FWO DMP Template - Flemish Standard Data Management Plan

Version KU Leuven

Project supervisors (from application round 2018 onwards) and fellows (from application round 2020 onwards) will, upon being awarded their project or fellowship, be invited to develop their answers to the data management related questions into a DMP. The FWO expects a **completed DMP no later than 6 months after the official start date** of the project or fellowship. The DMP should not be submitted to FWO but to the research co-ordination office of the host institute; FWO may request the DMP in a random check.

At the end of the project, the **final version of the DMP** has to be added to the final report of the project; this should be submitted to FWO by the supervisor-spokesperson through FWO's e-portal. This DMP may of course have been updated since its first version. The DMP is an element in the final evaluation of the project by the relevant expert panel. Both the DMP submitted within the first 6 months after the start date and the final DMP may use this template.

The DMP template used by the Research Foundation Flanders (FWO) corresponds with the Flemish Standard Data Management Plan. This Flemish Standard DMP was developed by the Flemish Research Data Network (FRDN) Task Force DMP which comprises representatives of all Flemish funders and research institutions. This is a standardized DMP template based on the previous FWO template that contains the core requirements for data management planning. To increase understanding and facilitate completion of the DMP, a standardized **glossary** of definitions and abbreviations is available via the following link.

1. General Project Information	
Name Grant Holder & ORCID	Charlotte Janssens (https://orcid.org/0000-0002-0252-9843)
Contributor name(s) (+ ORCID) & roles	Miet Maertens (https://orcid.org/0000-0001-7245-0375), promotor
Project number [1] & title	ZKE4414, Modelling of agricultural markets and market impacts across spatial and temporal scales: Application to climate change scenario studies
Funder(s) GrantID [2]	12ATF24N
Affiliation(s)	☒ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: ROR identifier KU Leuven: 05f950310
Please provide a short project description	The project has three objectives that each relate to the impact of intra- and international agricultural markets on climate change mitigation and adaptation within the Agriculture, Forestry and Land use sector in Africa. The first one is to develop a spatial model of agricultural supply chains and markets. The second one is to understand the macro-economic conditions governing trade responses to climate extremes. The third one is to study the micro-level distributive impacts of market-based adaptation to climatic variability. The project will lead to technical innovations in spatial, micro-simulation, and dynamic general equilibrium modelling and to new, policy relevant knowledge on the impact of agricultural markets on climate change mitigation and adaptation.

[1] "Project number" refers to the institutional project number. This question is optional. Applicants can only provide one project number.

[2] Funder(s) GrantID refers to the number of the DMP at the funder(s), here one can specify multiple GrantIDs if multiple funding sources were used.

2. 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 [3].

				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
D1	Dataset WP1: spatial explicit data on agri-food production, consumption, and trade in 15 ECOWAS countries; spatial quantitative model; GLOBIOM model	Reuse existing data	Digital	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☒ Model ☐ Software ☐ Other:	Numerical data: .csv, .qs Model: .gms, .R, .py	☐ < 1 GB ☐ < 100 GB ☒ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	NA
D2_a	Dataset WP2 (adaptation): global country-level data on trade, stocks, crop yields, food aid, policies etc. of 4 main international crops (wheat, soya, maize, rice)	Reuse existing data	Digital	☒ Numerical ☒ Model	Numerical data: .csv, .qs Model: .gms, .R, .py	< 100 GB	NA
D2_m	Dataset WP2 (mitigation; research visit PSAE & ENS): global country-level data on production, consumption, trade,	Reuse existing data	Digital	☒ Numerical ☒ Model	Numerical data: .csv, .qs Model: .gms, .R, .py	< 1 TB	NA

	GHG emissions, and land use of agri-food products						
D3	Dataset WP3: household-level data on food consumption, assets, and agricultural market participation from the Tanzania National Panel Survey (TZNPS) and grid-level data on SPEI index	Reuse existing data	Digital	☒ Numerical ☒ Model	Numerical data: .csv, .qs Model: .gms, .R, .py	< 100 GB	NA

GUIDANCE:
The data description forms the basis of your entire DMP, so make sure it is detailed and complete. It includes digital and physical data and encompasses the whole spectrum ranging from raw data to processed and analysed data including analysis scripts and code. Physical data are all materials that need proper management because they are valuable, difficult to replace and/or ethical issues are associated. Materials that are not considered data in an RDM context include your own manuscripts, theses and presentations; documentation is an integral part of your datasets and should described under documentation/metadata.
RDM Guidance on data

[3] Add rows for each dataset you want to describe.

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.	D1: • FAOSTAT (https://www.fao.org/faostat/en/#data) • Karg et al. 2023 (https://www.nature.com/articles/s41597-023-02163-6) • IFPRI MapSPAM (https://dataverse.harvard.edu/dataverse/harvestchoice) • Global Roads Inventory Project (https://data.apps.fao.org/catalog/dataset/global-roads-inventory-project-grip-dataset) • GPW UN WPP-Adjusted Population Density v4.11 (https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-density-adjusted-to-2015-unwpp-country-totals-rev11) • ILO (https://ilostat.ilo.org/data/) • CEPII BACI (http://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele_item.asp?id=37) • West African Association for Cross-Border Trade in Agro-Forestry-pastoral and Fisheries Products (http://www.eco-icbt.org/) • FAO Food Price Monitoring and Analysis tool (https://fpma.fao.org/giews/fpmat4/#/dashboard/home) • FEWS food price data (https://fews.net/staple-food-price-data) • WFP food price data (https://data.humdata.org/dataset/global-wfp-food-prices?) • GLOBIOM model (https://globiom.org/index.html) D2_a: • FAOSTAT (https://www.fao.org/faostat/en/#data) • CEPII MAcMap-HS6 (http://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele_item.asp?id=12#:~:text=CEPII%20%2D%20MAcMap%2DHS6&text=MAcMap%2DHS6%20(Market%20Access%20Map,trade%20preferences%20exhaustively%20into%20account.) • CEPII BACI (http://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele_item.asp?id=37) • FAPDA (https://fapda.apps.fao.org/fapda/#main.html) • IMF Global Dept Database (https://www.imf.org/external/datamapper/datasets/GDD) • World Bank Development Indicators (https://databank.worldbank.org/source/world-development-indicators) • OECD Trade Facilitation Indicators (https://sim.oecd.org/default.ashx?ds=TFI)

	D2_m: • FAOSTAT (https://www.fao.org/faostat/en/#data): Emissions, Land use, Fertilizer use • FABIO input-output model v1.2 (https://github.com/fineprint-global/fabio) • Ecoinvent v3.10 (https://ecoinvent.org/database/) • EXIOBASE (https://zenodo.org/records/5589597) • Spatial map of nitrogen use (https://data.tpdc.ac.cn/en/data/1e31ad1f-5904-4b49-8796-b379c37e7dba) • OSCAR v3.3 (via personal contact with Thomas Gasser; dataset partially available at https://zenodo.org/records/10716491) D3: • Tanzania National Panel Survey Waves 1 – 5 (https://microdata.worldbank.org/index.php/catalog/5639) • Peng et al. 2019 (https://catalogue.ceda.ac.uk/uuid/bbdfd09a04304158b366777eba0d2aeb)
Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.	☐ Yes, human subject data; provide SMEC or EC approval number: ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☒ No Additional information:
Will you process personal data[4]? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).	☐ Yes (provide PRET G-number or EC S-number below) ☒ No Additional information:

[4] See Glossary Flemish Standard Data Management Plan

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 ☒ No If yes, please comment:
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 to what data they relate and what restrictions are in place.	☒ Yes ☐ No If yes, please explain: There will be a user agreement between KU Leuven and IIASA related to the use of the GLOBIOM model (part of database D1). There will be no restrictions on dissemination of the model output related to this research project. In relation to the exploitation of the model beyond the research tasks in this project, restrictions will be put in place in the user agreement.
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 to what data they relate and which restrictions will be asserted.	☐ Yes ☒ No If yes, please explain:

3. 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). *RDM guidance on documentation and metadata.*	The data and code of each work package will be organized in a separate folder that will be tracked with the **version control system Git** from the initial start of the project onwards. The following approach will ensure that the data is understandable and the research is reproducible (following guidelines): • Open-source datasets are directly downloaded from their respective online repositories. • Datasets with restricted access will be stored on KU Leuven OneDrive. • All scripts for data cleaning, processing, and analysis are numbered and organised with a run-all script (a single script that that executes all code from beginning to end). • Code is written following common style guidelines from the research community (e.g. https://style.tidyverse.org/ for R, https://www.tug.org/TUGboat/tb32-3/tb102verna.pdf for LaTeX, https://github.com/christophe-gouel/gams-style-guide for GAMS). • The version of R packages used in each project are saved with Renv (https://rstudio.github.io/renv/articles/renv.html). • A README.txt file will describe the data, the software versions, and the instructions for replication.
Will a metadata standard be used to make it easier to **find and reuse the data**? If so, please specify which metadata standard will be used. If not, please specify which metadata will be created to make the data easier to find and reuse. *REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS.*	☐ Yes ☒ No If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used: If no, please specify (where appropriate per dataset or data type) which metadata will be created: A metadata will be created for each dataset (D1, D2_a, D2_m, D3) based on the metadata model from KU Leuven RDR (https://www.kuleuven.be/rdm/en/rdr/full-metadata-model).

4. Data Storage & Back-up during the Research Project	
Where will the data be stored? *Consult the interactive KU Leuven storage guide to find the most suitable storage solution for your data.*	☐ Shared network drive (J-drive) ☐ Personal network drive (I-drive) ☒ OneDrive (KU Leuven) ☐ Sharepoint online ☐ Sharepoint on-premis ☐ Large Volume Storage ☐ Digital Vault ☒ Other: Github https://github.com/ Note: Datasets which are not elsewhere available on online repositories will be stored on KU Leuven Onedrive. All scripts for data cleaning, processing and analysis will be stored on GitHub. The open-source datasets that are re-used from existing studies or databases will be accessed directly from their online repositories and thus do not need to be stored separately.
How will the data be backed up? *WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS?*	☒ Standard back-up provided by KU Leuven ICTS for my storage solution ☐ Personal back-ups I make (specify) ☐ Other (specify)
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 (KU Leuven OneDrive offers 2 TB of data) ☐ No If no, please specify:

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons? *CLEARLY DESCRIBE THE MEASURES (IN TERMS OF PHYSICAL SECURITY, NETWORK SECURITY, AND SECURITY OF COMPUTER SYSTEMS AND FILES) THAT WILL BE TAKEN TO ENSURE THAT STORED AND TRANSFERRED DATA ARE SAFE.* *Guidance on security for research data*	During the research phase, the data and code are confidential and should only be shared with collaborators (inside and outside KU Leuven). The data and code will be securely stored on OneDrive (on personal OneDrive shared with specific collaborators) and on Github (on private repository shared with collaborators). When the research is finished, the data and code are made publicly available on the KU Leuven RDR repository.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	Data storage and backup provided by KU Leuven (OneDrive) are offered free of charge to KUL personnel. The only expected costs for data storage are 5 USD/month for the Git LFS feature, which will be covered with the FWO bench fee related to this project.

5. 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...). *Guidance on data preservation*	☒ All data will be preserved for 10 years according to KU Leuven RDM policy ☐ All data will be preserved for 25 years according to CTC recommendations for clinical trials with medicinal products for human use and for clinical experiments on humans ☐ Certain data cannot be kept for 10 years (explain)

Where will these data be archived (stored and curated for the long-term)? *Dedicated data repositories are often the best place to preserve your data. Data not suitable for preservation in a repository can be stored using a KU Leuven storage solution, consult the interactive KU Leuven storage guide.*	☒ KU Leuven RDR ☐ Large Volume Storage (longterm for large volumes) ☐ Shared network drive (J-drive) ☐ Other (specifiy):
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	The data repository provided by KU Leuven RDR are offered free of charge to KUL personnel. You can store up to 50GB per year and each individual file uploaded may not exceed 5GB.

6. Data Sharing and Reuse	
Will the data (or part of the data) be made available for reuse after/during the project? Please explain per dataset or data type which data will be made available. *NOTE THAT 'AVAILABLE' DOES NOT NECESSARILY MEAN THAT THE DATA SET BECOMES OPENLY AVAILABLE, CONDITIONS FOR ACCESS AND USE MAY APPLY. AVAILABILITY IN THIS QUESTION THUS ENTAILS BOTH OPEN & RESTRICTED ACCESS. FOR MORE INFORMATION: HTTPS://WIKI.SURFNET.NL/DISPLAY/STANDARDS/INFO-EU-REPO/#INFOEUREPO-ACCESSRIGHTS*	☐ Yes, as open data ☒ Yes, as embargoed data (temporary restriction) ☐ Yes, as restricted data (upon approval, or institutional access only) ☐ No (closed access) ☐ Other, please specify:
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 per dataset or data type where appropriate.	☐ Yes, privacy aspects ☐ Yes, intellectual property rights ☐ Yes, ethical aspects ☐ Yes, aspects of dual use ☐ Yes, other ☒ No If yes, please specify:
Where will the data be made available? If already known, please provide a repository per dataset or data type.	☒ KU Leuven RDR ☐ Other data repository (specify) ☐ Other (specify)
When will the data be made available?	☒ Upon publication of research results ☐ Specific date (specify) ☐ Other (specify)
Which data usage licenses are you going to provide? If none, please explain why. *A DATA USAGE LICENSE INDICATES WHETHER THE DATA CAN BE REUSED OR NOT AND UNDER WHAT CONDITIONS. IF NO LICENCE IS GRANTED, THE DATA ARE IN A GREY ZONE AND CANNOT BE LEGALLY REUSED. DO NOTE THAT YOU MAY ONLY RELEASE DATA UNDER A LICENCE CHOSEN BY YOURSELF IF IT DOES NOT ALREADY FALL UNDER ANOTHER LICENCE THAT MIGHT PROHIBIT THAT.* *Check the RDR guidance on licences for data and software sources code or consult the License selector tool to help you choose.*	☒ CC-BY 4.0 (data) ☐ Data Transfer Agreement (restricted data) ☐ MIT licence (code) ☒ GNU GPL-3.0 (code) ☐ Other (specify)

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here. *INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.*	☒ Yes, a PID will be added upon deposit in a data repository ☐ My dataset already has a PID ☐ No
What are the expected costs for data sharing? How will these costs be covered?	There are no expected costs for data sharing.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	Charlotte Janssens
Who will manage data storage and backup during the research project?	Charlotte Janssens
Who will manage data preservation and sharing?	Charlotte Janssens
Who will update and implement this DMP?	Charlotte Janssens