may be shared in the context of collaboration with other researchers (e.g. co-authors) where needed. Standard good cyber security practice will be followed to prevent unauthorized access to the data (cloud storage protected by 2-factor authorization, accessing servers through (virtual) private networks, etc.). Since the data is not sensitive (no ethical/legal/privacy-related concerns; see section 2), no additional measures to protect the data against unauthorized access are deemed necessary.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?	The amount of data collected is very small (see section 2), and its storage thus completely covered by the host institution's offered cloud storage, at no extra cost.

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*** * Once research output has been fully integrated in a research paper (manuscript), unpublished notes whose content has been completely subsumed by the research paper will not be kept, as they are of no value.
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.*	☐ **Other** (specifiy): Source files of manuscripts (AAM) will be kept indefinitely on the reputable preprint server arXiv, freely available, together with their compiled (pdf) versions. Unpublished notes which have been fully integrated in a research paper, will not be archived. Unpublished notes on research carried out during this project but not yet published at the end of the fellowship, will be retained by myself on a private server as mentioned above, and also handed over to the supervisor of the project, who will take responsibility of their continued safekeeping on KU Leuven (cloud) storage, until the research has been published.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	The amount of data collected is very small (see section 2), and its storage thus completely covered by the host institution's cloud storage, at no extra cost.

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** Source files of manuscripts (AAM) will be kept indefinitely on the reputable preprint server arXiv. Unpublished notes will not be made available.
If access is restricted, please specify who will be able to access the data and under what conditions.	NA
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.	☐ **No**
Where will the data be made available? If already known, please provide a repository per dataset or data type.	☐ **Other data repository** (specify) **https://www.arxiv.org/**
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. *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.*	☐ **Other** (specify) Source files of manuscripts will be uploaded to arXiv under a non-exclusive perpetual license to distribute. Although source files will be freely available, it is not the intended usage that they are reused verbatim by other researchers.
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** (DOI) **will be added upon deposit in a data repository**
What are the expected costs for data sharing? How will these costs be covered?	No expected costs; the preprint server arXiv is managed by Cornell University and funded by donations of research institutions and non-profit organisations. No charges are levied on author submissions.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	**Nicolas Daans**
Who will manage data storage and backup during the research project?	**Nicolas Daans**
Who will manage data preservation and sharing?	**Nicolas Daans** until the end of the project, afterwards to be coordinated between **Nicolas Daans** and **Raf Cluckers**
Who will update and implement this DMP?	**Nicolas Daans**