DMP title

Project Name My plan (KU Leuven DMP)
Project Identifier u0130472
Grant Title 1S43222N
Principal Investigator / Researcher Suzanna Cuypers
Project Data Contact Suzanna Cuypers, 0479016773, suzanna.cuypers@kuleuven.be
Description Urban sprawl in Flanders has led to an increase in sealed surfaces, a more complex mobility, and longer infrastructure networks. The biggest financial consequence of urban sprawl is the loss of open space, which leads to a groundwater deficit. To remedy this, Flanders has set the goal to zero net land take by 2040. This project will help urban planners by exposing the effects on the landscape of current zoning laws at highly detailed parcel level. Concretely, machine learning networks will be developed to detect and segment the physical impact of zoning laws from high-resolution imagery. Machine learning networks that are capable of 3D object detection and image segmentation will allow for high-resolution mapping of physical impacts of zoning law changes and help interpret building functions. Change detection algorithms will process this output and compare time frames to identify changes and expose the evolution of land use. This involves both multiclass and simultaneous processing of time series. The most novel element of this project is the â€˜meta networkâ€™, which will deduce from a limited number of samples the objects that are indicative of and are the result of a zoning law change. This allows for the dynamic generation of output layers for a given zoning law, which illustrate the geographical impact of previous law changes and predict the geographical consequences of new law proposals.
Institution KU Leuven

1. Data Description
What data will you collect or create? Fill out the table below and/or describe.

Type of data	Format	Volume	How created
Lidar data	.e57	1Tb	ALS, SAR, UAV
DSM, DEM	.tiff	50 Gb	Lidar data
Processed mesh geometry	.obj	10Gb	Processed from lidar/photogrammetric inputs
WMS, WFS, GIS vector data	.shp	200 Gb	GeoPunt, CADGIS, publicly available
Orthomosaics	.tiff	200Gb	Orthomosaics processed from aerial and satellite imagery

Do you intend to reuse existing data?
Yes, I will primarilly be using publicly available data from GeoPunt, CADGIS, and other governmental data portals.

Do you use personal data (i.e. all data possibly identifying an individual)?

- No

The highest resolution I will be able to obtain will have a ground samping distance of 1m. This is not sufficient to identify individuals.

2. Documentation and Metadata
Describe the documentation that will be created for the data. This section deals with the way in which you will document how the dataset was created and subsequently processed.

1. An excel file will contain a sheet per georeferenced datatype: orthoimages (satellite and aerial), DSM, shapefiles, ... Each sheet contains the name of the folder where the imagery is saved, the municipalities the data covers, and whether it is downloaded and backup on the campus harddrives (KU Leuven protected storage).

2. Produced and adapted algoritms will be written on the campus harddrives and on a Linux pc. They will be backed up in the research group's Github and Gitlab. Some algorithms will be saved and run on Google drive.
3. The generated data will be saved on the local drive of the Linux pc, in Google drive, and, when it concerns a final result, on the campus harddrive. The head of the lab is in charge of data security.

Describe the metadata for the data. This section deals with metadata: information contained in your dataset about the research data.

As described above, and Excel file will contain the source infomation of gathered data. Output data from algorithms will be stored per run in output.txt files. We will also be using tools such as TensorBoard to monitor and compare algorithms systematically. A powerpoint file is used to record progress as well as where data can be found. We will include results with specific visual examples in the file.

3. Ethical, Legal and Privacy Issues
Are there any ethical issues concerning the creation and/or use of the data?

no

Did you consider all issues about copyrights and IPR?

yes

Are the collected data considered to be â€œdata containing personal informationâ€ and are all the requirements about the collection of these data met?

no
yes

4. Data storage and Backup during Research
How and where will the data be stored during research?

- Centrally on storage facilities of the research unit

Which back-up procedures are in place?

1. The time-stamped master copy of the data will be kept on our research unit central storage facility which is raid protected. Copies can be made and kept on personal devices.
2. Onedrive and Google drive are used to locally backup research and management files
3. The research group's Github and Gitlab will be used to backup algorithms, functions, definitions, and other code contents.

Describe the data security procedures and who has access to the data.

The research group's raid protected storage sits behind KUL firewalls. The code repositories are SSH key and password protected. The head of the lab is in charge of data security.

Only members of the research group have access to the data.

5. Data selection and Preservation after Research
What is the long-term preservation plan for these dataset(s)?

The data will be stored on the university's central servers (with automatic back-up procedures) for at least 10 years, conform the KU Leuven RDM policy.

Data Selection: Which data will have long time value for the research and will be preserved?

The modified algorithms and developed tools have long term value and will be preserved on the Git repositories.

6. Data Sharing
Are there any restrictions for sharing the data?

1. The gathered remote sensing data and inventories are publicly available and have no restrictions on sharing.
2. In case new data is gathered, this will not be published if it contains personal information.
3. The data will be used for research purposes only.

If there are no restrictions, which mechanisms will be in place to assure that the data

If there are no restrictions, which mechanisms will be in place to assure that the data are discoverable, accessible and intelligible?

At the end of the project, a visualization tool will be created to process the input data and generate outputs on the fly.

How will you share the data?

- Repository
- Website
- Publication

The produced outputs, such as statistics, will be shared in publications. However, in this project, the methodology is of higher relevance since the visualization tool can run the algorithm and produce the desired results on the fly. It will be made publicly available after the stakeholders have approved it.

With whom will the data be shared?

- Within the research unit only

The generated output maps will first be shared with the research group only. Some will be used as figures in publications or on social media, but they will be made anonymous by not sharing the location. The generated algorithms will be made public on Github.

7. Responsabilities and Resources

Who is responsible for Data Management during the project? This will be the person who might receive questions on the data management aspects of the research project.

The head of the research group: Maarten Vergauwen

Which additional resources are needed for the execution of the Data Management Plan?

Did you read the KU Leuven Data Management Policy? (find the link to the policy in the guidance).

- Yes