

FWO DMP Template - Flemish Standard Data Management Plan

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	Davide Mattiolo, ORCID: 0000-0001-5876-0887
Contributor name(s) (+ ORCID) & roles	Prof. Jan Goedgebeur, ORCID: 0000-0001-8984-2463 (Supervisor)
Project number ¹ & title	Project number: 1268323N Title: Connections between edge-colorings, factors and flows in cubic graphs
Funder(s) GrantID ²	Research Foundation - Flanders (FWO)
Affiliation(s)	<input checked="" type="checkbox"/> KU Leuven <input type="checkbox"/> Universiteit Antwerpen <input type="checkbox"/> Universiteit Gent <input type="checkbox"/> Universiteit Hasselt <input type="checkbox"/> Vrije Universiteit Brussel <input type="checkbox"/> Other: Provide ROR ³ identifier when possible:

¹ "Project number" refers to the institutional project number. This question is optional since not every institution has an internal project number different from the GrantID. 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.

³ Research Organization Registry Community. <https://ror.org/>

Please provide a short project description	<p>Many old conjectures regarding the structure of graphs, such as the 5-flow Conjecture, are stated for large classes of graphs. Classical results of the last century prove that most of such open problems are true if and only if they are true for all cubic graphs (i.e., graphs where all vertices have degree 3). In this project we aim to tackle open problems on cubic graphs which are related to some major conjectures in the field. In the first research line, our main objective is to prove the existence of normal 6- (and 5-)edge-colorings on families of cubic graphs. These are proper edge-colorings of cubic graphs with the property that, for every edge e, the number of colors of the edges adjacent to e is either 2 or 4. In the second research line we aim to solve problems about perfect matchings in cubic graphs. Our main objectives here are to prove new connections between conjectures in this field and to verify them on several subclasses of cubic graphs. The third research line is devoted to vector flows in cubic graphs. Vector flows are a generalization of the classical concept of flow, where each flow value is a n-dimensional vector with some constraints on its norm. The first natural step is the case $n=2$. The major objectives in this research line are to investigate structural properties of 2-dimensional vector flows in cubic graphs and to design an algorithm that, given a graph, computes the best possible 2-dimensional vector flow that can be defined on that graph.</p>
--	---

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 data⁴.

Dataset Name	Description	New or Reused	Digital or Physical	ONLY FOR DIGITAL DATA	ONLY FOR DIGITAL DATA	ONLY FOR DIGITAL DATA	ONLY FOR PHYSICAL DATA
				Digital Data Type	Digital Data Format	Digital Data Volume (MB, GB, TB)	Physical Volume
		<input type="checkbox"/> Generate new data <input type="checkbox"/> Reuse existing data	<input type="checkbox"/> Digital <input type="checkbox"/> Physical	<input type="checkbox"/> Observational <input type="checkbox"/> Experimental <input type="checkbox"/> Compiled/aggregated data <input type="checkbox"/> Simulation data <input type="checkbox"/> Software <input type="checkbox"/> Other <input type="checkbox"/> NA	<input type="checkbox"/> .por <input type="checkbox"/> .xml <input type="checkbox"/> .tab <input type="checkbox"/> .csv <input type="checkbox"/> .pdf <input type="checkbox"/> .txt <input type="checkbox"/> .rtf <input type="checkbox"/> .dwg <input type="checkbox"/> .tab <input type="checkbox"/> .gml <input type="checkbox"/> other: <input type="checkbox"/> NA	<input type="checkbox"/> < 100 MB <input type="checkbox"/> < 1 GB <input type="checkbox"/> < 100 GB <input type="checkbox"/> < 1 TB <input type="checkbox"/> < 5 TB <input type="checkbox"/> < 10 TB <input type="checkbox"/> < 50 TB <input type="checkbox"/> > 50 TB <input type="checkbox"/> NA	
Papers	Manuscripts in which the new results will be presented	Generate new data	Digital	Other: papers	.pdf	<1 GB	
Source code	Computer programs that	Generate new data. Reuse existing data.	Digital	Software	C++, Java, and/or Python code	< 100 MB	

⁴ Add rows for each dataset you want to describe.

	we will develop						
Graphs	New list of graphs that we will generate	Generate new data. Reuse existing data.	Digital	Graphs in graph6 or adjacency list format	.txt	< 100 GB	

GUIDANCE:

DATA CAN BE DIGITAL OR PHYSICAL (FOR EXAMPLE BIOBANK, BIOLOGICAL SAMPLES, ...). DATA TYPE: DATA ARE OFTEN GROUPED BY TYPE (OBSERVATIONAL, EXPERIMENTAL ETC.), FORMAT AND/OR COLLECTION/GENERATION METHOD.

EXAMPLES OF DATA TYPES: OBSERVATIONAL (E.G. SURVEY RESULTS, SENSOR READINGS, SENSORY OBSERVATIONS); EXPERIMENTAL (E.G. MICROSCOPY, SPECTROSCOPY, CHROMATOGRAMS, GENE SEQUENCES); COMPILED/AGGREGATED DATA⁵ (E.G. TEXT & DATA MINING, DERIVED VARIABLES, 3D MODELLING); SIMULATION DATA (E.G. CLIMATE MODELS); SOFTWARE, ETC.

EXAMPLES OF DATA FORMATS: TABULAR DATA (.POR, .SPSS, STRUCTURED TEXT OR MARK-UP FILE XML, .TAB, .CSV), TEXTUAL DATA (.RTF, .XML, .TXT), GEOSPATIAL DATA (.DWG, .GML, ..), IMAGE DATA, AUDIO DATA, VIDEO DATA, DOCUMENTATION & COMPUTATIONAL SCRIPT.

DIGITAL DATA VOLUME: PLEASE ESTIMATE THE UPPER LIMIT OF THE VOLUME OF THE DATA PER DATASET OR DATA TYPE.

PHYSICAL VOLUME: PLEASE ESTIMATE THE PHYSICAL VOLUME OF THE RESEARCH MATERIALS (FOR EXAMPLE THE NUMBER OF RELEVANT BIOLOGICAL SAMPLES THAT NEED TO BE STORED AND PRESERVED DURING THE PROJECT AND/OR AFTER).

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.	We will possibly reuse existing generators or existing lists of snarks. The sources are listed here. <ol style="list-style-type: none"> 1. http://users.cecs.anu.edu.au/~bdm/nauty/ 2. https://houseofgraphs.org/
---	---

⁵ These data are generated by combining multiple existing datasets.

<p>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, please describe these issues further and refer to specific datasets or data types when appropriate.</p>	<p> <input type="checkbox"/> Yes, human subject data <input type="checkbox"/> Yes, animal data <input type="checkbox"/> Yes, dual use <input checked="" type="checkbox"/> No If yes, please describe: </p>
<p>Will you process personal data⁶? If so, briefly describe the kind of personal data you will use. Please refer to specific datasets or data types when appropriate. If available, add the reference to your file in your host institution's privacy register.</p>	<p> <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No If yes: <ul style="list-style-type: none"> - Short description of the kind of personal data that will be used: - Privacy Registry Reference: </p>
<p>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.</p>	<p> <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No If yes, please comment: </p>
<p>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.</p>	<p> <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No If yes, please explain: </p>

⁶ See Glossary Flemish Standard Data Management Plan

<p>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.</p>	<p><input type="checkbox"/> Yes</p> <p><input checked="" type="checkbox"/> No</p> <p>If yes, please explain:</p>
---	--

3. Documentation and Metadata	
<p>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).</p>	<ul style="list-style-type: none"> - All source code developed in this project will be released under an open source license and will be hosted on a public GitHub repository and will include a README file with clear instructions on how to use the software. The code will also contain comments so other researchers can also build upon the software, if wanted. - The new lists of graphs will be uploaded to the online searchable database of interesting graphs called the House of Graphs (https://houseofgraphs.org). - All manuscripts will be uploaded to the preprint server arXiv.org.

<p>Will a metadata standard be used to make it easier to find and reuse the data?</p> <p>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.</p> <p><i>REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS.</i></p>	<p><input type="checkbox"/> Yes <input checked="" type="checkbox"/> No</p> <p>If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used:</p> <p>If no, please specify (where appropriate per dataset or data type) which metadata will be created:</p> <ul style="list-style-type: none"> - All software will include a README with instructions on how to use the software. - The new graphs will be uploaded to the House of Graphs (https://houseofgraphs.org) where metadata will be added on how the graphs were obtained and where they can be downloaded in the most used graph formats (e.g. graph6 and as adjacency lists).
---	--

4. Data Storage & Back-up during the Research Project

<p>Where will the data be stored?</p>	<p>All data will be stored on the KU Leuven OneDrive storage and the group's network drives. Moreover, the data will also be available as follows:</p> <ol style="list-style-type: none"> 1) Manuscripts will be uploaded to arXiv.org; 2) The developed source code will be released under an open-source license and hosted on a public GitHub repository; 3) The new lists of graphs will be uploaded to the online searchable database of interesting graphs called the House of Graphs (https://houseofgraphs.org).
---------------------------------------	---

<p>How will the data be backed up?</p> <p><i>WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS? DESCRIBE THE LOCATIONS, STORAGE MEDIA AND PROCEDURES THAT WILL BE USED FOR STORING AND BACKING UP DIGITAL AND NON-DIGITAL DATA DURING RESEARCH.⁷</i></p> <p><i>REFER TO INSTITUTION-SPECIFIC POLICIES REGARDING BACKUP PROCEDURES WHEN APPROPRIATE.</i></p>	<p>All data will be stored on the KU Leuven OneDrive storage and the group's network drives, which are automatically backed up by the department.</p> <p>The manuscripts will also be available on arXiv.org and the code on GitHub. The new graphs will be uploaded to the House of Graphs, which was developed and is still maintained by Prof. Jan Goedgebeur and which is automatically backed up by the department.</p>
<p>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.</p>	<p><input checked="" type="checkbox"/> Yes <input type="checkbox"/> No</p> <p>If yes, please specify concisely: Every KU Leuven employee has 2 TB of available storage space and the department has large network drives (which are automatically backed up) for each group. This is more than enough space to store our manuscripts, source code, and new graphs.</p> <p>If no, please specify:</p>
<p>How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?</p> <p><i>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.⁷</i></p>	<p>The KU Leuven One Drive is completely personal and not accessible by other people. Similarly, the folders where the project's data will be stored on the group's network drives will only be accessible by the members who are involved in this project.</p> <p>In any case, we are not working with sensible data, and we will release everything as open source. So, this problem does not apply here.</p>
<p>What are the expected costs for data storage and backup during the research project? How will these costs be covered?</p>	<p>No costs.</p>

⁷ Source: Ghent University Generic DMP Evaluation Rubric: <https://osf.io/2z5g3/>

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...).	All produced data, i.e. manuscripts, source code and generated graphs, will be retained for at least five years after the end of the project. Prof. Jan Goedgebeur will be responsible for the data retention for at least five years after the end of the project.
Where will these data be archived (stored and curated for the long-term)?	All data will be kept in the personal KU Leuven OneDrive of Prof. Goedgebeur and in the group's network storage of the department of Computer Science.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	No costs are expected. The storage is provided by KU Leuven and/or the department. If in the future the department would ask us to pay for this storage, we expect that this won't cost more than about 100 euro per year.

6. Data Sharing and Reuse

<p>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.</p> <p><i>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/#INFOEU-REPO-ACCESSRIGHTS</i></p>	<p><input checked="" type="checkbox"/> Yes, in an Open Access repository <input type="checkbox"/> Yes, in a restricted access repository (after approval, institutional access only, ...) <input type="checkbox"/> No (closed access) <input type="checkbox"/> Other, please specify:</p> <p>All data will be made available as follows:</p> <ol style="list-style-type: none"> 1. Manuscripts will be uploaded to arXiv.org; 2. The developed source code will be released under an open-source license and hosted on a public GitHub repository; 3. The new graphs will be uploaded to the House of Graphs.
<p>If access is restricted, please specify who will be able to access the data and under what conditions.</p>	
<p>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.</p>	<p><input type="checkbox"/> Yes, privacy aspects <input type="checkbox"/> Yes, intellectual property rights <input type="checkbox"/> Yes, ethical aspects <input type="checkbox"/> Yes, aspects of dual use <input type="checkbox"/> Yes, other <input checked="" type="checkbox"/> No</p> <p>If yes, please specify:</p>
<p>Where will the data be made available? If already known, please provide a repository per dataset or data type.</p>	<p>All data will be available at the following links:</p> <ul style="list-style-type: none"> • Manuscripts: https://arxiv.org/ • Source codes: https://github.com/ • Generated graphs: https://houseofgraphs.org/

<p>When will the data be made available?</p> <p><i>THIS COULD BE A SPECIFIC DATE (DD/MM/YYYY) OR AN INDICATION SUCH AS 'UPON PUBLICATION OF RESEARCH RESULTS'.</i></p>	<p>All data will be made available upon publication of the research results.</p>
<p>Which data usage licenses are you going to provide? If none, please explain why.</p> <p><i>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.</i></p> <p><i>EXAMPLE ANSWER: E.G. "DATA FROM THE PROJECT THAT CAN BE SHARED WILL BE MADE AVAILABLE UNDER A CREATIVE COMMONS ATTRIBUTION LICENSE (CC-BY 4.0), SO THAT USERS HAVE TO GIVE CREDIT TO THE ORIGINAL DATA CREATORS." ⁸</i></p>	<p>All data will be made available as open access under Creative Commons Licence, see https://creativecommons.org/licenses/</p> <p>The source code will be publicly available as open source software under the GNU GPLv3 license.</p>
<p>Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here.</p> <p><i>INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.</i></p>	<p><input type="checkbox"/> Yes <input checked="" type="checkbox"/> No If yes:</p>
<p>What are the expected costs for data sharing? How will these costs be covered?</p>	<p>No costs.</p>

⁸ Source: Ghent University Generic DMP Evaluation Rubric: <https://osf.io/2z5g3/>

7. Responsibilities

Who will manage data documentation and metadata during the research project?	Davide Mattiolo
Who will manage data storage and backup during the research project?	Davide Mattiolo
Who will manage data preservation and sharing?	Davide Mattiolo and Prof. Jan Goedgebeur
Who will update and implement this DMP?	Davide Mattiolo and Prof. Jan Goedgebeur