
Plan Overview

A Data Management Plan created using DMPOnline

Title: DroneML

Creator: Jitte Waagen

Principal Investigator: Jitte Waagen

Data Manager: Jitte Waagen

Project Administrator: Jitte Waagen

Affiliation: University of Amsterdam (Universiteit van Amsterdam)

Funder: Netherlands Organisation for Scientific Research (NWO)

Template: Data Management Plan NWO (September 2020)

ORCID iD: 0000-0001-8884-3704

Project abstract:

Archaeological remote sensing is of paramount importance to face challenges of climate change, agricultural intensification, and landscape development that are threatening buried archaeological remains. Drones using state-of-the-art sensor techniques strongly impact the possibilities of archaeologists by enabling them to trace ground and subsurface human activity in the past. This innovative technology however also results in huge high-resolution and multimodal datasets. Therefore, there is an urgent need for computer-aided inspection of these complex datasets. DroneML aims to develop software that can rapidly screen multiple feature types and multiple input layers simultaneously, to enable rapid processing of large datasets for subsequent manual assessment of identified features. The technology that would be the result of DroneML will hugely facilitate the work of archaeologists, as well as widen research possibilities, both within the field of heritage as well as any other discipline making use of remote sensing.

ID: 125768

Start date: 01-01-2024

End date: 01-07-2025

Last modified: 07-06-2023

Grant number / URL: <https://www.esciencecenter.nl/calls-for-proposals/open-escience-call-2023-oec-2023/>

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any

relationship to, your project or proposal

DroneML

General Information

Name applicant and project number

Jitte Waagen

Name of data management support staff consulted during the preparation of this plan and date of consultation.

Hanna Fricke (Data Steward, Faculty of Humanities, UvA)

1. What data will be collected or produced, and what existing data will be re-used?

1.1 Will you re-use existing data for this research?

If yes: explain which existing data you will re-use and under which terms of use.

- Yes

The project will re-use existing datasets from the 4D Research Lab (Drone). All of these datasets are open access (CC-BY) and there are no constraints on their use/re-use.

1.2 If new data will be produced: describe the data you expect your research will generate and the format and volumes to be collected or produced.

The data this project is going to produce is research software and new datasets of identified archaeological anomalies. This will be both code made available through the eScience Center Directory (<https://www.esciencecenter.nl/research-software-directory/>) as well as raster/vector data representing archaeological anomalies, shp/tif.

1.3. How much data storage will your project require in total?

- 10 – 100 GB

Produced raster and vector files will vary between a few kb's and potentially larger files, such as 500-1000 mb.

2. What metadata and documentation will accompany the data?

2.1 Indicate what documentation will accompany the data.

The 4D Research Lab will produce a report on the production of the DroneML Research Software, which will be made available

through the 4DRL Report Series (open access paper series in PDF format).

The coding for the DroneML research software app will be documented using the protocols and standards of the eScience Center.

The archaeological anomalies stored as vector or raster files will be added to the specific project repositories, and provided with the metadata description which is standard for the 4D Research Lab Report Series.

All this guarantees that the data, software and methods are published FAIR, stored long-term and are available for third-parties, without any restrictions.

2.2 Indicate which metadata will be provided to help others identify and discover the data.

The 4D Research Lab Report Series has a high standard of metadata. The report series is open access available in Figshare, provided with a DOI and licensed as CC-BY.

The code of the DroneML research software, and how to implement it, will be documented within the eScience Center Directory (<https://www.esciencecenter.nl/research-software-directory/>), thereby maintaining full open access to the metadata.

3. How will data and metadata be stored and backed up during the research?

3.1 Describe where the data and metadata will be stored and backed up during the project.

- Institution networked research storage

MS Teams/Sharepoint - General group storage

3.2 How will data security and protection of sensitive data be taken care of during the research?

- Not applicable (no sensitive data)

4. How will you handle issues regarding the processing of personal information and intellectual property rights and ownership?

4.1 Will you process and/or store personal data during your project?

If yes, how will compliance with legislation and (institutional) regulation on personal data be ensured?

- No

4.2 How will ownership of the data and intellectual property rights to the data be managed?

The eScience Center will be the owner of the developed research software. The 4D Research Lab will be the owner of the processed datasets and outcomes.

5. How and when will data be shared and preserved for the long term?

5.1 How will data be selected for long-term preservation?

- All data resulting from the project will be preserved for at least 10 years

5.2 Are there any (legal, IP, privacy related, security related) reasons to restrict access to the data once made publicly available, to limit which data will be made publicly available, or to not make part of the data publicly available?

If yes, please explain.

- No

5.3 What data will be made available for re-use?

- All data resulting from the project will be made available

5.4 When will the data be available for re-use, and for how long will the data be available?

- Data available upon completion of the project

5.5 In which repository will the data be archived and made available for re-use, and under which license?

Software:

The research software developed as part of DroneML will be made available by the eScience Center in its own research software directory: <https://www.esciencecenter.nl/research-software-directory/>. The license will be coordinated with the eScience Center, but an open access format is preferred, such as CC-BY or Apache 2.0. This ensures that the software will be attributed a DOI as well as a license, and easily findable and available for further (re-)use.

Data:

The data produced by DroneML will be made available through the existing digital infrastructures of the 4DRL and TPW, both guaranteed by the Amsterdam Institute of Humanities Research, UvA. The individual raster or vector datasets will also be persistently deposited individually as single files in Figshare, with a license (CC-BY) and DOI.

5.6 Describe your strategy for publishing the analysis software that will be generated in this project.

The DroneML research software is intended as a Quantum GIS plugin, and will be made available through the common QGIS Official Plugin Repositories.

6. Data management costs

6.1 What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?

This will fall under the scheduled work packages 1-7 of the project plan, and is therefore covered by the eScience Center time dedication to the project, as well as the lead applicants (0.1 FTE) time dedication.