1. General Project Information	
Name Grant Holder & ORCID	Céline Van Migerode – 0000-0002-4023-7665
Contributor name(s) (+ ORCID) & roles	**Supervisor:** Ben Derudder – 0000-0001-6195-8544 **Co-supervisor:** Ate Poorthuis – 0000-0002-3808-7493
Project number [1] & title	3H220428 – Leveraging ambiguity in urban morphologies
Funder(s) GrantID [2]	11P4224N
Affiliation(s)	☑ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: ROR identifier KU Leuven: 05f950310

[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.

Please provide a short project description	Key concepts in urban studies such as cities, suburbanisation and sprawl often lack a commonly accepted definition and methodological framework guiding their operationalisation. This issue – described as 'ambiguity' – has major implications for both policy and research. It leads to arbitrariness in funding decisions and hampers comparative urban research, thus limiting our understanding of urban processes on a global scale. Rather than approaching ambiguity as a problem, I propose to leverage ambiguity as it can provide insight in the fundamental characteristics of inherently complex urban morphologies. This can only be done by making ambiguity measurable and tangible. To this end, the project uniquely and explicitly quantifies different sources of ambiguity in urban concepts by assessing the consistency in operationalisations on a global scale. I will use three specific case studies: (i) the definition of cities and towns provided by the Degree of Urbanisation project – one of the most widely used definitions in policy, (ii) the assessment of urban sprawl and (iii) the delineation of urban areas. In addition, I will illustrate with concrete examples how the resulting quantification of ambiguity can be leveraged in future policy and research. By making ambiguity in urban concepts tangible, the project facilitates a more balanced knowledge production and enhances interpretation and understanding of 'the urban' in a global context, with room for regional and local nuance.

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.

				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
Global Human Settlement Layer	Geographical grid with (1) population counts, (2) proportion of built-up area and (3) proportion of land surface.	☐ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	GeoTiFF	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
WorldPop	Geographical grid with population counts	☐ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	GeoTiFF	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
VIIRS night-time lights from Earth Observation Group	Geographical grid with night-time light emission	☐ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual	GeoTiFF	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB	

				☐ Model ☐ Software ☐ Other:		☐ NA	
Flexurba package	R package constructed in WP1.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☒ Software ☐ Other:	Text files with R code	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Set of urban delineations generated in WP1.	Urban delineations constructed by varying criteria in Degree of Urbanisation algorithm.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	GeoTiFF	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Code repository for WP1.	R code to construct and analyse delineations in WP1.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	Text files with R code	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Set of urban delineations	Urban delineations constructed by	☒ Generate new data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images	GeoTiFF and GeoJSON	☐ < 1 GB ☒ < 100 GB	

generated in WP3.	varying urban proxy and data source.	☐ Reuse existing data		☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:		☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Code repository for WP3.	R code to construct and delineations in WP3.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	Text files with R code	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Set of sprawl delineations generated in WP2.	Sprawl delineations constructed by different indicators.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	GeoTiFF and GeoJSON	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Code repository for WP2.	R code to construct and delineations in WP2.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software	Text files with R code	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	

				☐ Other:			

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

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.	Global Human Settlement Layer: https://human-settlement.emergency.copernicus.eu/download.php WorldPop: https://www.worldpop.org/ VIIRS night-time lights of Earth Observation Group: https://eogdata.mines.edu/products/vnl/#v1
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[3]? 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:

[3] 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:
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*.	• Analysis and/or calculations on existing data sets will be conducted in the R programming language. All code will be provided by extensive documentation. I will employ computational notebooks, such as Quarto for this purpose. These notebooks allow to combine both text, code and figures in a clear and organised manner. The notebooks with documentation of the workflow will be openly made available in a code repository according to the FAIR principle (Findable, Accessible, Interoperable and Reusable) to ensure reproducibility. • When creating new data (such as the urban delineations), there will always be an accompanying JSON-file generated with the metadata (see below).
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: • Metadata about the urban/sprawl delineations generated in WP1, WP2 and WP3 will be stored in JSON files. The JSON-files will contain the source data and parameters/configurations used to generate the delineation and the timestamp of creation.

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: • All code will be stored using Git version control (through KU Leuven GitLab service). • Data and metadata will be stored on the active data management platform (ManGO) and on a server environment provided by the research group of the co-supervisor (Ate Poorthuis)
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): • Git version control will be used to store the code repositories. The Git service includes online back-ups. • The data and metadata will be stored both on OneDrive, on a server environment provided by the research group of the co-supervisor (Ate Poorthuis) and on KU Leuven's ManGO facility. This three-double storage ensures back-up of the data and metadata.
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 ☐ No • The expected size of all research datasets is between 100GB and 200GB. OneDrive for Business Cloud storage allows for a maximum size of 2 TB. KU Leuven's ManGO facility allows for data storage up to 1 TB.

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*	• Both OneDrive and ManGO is only accessible through the KU Leuven multifactor authentication procedure. Editing and viewing rights in OneDrive can be set within KU Leuven. As a default, the datasets will be made inaccessible to those not involved with the research project. ManGO projects are only accessible to project members by definition. Both data storage solutions are thus highly secured.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	• There are no expected costs. Both the OneDrive and ManGO facility are free for the expected size of the research 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?	• There are no expected costs as the KU Leuven RDR is free to use.

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: • The code repository for WP1, WP2 and WP3 will be shared in a GitLab repository upon publication of the corresponding research articles. • In addition, both the code repositories and the urban/sprawl delineation for WP1, WP2 and WP3 will be published on the KU Leuven RDR.
If access is restricted, please specify who will be able to access the data and under what conditions.	N/A

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?	No costs are expected.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	The PhD researcher (Céline Van Migerode) will manage data documentation and metadata during the research, supported by the supervisor (Ben Derudder) and co-supervisor (Ate Poorthuis).
Who will manage data storage and backup during the research project?	The PhD researcher (Céline Van Migerode) will manage data storage and backup during the research, supported by the supervisor (Ben Derudder) and co-supervisor (Ate Poorthuis).
Who will manage data preservation and sharing?	The PhD researcher (Céline Van Migerode) will manage data preservation and sharing during the research, supported by the supervisor (Ben Derudder) and co-supervisor (Ate Poorthuis).
Who will update and implement this DMP?	The PhD researcher (Céline Van Migerode) will update and implement this DMP.