

---

## ROSANA: RObust SemAntic NAVigation in orchards

*A Data Management Plan created using DMPonline.be*

**Creator:** Lobna Joualy

**Affiliation:** KU Leuven (KUL)

**Funder:** Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

**Template:** FWO DMP (Flemish Standard DMP)

**Grant number / URL:** 1SHHJ24N

**ID:** 206036

**Start date:** 01-11-2023

**End date:** 31-10-2027

### **Project abstract:**

Many industrial sectors (e.g. automotive) have undergone strong automation in the past decades. However, within the agricultural sector and more specifically for the fruit industry, such automation is still lagging behind due to the cost and limitations of existing automation solutions. The decreasing availability of seasonal workers and increasing cost of manual labour threatens the profitability of the Belgian fruit sector. This PhD aims to support this sector by developing affordable yet robust mobile navigation techniques in the challenging outdoor setting of orchards. The overall goal of this project is to make several step changes towards a fully autonomous mobile robot capable of robustly navigating outdoors despite changing illumination, seasonal and weather conditions. The platform will be able to position itself and navigate both globally in the orchard and relatively to trees or branches. To achieve this, the project pursues three main objectives: (1) robust outdoor sensing and dataset construction; these sensors and datasets will be used in the next two objectives; (2) lifelong localisation and mapping with respect to trees and orchards using a semantic, adaptable map of the orchard; (3) outdoor semantic navigation by adopting semantic knowledge about recognised objects when planning robot trajectories, and by storing the effects of robot control parameters into the map.

**Last modified:** 11-04-2024

# ROSANA: RObust SemAntic NAVigation in orchards

## FWO 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
		<i>Please choose from the following options:</i> <ul style="list-style-type: none"> <li>• Generate new data</li> <li>• Reuse existing data</li> </ul>	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> <li>• Digital</li> <li>• Physical</li> </ul>	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> <li>• Observational</li> <li>• Experimental</li> <li>• Compiled/aggregated data</li> <li>• Simulation data</li> <li>• Software</li> <li>• Other</li> <li>• NA</li> </ul>	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> <li>• .por,</li> <li>• .xml,</li> <li>• .tab,</li> <li>• .csv,.pdf,</li> <li>• .txt, .rtf,</li> <li>• .dwg,</li> <li>• .gml, ...</li> <li>• NA</li> </ul>	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> <li>• &lt;100MB</li> <li>• &lt;1GB</li> <li>• &lt;100GB</li> <li>• &lt;1TB</li> <li>• &lt;5TB</li> <li>• &lt;10TB</li> <li>• &lt;50TB</li> <li>• &gt;50TB</li> <li>• NA</li> </ul>	
KB4RL	Knowledge Base that provides the mechanisms to store and retrieve information about actions, objects, the environment, the physics, their properties and relations.	Generate new data and Reuse existing data	Digital	Software	.ttl	<100MB	
Python scripts	The code is written in Python.	Generate new data	Digital	Software	.py	<100MB	
ROS2 package	The ROS2 package contains all the code relative to the simulation and Reinforcement Learning training.	Generate new data	Digital	Software	.py .xml .urdf .txt .yaml	<1GB	

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:

I plan to extend my knowledge base with knowledge graphs available on the web. However, I have not yet identified these external sources.

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.**

- 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

**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.**

- 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 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).**

- For the knowledge base I will provide a text document explaining the syntax used in it.
- The scripts will be commented to make the code easy to understand.
- The code, published on GitLab, will be accompanied by a README.md file that introduces the work and explains how to use the scripts.
- The ROS2 packages will also be published on GitLab and will be accompanied by a README.md file.

**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

For the Knowledge Base I will use the RDF format.

FAIRsharing.org: RDF/XML; RDF/XML Syntax Specification, DOI: 10.25504/FAIRsharing.5cae72, Last Edited: Thursday, May 4th 2023, 15:18, Last Accessed: Friday, March 29th 2024, 11:57

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

#### **Where will the data be stored?**

All the data will be stored on the shared OneDrive repository of research group ACRO, which is part of KU Leuven's OneDrive server. The algorithms and code will be stored and managed as repositories on the KU Leuven GitLab server.

#### **How will the data be backed up?**

- Everything on my laptop is synced with OneDrive, which automatically backs up my files.
- The scripts are uploaded to KU Leuven GitLab, which serves as a backup storage for the code. This also enables access to all previous versions of the code.

**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

The storage capacity required for the Knowledge Base and scripts is not very large, so the storage capacity is sufficient.

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

- OneDrive and GitLab are secured by KU Leuven and are not publicly accessible.
- My laptop is password-protected.

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

Everything used is provided by the University, and no additional costs are required for any of the storages.

### **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...).**

All data will be retained for a period of 5 years after the end of the project.

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

After the end of my PhD, the data and documents will be property of ACRO and therefore also of KU Leuven. Hence, the data will remain on secured servers. For ACRO, the repositories are managed by Prof. dr. Ir. Nikolaos Tsiogkas, Prof. dr. ir. Eric

Demeester and Prof. dr. ing. Karel Kellens, who are the responsible professors of ACRO.

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

No costs are required for data preservation.

## **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 scripts and knowledge base will be published and made publicly available.

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

/

**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.**

- No

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

The data will be made available via a Open Access GitLab repository.

**When will the data be made available?**

Once the scripts and the knowledge base are functional, they will be made available.

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

MIT License.

**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 costs are expected for data sharing.

## **6. Responsibilities**

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

Lobna Joualy

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

Lobna Joualy

**Who will manage data preservation and sharing?**

Lobna Joualy

**Who will update and implement this DMP?**

Lobna Joualy