

Data Management

C1-BETWEEN LINEAR AND NONLINEAR: NEW TENSOR PERSPECTIVES

ADMIN DETAILS

Project Name: C1-Between linear and nonlinear: new tensor perspectives

Principal Investigator / Researcher: Lieven De Lathauwer

Institution: KU Leuven

1. GENERAL INFORMATION

Name of the project lead (PI)

Lieven De Lathauwer

Internal Funds Project number & title

C14/22/096

Between linear and nonlinear: new tensor perspectives

2. DATA DESCRIPTION

2.1. Will you generate/collect new data and/or make use of existing data?

- Generate new data

2.2. What data will you collect, generate or reuse? Describe the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a numbered list or table and per objective of the project.

Random data will be generated using scripts to assess the performance of the various algorithms developed in this project. The data are therefore stored as scripts with fixed random number generators.

Type	synthetic numerical data
Format	.mat
Size	negligible
How created	by running numerical script

3. ETHICAL AND LEGAL ISSUES

3.1. Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to the file in KU Leuven's Record of Processing Activities. Be aware that registering the fact that you process personal data is a legal obligation.

No personal data will be used.

3.2. 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, add the reference to the formal approval by the relevant ethical review committee(s).

NA

3.3. Does your research possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

NA

3.4. Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions regarding reuse and sharing are in place?

There are no 3rd party agreements restricting dissemination and exploitation of the data.

4. DOCUMENTATION AND METADATA

4.1. What documentation will be provided to enable understanding and reuse of the data collected/generated in this project?

The seeds for the random generator are provided.

4.2. Will a metadata standard be used? If so, describe in detail which standard will be used. If not, state in detail which metadata will be created to make the data easy/easier to find and reuse.

Metadata with specific information about the numerical runs will be stored in associated text files.

5. DATA STORAGE AND BACKUP DURING THE PROJECT

5.1. Where will the data be stored?

The algorithms and codes will be kept on Gitlab repository.

When a work package is finished, the dataset moves to the STADIUS Dataset Server which is regularly backed up.

5.2. How will the data be backed up?

The storage used by the research group is backed up daily. The backup process is managed by the IT division of the department.

5.3. 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.

There is sufficient storage on the servers of our research unit. The capacity is being evaluated regularly.

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

The storage facilities of the research unit are available for the researchers for free.

5.5. Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The data stored in the ESAT servers has access regulated by an access control list (ACL) that grants: read-write access to the project owner read-only access to specific users The ACL is managed by the project owner. Client computers can access the data using: SMB2 (or higher) from specific IP ranges NFSv4 from specific (IT managed) systems.

6. DATA PRESERVATION AFTER THE END OF THE PROJECT

6.1. Which data will be retained for the expected 10 year period after the end of the project? If only a selection of the data can/will be preserved, clearly state why this is the case (legal or contractual restrictions, physical preservation issues, ...).

All data relevant for the project will be stored the period of 10 years. If the data generated during the project is no longer in use after 10 years, it will be removed from the STADIUS servers and moved to offline archive, or removed completely.

6.2. Where will these data be archived (= stored for the long term)?

The data will be stored at research unit's servers for at least 10 years, conform the KU Leuven RDM policy.

6.3. What are the expected costs for data preservation during these 10 years? How will the costs be covered?

The research unit has a storage capacity allowing to store data at no extra costs to the project.

7. DATA SHARING AND RE-USE

7.1. Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions or because of IP potential)?

There are no factors preventing from sharing the data with a 3rd party.

7.2. Which data will be made available after the end of the project?

Selected publications will be added to Tensorlab+, our reproducible research repository for tensor computations. This includes seeds for random number generators used in numerical experiments.

7.3. Where/how will the data be made available for reuse?

Tensorlab+: <https://www.tensorlabplus.net/>

7.4. When will the data be made available?

- Upon publication of the research results

Upon publication of the research results - or shortly after.

7.5. Who will be able to access the data and under what conditions?

Tensorlab+ is freely available for academic research (non-profit).

License: <https://www.tensorlabplus.net/license.html>

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

There are no additional costs of data sharing foreseen.

8. RESPONSIBILITIES

8.1. Who will be responsible for the data documentation & metadata?

The first author of a publication is responsible for managing the data during his/her employment at KU Leuven; afterwards, the responsibility moves to the promoter (Lieven De Lathauwer). A data management officer is assigned to guide this process. At this moment, the officer is dr. Nico Vervliet. The PI, Lieven De Lathauwer will be end responsible for data documentation and metadata.

8.2. Who will be responsible for data storage & back up during the project?

The data storage and back up is the responsibility of the IT support of the faculty.

8.3. Who will be responsible for ensuring data preservation and sharing?

The PI, Lieven De Lathauwer will be end responsible for the data preservation and reuse.

8.4. Who bears the end responsibility for updating & implementing this DMP?

The end responsibility for updating and implementing the DMP is with the PI Lieven De Lathauwer.