DMP title

Project Name INTERACTIONS IN MOBILITY-AS-A-SERVICE ECO-SYSTEMS - DMP title
Project Identifier 1S44922N
Grant Title 1S44922N
Principal Investigator / Researcher Gaurav Malik
Project Data Contact gaurav.malik@kuleuven.be
Description SAP project code: 3E200789
Institution KU Leuven

1. General Information

Name applicant

Gaurav Malik

FWO Project Number & Title

Title: INTERACTIONS IN MOBILITY-AS-A-SERVICE ECO-SYSTEMS

Project reference nr: 1S44922N
SAP project code: 3E200789

Affiliation

- KU Leuven

2. Data description

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

- Generate new data
- Reuse existing data

Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).

Used Data:

- Python code: DyntaPy traffic assignment source code developed within the group with open source license.
- Belgian traffic demand available via MOW: namely xls files and shape (.shp) files 1-5 GB.
- CASADI (https://web.casadi.org/): It has a permissive license such that we are allowed to freely use them in a software project with not necessarily the same license.

Generated Data:

- MATLAB and Python codes for optimisation of transport policy variables like tolls.
- PTV VISUM Static Traffic Assignment Models of different settings and cities. 1-50 GB
- xls files for traffic volumes in response of different policy variable values for different cities. 1-10 GB

3. Legal and ethical issues

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.

- No

Privacy Registry Reference: Not Applicable
Short description of the kind of personal data that will be used: Not Applicable

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)

- No

Does your work 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?

- Yes

We work with open source licenses with the restriction that everyone who uses our data or tools must also make their tools or data available open source as well. For commercial use or other use types that are excluded from the open-source license, license needs to be negotiated with KUL/LRD.

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 are in place?

- No

4. Documentation and metadata

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

Raw simulation data will be collected per simulation test, including a txt file with a clear description of what the data represent and how they were generated. This text file will also be available in the GIT repository for anyone willing to re-use the data internally or externally. The input files used for the simulation will be kept in the same folder.

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

- No

Most simulation data will be stored in CSV and xls files and will be accompanied by metadata headers. We will ensure all metadata are described consistently during the project and well documented. This will facilitate the reuse of our datasets, both internally and externally.

5. Data storage and backup during the FWO project

Where will the data be stored?

1. All the tools and data are regularly committed to the Git repository of the group i.e. ITSCrealab and/or to KUL OneDrive folders during the active research phase.
2. In addition, a copy is stored on researcher's work computer which has already been purchased prior to FWO grant .
3. The outcome data will be stored on Gitlab, split in public (for open-source output) and private folders (for internal reuse).

How is backup of the data provided?

KU Leuven provided Gitlab and OneDrive have adequate backup strategies in place. In case of data loss or damage, we can easily revert to our software or data.

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

Yes. KUL OneDrive and Gitlab quota are sufficient for the entirety of the project.

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

Currently, existing quota are sufficient. KUL offers large-volume storage and archiving services (https://admin.kuleuven.be/icts/onderzoek/dlm) at rates ~€99/TB, but it is not expected that

these will be required for this project as data volumes remain small.

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

Our Git repository can only be modified by the researcher or the group GIT managers.

Only the researcher has access to his work computer. KUL works with multi-factor authentication (https://set.kuleuven.be/set-it/news/for-security-choose-multi-factor-authentication) and is permanently monitoring and updating its web security performance and policy.

6. Data preservation after the FWO project

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).

The GIT repository of the group stays in tact irrespective of people finishing their research and moving on to other projects.

All data related to publications will necessarily be retained along with other important unpublished data. It will also be made publicly available. Currently the group works with public Gitlab pages (https://gitlab.kuleuven.be/ITSCreaLab); this policy can be updated depending on the technological evolution and new options that may come available.

Where will the data be archived (= stored for the longer term)?

GIT Repository, KU Leuven shared drive and as a back up, work computer/hard drive of the researcher will be submitted.

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?

Negligible costs covered by resources of the research group.

7. Data sharing and reuse

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)?

- No

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

All data related to publications including the model setups as well as their results (csv, xls) will be made publicly available along with other important unpublished data.

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

- In an Open Access repository

Full data will be shared publicly via the public folders of Gitlab repository of the group. (https://gitlab.kuleuven.be/ITSCreaLab).

When will the data be made available?

- Immediately after the end of the project
- Upon publication of the research results

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

Datasets will be published under GNU General Public License version 3 (https://www.gnu.org/licenses/gpl-3.0.en.html) which is an open source software license. Only condition to reuse (add or modify) the data/code would be to release anything developed using this data also under the same license.

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

None

8. Responsibilities

Who will be responsible for data documentation & metadata?

Gaurav Malik

Chris Tampère

Who will be responsible for data storage & back up during the project?
Gaurav Malik

Who will be responsible for ensuring data preservation and reuse ?
Chris Tampère

Who bears the end responsibility for updating & implementing this DMP?
The supervisor bears the end responsibility of updating & implementing this DMP.