1. General Project Information	
Name Grant Holder & ORCID	**Promoter: Prof. Tara Grauwet (ORCID 0000-0002-5156-4792)**
Contributor name(s) (+ ORCID) & roles	**Co-promoter: Prof. Ann Van Loey (ORCID 0000-0002-9164-7760)**
Project number[1] & title	C14/22/101 INGENIOUS HETEROGENEOUS BEANS: Unravelling variation during hydrothermal processing and digestion
Funder(s) GrantID[2]	D-2022-1707
Affiliation(s)	☒ KU Leuven
Please provide a short project description	Based on literature and own research, a large heterogeneity in processing and digestive behaviour of common beans has been described including differences between different varieties, between different seeds from a given batch, and even at cellular level. From a processing technology as well as food nutrition point of view, it is of primary interest to get more fundamental insight into the reasons for these heterogeneities at the level of processing and digestion. For this purpose, we will develop an innovative approach based on selecting common bean varieties (fresh and aged) and sorting individual seeds and cells into more homogenous subfractions to study in detail (i) the relation between the cooking behaviour and composition and structure, and (ii) the relation between the digestion properties and composition and structure as affected by the cooking process. This will be done by combining ex situ and in situ microscopic approaches complemented with (bio)chemical characterisations.

[1] "Project number" refers to the institutional project number. This question is optional since not every institution has an internal project number different from the GrantID. 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
Existing data	Published data for use in literature review	☒ Reuse existing data	☐ Digital	☒ Compiled/ aggregated data	☒ .pdf	☒ < 100 MB	
Activity 0	Sourcing of bean varieties	☒ Generate new data	☒ Physical				Storage at -40 °C (2 large freezer units)
Activity 1.1: Bean characterization	- Compositional characterization (spectrophotometric and chromatography methods)	☒ Generate new data	☒ Digital	☒ Experimental	☒ .xlsx ☒ .pdf	☒ < 100 MB	
	- Microscopic characterization	☒ Generate new data	☒ Digital	☒ Experimental ☒ Software	☒ .jpeg ☒ .tiff ☒ .xslx	☒ < 1 GB	
	- Cooking behaviour	☒ Generate new data	☒ Digital	☒ Experimental ☒ Software	☒ .xlsx ☒ .jmp ☒ .sas	☒ < 100 MB	
Activity 1.2: Bean aging	- Generation of aged beans	☒ Generate new data	☒ Physical				Storage at -40 °C (1 large freezer unit)
	- Cooking behaviour	☒ Generate new data	☒ Digital	☒ Experimental ☒ Software	☒ .xlsx ☒ .jmp ☒ .sas	☒ < 100 MB	

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

Activity 2.1: Seed sorting	- Bean classification according to cooking behaviour	☒ Generate new data	☒ Digital ☒ Physical	☒ Experimental	☒ .xlsx	☒ < 100 MB	Storage at -40 °C (<1 freezer unit)
	- Generation of intact cotyledon cells	☒ Generate new data	☒ Digital ☒ Physical	☒ Experimental	☒ .xlsx ☒ .jpeg ☒ .tiff	☒ < 100 MB	Storage at -40 °C (<1 freezer unit)
Activity 2.2: Cell sorting and *in vitro* digestion	- Sorting of intact and digested cotyledon cells	☒ Generate new data	☒ Digital ☒ Physical	☒ Experimental	☒ .xlsx	☒ < 100 MB	Storage at -40 °C (<1 freezer unit)
	- *In vitro* digestion of intact cotyledon cells	☒ Generate new data	☒ Digital ☒ Physical	☒ Experimental ☒ Software	☒ .xlsx ☒ .jmp ☒ .sas	☒ < 100 MB	Storage at -40 °C (<1 freezer unit)
Activity 3.1: Microscopic *ex situ* evaluation	- Microscopic characterization	☒ Generate new data	☒ Digital	☒ Experimental ☒ Software	☒ .jpeg ☒ .tiff ☒ .xslx	☒ < 100 GB	
Activity 3.2: Microscopic *in situ* evaluation during cooking	- Generation of intact cotyledon cells	☒ Generate new data	☒ Digital ☒ Physical	☒ Experimental	☒ .xlsx ☒ .jpeg ☒ .tiff	☒ < 100 MB	Storage at -40 °C (<1 freezer unit)
	- Microscopic characterization	☒ Generate new data	☒ Digital	☒ Experimental ☒ Software	☒ .jpeg ☒ .tiff ☒ .xslx	☒ < 100 GB	
Activity 3.3: Microscopic *in situ* evaluation during digestion	- Microscopic characterization	☒ Generate new data	☒ Digital	☒ Experimental ☒ Software	☒ .jpeg ☒ .tiff ☒ .xslx	☒ < 100 GB	
Activity 4.1: Biochemical characterization of digestion	- Chemical characterization (spectrophotometric and chromatography methods)	☒ Generate new data	☒ Digital	☒ Experimental ☒ Software	☒ .xlsx ☒ .pdf ☒ .jmp ☒ .sas	☒ < 100 MB	

Activity 4.2: Biochemical characterization of cell wall biopolymers	- Chemical characterization (spectrophotometric, spectroscopic and chromatography methods)	☒ Generate new data	☒ Digital	☒ Experimental	☒ .xlsx ☒ .pdf	☒ < 100 MB	

GUIDANCE:

DATA CAN BE DIGITAL OR PHYSICAL (FOR EXAMPLE BIOBANK, BIOLOGICAL SAMPLES, …). DATA TYPE: DATA ARE OFTEN GROUPED BY TYPE (OBSERVATIONAL, EXPERIMENTAL ETC.), FORMAT AND/OR COLLECTION/GENERATION METHOD.

EXAMPLES OF DATA TYPES: OBSERVATIONAL (E.G. SURVEY RESULTS, SENSOR READINGS, SENSORY OBSERVATIONS); EXPERIMENTAL (E.G. MICROSCOPY, SPECTROSCOPY, CHROMATOGRAMS, GENE SEQUENCES); COMPILED/AGGREGATED DATA[4] (E.G. TEXT & DATA MINING, DERIVED VARIABLES, 3D MODELLING); SIMULATION DATA (E.G. CLIMATE MODELS); SOFTWARE, ETC.

EXAMPLES OF DATA FORMATS: TABULAR DATA (.POR,. SPSS, STRUCTURED TEXT OR MARK-UP FILE XML, .TAB, .CSV), TEXTUAL DATA (.RTF, .XML, .TXT), GEOSPATIAL DATA (.DWG,. GML, ..), IMAGE DATA, AUDIO DATA, VIDEO DATA, DOCUMENTATION & COMPUTATIONAL SCRIPT.

DIGITAL DATA VOLUME: PLEASE ESTIMATE THE UPPER LIMIT OF THE VOLUME OF THE DATA PER DATASET OR DATA TYPE.

PHYSICAL VOLUME: PLEASE ESTIMATE THE PHYSICAL VOLUME OF THE RESEARCH MATERIALS (FOR EXAMPLE THE NUMBER OF RELEVANT BIOLOGICAL SAMPLES THAT NEED TO BE STORED AND PRESERVED DURING THE PROJECT AND/OR AFTER).

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.	Literature information might be combined at the level of literature review papers. Of course, at that level, literature sources will be always cited by references as part of the text.
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, please describe these issues further and refer to specific datasets or data types when appropriate.	☐ Yes, human subject data ☐ Yes, animal data ☐ Yes, dual use ☒ No If yes, please describe:

[4] These data are generated by combining multiple existing datasets.

Will you process personal data[5]? If so, briefly describe the kind of personal data you will use. Please refer to specific datasets or data types when appropriate. If available, add the reference to your file in your host institution's privacy register.	☐ Yes ☒ No If yes: - Short description of the kind of personal data that will be used: - Privacy Registry Reference:
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:

[5] See Glossary Flemish Standard Data Management Plan

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).	***Physical data:*** A descriptive map indicating where each sample or material generated is stored will be maintained as a *.xlsx or *.docx file. ***Digital data:*** -Protocols/experimental and evaluation procedures will be clearly written and maintained in *.docx /*.pdf format. -Meaningful and descriptive data file names will be used: for raw data as well as analysed data. -Where initial raw data is exported into a new format: reference to initial data (location) will be made. -Steps involved in data analysis and relevant analysis will be documented in *.docx file. -For published material, a metadata will be available: with folders showing the published material, associated processed and raw data files (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 For each published article, a metadata will be created to ease data retrieval and reuse. This metadata will be based on the Dublin core metadata standard. The metadata will be named based on the title of article, authors, year of publication. Files associated with the metadata will include: -The accepted version of the article as available in KU Leuven, Lirias: *.pdf file. -The published article as available online in the peer reviewed journal: *.pdf file; wherein doi-number, journal of publication and abstract, etc can be found. -The peer review history of the article; comments from reviewers, rebuttal to the reviewers: *.pdf file/*.docx. -An excel sheet of the data used to generate the graphs and tables presented in the published article: *.xlsx. A folder (only for internal use) with all excel sheets (*.xlsx) related to the final graphs in the article: therein, the raw data (including the equipment used, location of the raw/equipment generated data) will be referenced. The metadata and the referenced files will be stored on the Archive/'K:' network drive

4. Data Storage & Back-up during the Research Project	
Where will the data be stored?	***Physical data:*** Raw materials and samples generated during experiments will be stored in freezers or desiccators until the associated data has been published or until the end of the project (end of the associated PhD). For samples that are difficult to regenerate, a longer storage period will be ensured. ***Digital data:*** Data will be stored on KU Leuven network drives: - Personal/'I:' network drive or OneDrive for Business storage (2 TB available to every PhD researcher – not for strictly confidential or personal data) for daily personal use and data storage. - Archive/'K:' network drive will be used for storage of published data. Data which was stored on the I: network drive and OneDrive storage will be transferred to K: network drive when personnel is leaving the research unit.
How will the data be backed up? *WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS? DESCRIBE THE LOCATIONS, STORAGE MEDIA AND PROCEDURES THAT WILL BE USED FOR STORING AND BACKING UP DIGITAL AND NON-DIGITAL DATA DURING RESEARCH.*[6] *REFER TO INSTITUTION-SPECIFIC POLICIES REGARDING BACKUP PROCEDURES WHEN APPROPRIATE.*	The network drives are automatically backed-up by IT, KU Leuven: https://icts.kuleuven.be/sc/english/storage/desktop-file-storage. OneDrive for Business data is stored in the cloud in which version history (up to 100 versions per file) is enabled: https://icts.kuleuven.be/sc/english/storage/onedrive.

[6] Source: Ghent University Generic DMP Evaluation Rubric: https://osf.io/2z5g3/

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 - 50 GB of Personal/'I:' data storage for every researcher is available in our research unit. In addition, 2 TB of OneDrive for Business storage space is available for free to every PhD researcher. - 1.3 TB of long-term data storage is available on the Archive/'K:' network drive folders of the research unit. If necessary, more storage space will be acquired from the university, with the cost being covered by the research unit/the project budget.
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.* [6]	- Data stored on the Personal/'I:' network drive is only accessible to a specific researcher. This data is automatically backed up by ICT, KU Leuven. Access to KU Leuven network drives is automatically secured by multifactor authentication. - Data stored on OneDrive for Business is only accessible to a specific researcher. Version history (up to 100 versions per file) is enabled. Multifactor authentication can be activated as additional security measure. OneDrive for Business will not be used for strictly confidential or personal data. - Final data files can also be stored on the Archive/'K:' network drive, which has restricted access (only professors and postdocs of the research group).
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	Storage capacity (on the Archive/'K:' network drive) can be increased for an annual fee of 100.86 Euro/TB. This data is automatically backed up by ICT, KU Leuven. Buying additional storage capacity can be covered by the consumables budget of the 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...).	- Only physical samples that are difficult to regenerate will be retained after the end of the project. - All digital data will be retained for more than the expected 5-year period, on the Archive/'K:' drive.
Where will these data be archived (stored and curated for the long-term)?	The data will be copied to the Archive/'K:' network drive (with automatic back-up procedures) and stored for at least 5 years. To date, 1.3 TB storage is available, however if needed, additional data space can be obtained.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	Annual fee of 100.86 Euro/TB. The research unit/project consumables budget will be used to cover this cost.

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*	Data in the form of figures and tables in peer reviewed publications will be made available: ☐ Yes, in an Open Access repository ☒ Yes, in a restricted access repository (after approval, 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.	- Only researchers and professors participating in the project will have access to the data prior to data publication. - Access of data will be through peer reviewed journals, conference presentations and proceedings, repositories. - Open access opportunities (major focus on green open access routes) will be considered to enhance visibility of our research.
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
Where will the data be made available? If already known, please provide a repository per dataset or data type.	- Data of completed work will be published in academic peer reviewed journals and will as such be available in existing and relevant repositories (e.g. the internal KU Leuven repository: Lirias). Open access opportunities (major focus on green open access routes) will be considered to enhance visibility of our research. - Unpublished data will be available on network drives with restricted access (as explained earlier).

When will the data be made available? *This could be a specific date (dd/mm/yyyy) or an indication such as 'upon publication of research results'.*	After publication of data in academic journals.
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.* *Example Answer: E.g. "Data from the project that can be shared will be made available under a Creative Commons Attribution license (CC-BY 4.0), so that users have to give credit to the original data creators."* [7]	Data usage licenses will be provided according to the requirements for reuse defined by the original publisher.
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 ☒ No

[7] Source: Ghent University Generic DMP Evaluation Rubric: https://osf.io/2z5g3/

What are the expected costs for data sharing? How will these costs be covered?	- Within the research unit: the Shared/'J:' network drive will be used and this is freely accessible to all researchers. - Between collaborating research units: the Shared/'J:' network drive (within KU Leuven) or OneDrive for Business (external partners) will be used. - Within KU Leuven: Lirias, a free document repository will be used. - External data sharing through publication in peer reviewed journals. - Cost of data sharing at conferences will be covered by the project budget. - Cost for possible golden open access will be covered by the project consumables budget.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	The PhD students involved in the project.
Who will manage data storage and backup during the research project?	The PhD students involved in the project.
Who will manage data preservation and sharing?	The promoters of the project: Prof. Tara Grauwet and Prof. Ann Van Loey.
Who will update and implement this DMP?	The promoters of the project: Prof. Tara Grauwet and Prof. Ann Van Loey.