STRAUSS - sustainable collaborative urban supply chains

A Data Management Plan created using DMPonline.be

Creators: Pieter Vansteenwegen, Johan Joubert

Affiliation: KU Leuven (KUL)

Funder: Vlaams Agentschap Innoveren & Ondernemen (VLAIO)

Template: VLAIO cSBO DMP (Flemish Standard DMP)

Principal Investigator: Pieter Vansteenwegen, Johan Joubert

Grant number / URL: HBC.2023.0008

ID: 207666

Start date: 01-01-2024

End date: 31-12-2026

Project abstract:

The main aim of this project is to lay the scientific foundation for a suite of urban supply chain models and algorithms that will simultaneously support the decisions of three different types of actors: (1) a single company or supply chain actor (SCA), (2) a coalition of collaborating SCAs that jointly optimize their logistics operations, (3) the city, that can steer the logistics operations on its territory by issuing appropriate measures.

We focus on four urban logistics flows: (1) waste logistics, i.e., the removal of waste from within the city, (2) construction logistics, i.e., logistics operations related to building and demolishing real-estate, (3) hospitality and health industry, i.e., the logistics operations for delivering food and commodities to cafés, restaurants, hotels, hospitals, daycare centers, etc., and, (4) grocery delivery logistics, i.e, the delivery of (mainly) food from supermarkets to individual consumers. These flows were chosen not just because of their importance in an urban context, which is likely only to grow in the future, but also because they present important scientific challenges when it comes to modelling and optimizing them in the most appropriate way.

Last modified: 11-06-2024

STRAUSS - sustainable collaborative urban supply chains
VLAIO DMP (Flemish Standard DMP)

1. 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.

				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
		Please choose from the following options: • Generate new data • Reuse existing data	*Please choose from the following options:* • Digital • Physical	*Please choose from the following options:* • Observational • Experimental • Compiled/aggregated data • Simulation data • Software • Other • NA	*Please choose from the following options:* • .por, .xml, .tab, .cvs,.pdf, .txt, .rtf, .dwg, .gml, … • NA	*Please choose from the following options:* • <100MB • <1GB • <100GB • <1TB • <5TB • <10TB • <50TB • >50TB • NA	
Benchmark instances hospitality	Instances generated for specific optimization problems, inspired by real data or based on confidential data (see next)	Generate new data	Digital	Simulation data, aggregated data	.txt	<1GB	
Hospitality supply chain actor data	Anonymized data from hospitality supply chain actors about their operations (mostly confidential data)	Reuse existing data	Digital	Observational, simulation data	.txt, .xml	<1GB	
Hospitality city data	Anonymized data from cities (mostly confidential)	Reuse existing data	Digital	Observational, simulation data	.txt, .xml	<1GB	

Building data	Building and address register (GRAR) from Digitaal Vlaanderen	Reuse existing data	Digital	Observational	.shp, .dbf, .prj, .shx	<100GB	
Building functions - Leuven	Classification of main functional roles for individual buildings	Reuse existing data	Digital	Observational	.csv, .xlsx	<1GB	
Road interferences - Leuven	Applications (from organisations and individuals) for permissions to use/reserve space from the public domain ("Inname Openbare Domain", IOD)	Reuse existing data	Digital	Compiled/aggregated	.csv, .xlsx	<100GB	

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:

The existing data we will reuse comes from the databases and management systems of the companies and cities that want to collaborate in the project. This data is typically confidential and cannot be shared. Therefore, we also generate benchmark instances inspired by this data.
For the Flemish government data, see Digitaal Vlaanderen (https://www.vlaanderen.be/digitaal-vlaanderen/onze-oplossingen/gebouwen-en-adressenregister).

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- Yes

To estimate and model the mobility impact of construction and its related waste, as well as the delivery of goods by heavy vehicles (or cargo lifts), this project uses data from the "*Inname Openbare Domein*" (IOD), where individuals and companies have to apply for permission to reserve a portion of the public space (roads and sidewalks) for some specified period.
The data shared with us in this project, by the City of Leuven, is in the form of the applications for these reservations. Each application record includes the name, surname, email and telephone number of the applicant. The application also includes the detailed location of the proposed space reserved. Since residents are obliged to apply for such space reservations in the case of moving/deliveries of large appliances or furniture, the data can potentially link specific individuals with their place of residence. In this project, we are only interested in the type of space reservation, its location, its duration, and whether the application was approved or not. Still, the raw data we initially receive may include personal information. Should that be the case, the personal information will be permanently removed before processing the data. Only the cleaned data, without personal information, will be stored on GitLab and OneDrive, accessible only to the researchers involved.

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

The research codes, models and algorithms might have the potential for commercial
valorization. We work with open source licenses with the restriction that everyone who uses our data or tools/code must also make their tools/code or data available open source as well.
For commercial use or other use types that are excluded from the open-source
license, the license needs to be negotiated with KUL/LRD.
None of the data used for research purposes will be directly exposed in the development of the code or models.

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 in the comment section to what data they relate and what restrictions are in place.

- Yes

The data provided by the companies or cities will probably be confidential and cannot be shared with other researchers without permission from the data owners. Still, specific (benchmark) instances might be made available for other researchers to verify our work and to build further on or improve our work.
Within the STRAUSS project, the availability and use of the personal travel data derived from the activity-based schedules FEATHERS (UHasselt) is guided by the "samenwerkingsovereenkomst" signed between the research parties.

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 in the comment section to what data they relate and which restrictions will be asserted.

- No

2. 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).

Re-used data will either refer to the original source or a relevant copy will be uploaded to the KU Leuven data management tool, which has required meta-data documentation fields. Relevant code (including README files) will be stored on one-drive and made available via websites (such as https://www.mech.kuleuven.be/en/cib/op) or in the access-restricted GitLab (https://gitlab.kuleuven.be) and will be documented according to documentation standards and the available time of the researcher to improve re-usability.

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- Yes

Whenever relevant and appropriate, RDR, the data management tool of KU Leuven, will be used. It has such a standard. For code, standard practice concerning readability and documentation will be used.

3. Data storage & back-up during the research project

Where will the data be stored?

The data will be stored on the research computer of the researcher, with a back up at the internal storage drives of the research group. Cloud services, such as KU Leuven OneDrive will be used as well. GitLab is used for code and the associated data (for testing and repeatability purposes).

How will the data be backed up?

The internal drive of the research group provided by the department is used as a backup, this drive has an auto backup. For code, GitLab works with version control. So, if necessary, a rollback to a previous version of the code and data states can be made as well.

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

KUL OneDrive and Gitlab quota are sufficient for the entire project.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Security checks are in line with KU Leuven regulations (e.g. dual factor authentication for accessing all platforms where data is shared, such as GitLab, MSTeams, OneDrive,...).
Some platforms, such as GitLab, furthermore use leveled access. This makes it impossible for non-authorised to access these files.

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

No additional cost are expected

4. 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...).

The benchmark instance data will be retained. KU Leuven research data management policy stipulates that all relevant data generated are retained for a period of minimally 10 years after the end of the project. This will be followed.
The GitLab repository of the group including all the developed codes stays intact 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. The possibility of publication of open-source tools as formal publications with permanent public identifier is currently being examined for other packages in the group and will become our standard for open-science publishing.

Where will these data be archived (stored and curated for the long-term)?

After the research is concluded, the data will be kept at RDR as well as the internal storage drive of the research group and the GitLab.

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

The stored data volume is assumed to be small enough to not incur additional costs.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in an Open Access repository

The benchmark instance data will be made publicly available. Either through GitLab, RDR, or online (example here: https://www.mech.kuleuven.be/en/cib/op).
All other data shared with this project will remain strictly confidential and inaccessible to unauthorised parties.

If access is restricted, please specify who will be able to access the data and under what conditions.

Data shared by third parties with the STRAUSS project will only be available and accessible to authorised researchers via authenticated channels.

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 in the comment section per dataset or data type where appropriate.

- Yes, Intellectual Property Rights

Data shared by the companies or cities cannot be made available publicly without their consent.

Where will the data be made available? If already known, please provide a repository per dataset or data type.

A GitHub/GitLab repository will be used, but is not made yet.
Via RDR, no details are available yet.
My research group has sites which publishes and stores research data.
Some examples:

- https://www.mech.kuleuven.be/en/cib/lp
- https://www.mech.kuleuven.be/en/cib/op

When will the data be made available?

Along with each publication, relevant code and data will be made public. Under the code of open science, all relevant data that can be shared will be shared, to increase verifiability and reproducibility of the methods and their conclusions.

Which data usage licenses are you going to provide? If none, please explain why.

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. It has 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 you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- Yes

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

No additional costs are expected.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

Pieter Vansteenwegen - Johan Joubert and their researchers

Who will manage data storage and backup during the research project?

Pieter Vansteenwegen - Johan Joubert and their researchers

Who will manage data preservation and sharing?

Pieter Vansteenwegen - Johan Joubert and their researchers

Who will update and implement this DMP?

Pieter Vansteenwegen - Johan Joubert

STRAUSS - sustainable collaborative urban supply chains
GDPR

GDPR

Have you registered personal data processing activities for this project?

- No

STRAUSS - sustainable collaborative urban supply chains
DPIA

DPIA

Have you performed a DPIA for the personal data processing activities for this project?

- Not applicable