
Plan Overview

A Data Management Plan created using DMPonline

Title: Harnessing root-microbe-mineral interactions for climate-friendly soils

Creator: Franciska de Vries

Principal Investigator: Franciska de Vries

Affiliation: University of Amsterdam (Universiteit van Amsterdam)

Funder: Netherlands Organisation for Scientific Research (NWO)

Template: Data Management Plan NWO (September 2020)

ORCID iD: 0000-0002-6822-8883

Project abstract:

Soils hold enormous amounts of carbon in the form of soil organic matter – only a small loss of this carbon can exacerbate climate change through increasing atmospheric CO₂ levels. Recent insights show that interactions between plant roots and soil microbes are the key to the formation of stable soil organic matter, but we don't know what the role is of different root systems. Here, using field experiments and greenhouse experiments, we will test which plants are effectively collaborating with microbes, and how we can use these plants to hold more carbon in our grasslands, under a changing climate.

ID: 97131

Start date: 01-05-2022

End date: 30-04-2027

Last modified: 13-04-2022

Grant number / URL: NWA.1527.20.001

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

Harnessing root-microbe-mineral interactions for climate-friendly soils

General Information

Name applicant and project number

Franciska de Vries, NWA.1527.20.001

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

Johannes de Groeve
Mariëtte van Selm
Karin Beijering

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

We will use the following existing data (WP1):

1. Non-publicly available data collected by our cooperation partner Eurofins. These data consist of soil properties, locations, and a timestamp. These data can only be published as aggregate statistics.
2. Farm level information from Statistics Netherlands. These data can only be published as aggregate statistics.
3. Hydrometeorological data. Temperature and rainfall data from KNMI, soil moisture from ERA5 and possibly private sources. These data will only be published as aggregate statistics.

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.

Data produced in T2.3, 2.4, 2.5 (WP2) and 3.1, 3.2, 3.3 (WP3) will consist of environmental data (soil properties, ecosystem fluxes and climatic variables), plant community data (species composition, biomass, functional traits) and soil community data (raw DNA sequences, OTU and taxonomy tables). All environmental, plant and soil community data will be stored in csv and txt format, while R code will

be stored in R format and QIIME2 pipeline in sh format.

We expect that the size of the data for each of the 6 tasks described above will be approximately 10GB, mostly because of the size of the fastq files containing the raw DNA sequences (each sample around 100MB in size).

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

- 10 – 100 GB

2. What metadata and documentation will accompany the data?

2.1 Indicate what documentation will accompany the data.

Each dataset will be accompanied by a README document explaining each file (methodology, set up, and variable names and units), by a file containing scripts and code of the statistical analyses used for the publication, and with a link to the associated publication where all methods for generating the datasets will be described in detail.

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

- All environmental, plant and soil community datasets, and R and QIIME2 code for reproducible analysis will be made openly available upon publication via the data repository FigShare (<https://uvaauas.figshare.com/>). These outputs (datasets/codes) will be described in detail via a README, including the methodology, set-up, variable names and description, units, geographic location and monitoring period of the study. To ensure proper metadata annotation and documentation we will follow established best-practices guidelines of IBED.
- Datasets and code associated to a publication may further be organized as a Figshare data collection. Such a collection will also include the link to the associated publication as well as links (e.g. doi) to any other associated resource used (e.g. NCBI raw DNA sequences).
- All raw DNA sequences from the project, will be deposited at NCBI. All necessary metadata following NCBI community standards will be provided. After publication NCBI datasets will have a unique identifier and link for accessing them – both (link/identifier) will be provided upon publication and will be added to the README file from the associated dataset on FigShare.

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

During the project, all data and metadata will be stored on University of Amsterdam cloud-based storage (Research Drive).

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?

Data for each WP will be co-owned by the PI, the WP-lead, and the postdoc (WP1 and WP2) or PhD (WP3).

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

Data storage in the FigShare repository is fully funded by the University of Amsterdam. The datasets will be kept on FigShare for 10 years and after this period will be moved to a long-term data repository from the University of Amsterdam, remaining openly accessible through the same DOI.

The raw DNA sequences from soil microbial communities will be kept at the NCBI repository permanently without any costs.

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 as soon as article is published

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

All data storage will be kept on FigShare for 10 years, and after this period will be moved to a long-term data repository from the University of Amsterdam, remaining openly accessible through the same DOI, under the CC-BY licence.

The raw DNA sequences from soil microbial communities will be kept at the NCBI repository permanently without any costs.

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

No analysis software will be generated.

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)?

Costs for data storage are fully met by the University of Amsterdam.