�·B-GOOD

Data Management Plan

Deliverable D9.2

28 November 2019

Lina De Smet[1], Marten Schoonman[2], Wim Verbeke[1], Dirk C. de Graaf[1]

1 UGENT, 2 BEEP

B-GOOD

Giving Beekeeping Guidance by cOmputatiOnal-assisted Decision making

_____

Prepared under contract from the European Commission

Grant agreement No. 817622
EU Horizon 2020 Research and Innovation action

Project acronym:	**B-GOOD**
Project full title:	**Giving beekeeping guidance by computational-assisted decision making**
Start of the project:	June 2019
Duration:	48 months
Project coordinator:	Prof. Dirk de Graaf
	Ghent University
	www.b-good-project.eu

Deliverable title:	Data Management Plan
Deliverable n°:	D9.2
Nature of the deliverable:	Report
Dissemination level:	Public

WP responsible:	WP9
Lead beneficiary:	UGENT

Citation:	De Smet, L., Schoonman, M., Verbeke, W. & de Graaf, D.C. (2019). *Data Management Plan*. Deliverable D9.2 EU Horizon 2020 B-GOOD Project, Grant agreement No. 817622.

Due date of deliverable: Month n°6
Actual submission date: Month n°6

Deliverable status:

Version	Status	Date	Author(s)
0.1	Draft	18 November 2019	De Smet, Schoonman, Verbeke, de Graaf UGENT; BEEP
0.2	Review	28 November 2019	Williams AU
1.0	Final	28 November 2019	De Smet, Schoonman, Verbeke, de Graaf UGENT; BEEP

_____

Table of contents

_____

Preface

This deliverable describes the first version of the Data Management Plan (DMP) which will be updated during the project. It includes the General Data Protection Regulation (GPDR) section as personal data will be collected in the project.

1. Data summary

What is the purpose of the data collection/generation and its relation to the objectives of the project?

The overall aim of the B-GOOD project is to pave the way towards healthy and sustainable beekeeping within the EU. A key to sustainable beekeeping is interdisciplinary problem-solving research by integrating socio-economics and the human-ecosystem equilibrium of beekeeping with bee health. Coordinated and harmonized flow of data from various sources and by testing and validating each component thoroughly will make the Health Status Index (HSI) fully operational. Another key to a sustainable beekeeping sector is a better understanding of its socio-economics, particularly within local value chains, its relationship with bee health and the human-ecosystem equilibrium of the beekeeping sector and to implement these insights into the data processing and decision making.

In WP1 different health components related to the bee colony will be collected under experimental and field conditions in order to contribute to the operationalization of the Health Status Index (HSI). Some data will be collected in an automatic or semi-automatic way while other data will be collected by laboratory analysis. In WP2, new tools will be developed and tested if they can contribute to the HSI. In a three tier process the valuable health components will be determined in order to end up with a validated HSI and related computational decision making. Next to the biological data, WP4 and WP8 will focus on the identification of socio-economic components of healthy and sustainable beekeeping. This data together with the ecological-environmental factors from WP5 will also be implemented in the operationalization of the HSI and decision making.

One of the core requirements of the B-GOOD project is to develop an EU-wide bee health data web portal, data portal in short. This will enable storing and sharing of bee health **datasets**. It may include raw data as well as analyzed and visualized data from both within B-GOOD generated data as well as data from other sources. The web portal will be developed by the B-GOOD partner BEEP foundation. The setup and management of the data platform will be done professionally for use during the project with the aim of continued use for many years after the end of the B-GOOD research project.

Next to the data portal a website (b-good-projects.eu) will be set in place. This website will provide an internal communication platform (ICP) which will be used for storage and sharing of **project files** (presentation, reports, etc.). The website is hosted by B-GOOD partner PENSOFT.

What types and formats of data will the project generate/collect and what is the expected size of the data (if known)?

Through the multidisciplinary approach of the B-GOOD project, a wide range of types and formats of data will be collected and generated. The overview is presented in Table 1.

_____

Table 1: Overview of data which will be generated during the B-GOOD project in the different work packages.

Work packages	Overview of the data that will be generated
WP1	• **Hive measurements:** continuous 15-minute data of about 23 sensors per hive: weight, temperature, humidity, 20 audio bands. Database: Application programming interface (API) • **Beekeeper inspection data:** about 20x per year a set of classified data of the status of the hives/bee colonies. Database: API • **Pathogen analysis + genotyping.** File-type: csv, 500KB-1MBytes each • **Photos** from results (GBs) and **scans of paper lab journal** (*.pdf) or **elab journal**.
WP2	• **Vibrational data:** vibrations produced by honeybees and other living organisms in the honeycomb. File type: *.flac. Number of records: one file per hour, 88 Mbytes each. • **Sound data:** sound detected at the landing board. File type: *.flac, same size as 'vibrational data'. • **Acceleration data** will be collected on a continuous basis using accelerometers. The data will consist of hour / half an hour-long FLAC files around 50mb / 22 mb each. Over the course of the project, 26000 audio files will be generated per hive, per apiary. • **Monitoring data from new tools**: csv files: unknow (big files) + html files (incorporation of genetic profiles)
WP4	• **Transcripts of interviews** with 40 stakeholders, Verbatim transcripts from 40 interviews (40 records). Word format (2 MB) and NVivo format for content analysis. • **Quantitative survey data** from 200 stakeholders; SPSS data file *.sav with 200 records (2Mb). • **Quantitative scoring data** from 25 stakeholders, Excel data file *.xls. • **Survey data from surveys** with beekeepers. Two datasets, one with 40 records, one with 600 records. SPSS format (*.sav), max. 2 MB per dataset. • Data will be produced on the **socio-economic characterization of stakeholders**, on their perceptions on the bee health status index, SWOT analysis and beekeepers' business models. The methodologies that will be used to collect such data include semi-structured interviews, surveys and participatory workshops. File types will include those readable by such programs as Excel, SPSS, maxqda.
WP4 + WP8	• **Actor contacts:** contact details of actors/stakeholders (scientists, NGO representatives and beekeepers) used for interviews, workshops and surveys, 200+. File type: *.csv, *.txt and *.docx • **Actor interview data**: primary data on actor networks, beekeeping practices and bee health, circa 40 interview transcripts. File type: .csv, .txt and *.docx • **Actor workshop data**: primary data on beekeeping objectives, beekeeping business strategies and bee health indicators, 5 workshop transcripts, File type: *.docx and *.txt • **Actor quantitative survey data**: primary data on actor networks, beekeeping practices and bee health, circa. 200 survey records. File type: *.csv, *.txt and *.docx

_____

Work packages	
WP5	• **Beekeeping strategy**: defining priority questions of the beekeeping masters when inspecting a colony to reach a diagnostic. File type: *.docx and *.txt • **Simulation results** to establish the relationship between environmental (WP3), biological (WP1 and WP2) and management drivers (WP4 – WP8) and bee health status: analysis of simulation results. Size and number of records depends on the number of scenarios used.

Will you re-use any existing data and, if so, how?

Table 2 gives an overview of existing data that will be re-used**.**

Table 2: Overview of the external data which will be used during the project in the different work packages. This list may be extended as the research/data may develop.

Work packages	Overview of the external data
WP1	• **Weather service** (JSON API) • **LoRa messages** (TTN JSON API)
WP3	• Input from **national data** : *.csv / *.shp / *.georeference data (input data but not own data)
WP4	• **FADN data:** Farm Accountancy Data Network; Source: FADN - European Commission, DG AGRI-E.3. Data on apiculture economics and honey production as far as available and relevant. • Secondary data such as macro-statistical information on the **general characteristics of the beekeeping sector**.

What is the origin of the data?

See Table 1 and Table 2. The exact data providers and how these data will be used will be clarified during the project and will depend on the needs of the research.

To whom might the data be useful ('data utility')?

Different kinds of data will be produced in the B-GOOD project. The raw data will be of use for the broader scientific community. The final application integrating decision support will be of benefit for beekeepers, stakeholders, policy makers and others.

2.1 FAIR data: Making data findable, including provisions for metadata

Are the data produced and/or used in the project discoverable with metadata?

Metadata from all files (dataset, protocol, project result, etc.) will be generated using the international standard for metadata. Digital structured metadata ensures the traceability of the data. As a basis we will use the DublinCore standard. The metadata will be automatically generated where possible and further completed by the owner/creator of the file. The used vocabularies as well as the consisted terming will be controlled as much as possible. It will contain a set of information useful to make the associated data 'FAIR', prior to data deposition and storage. Keywords will be included which makes retrieval easier. The metadata will be made available by the web portal and will be searchable.

_____

Are the data produced and/or used in the project identifiable and locatable by means of a standard identification mechanism?

We will try to make the data as much as possible identifiable and locatable by making use of persistent and unique identifiers such as Digital Object Identifiers (DOIs). DOI will be used to link the data to the corresponding publications.

What naming conventions do you follow?

A: For project files: they should have descriptive labels and have the following pattern: YYYYMMDD_content of file_vx.file-format. All files will be accompanied by essential metadata. The file name includes a version number (vx) in order to distinguish the different versions from the same data. The adjustments made between versions should be described in the associated metadata. All data will be stored on the Internal Communication Platform (ICP).

Each partner will store his active data files and will deposit them in the file repository after quality control. We encourage the different partners to use the file hierarchy

BGOOD - WPx - Taskx - all files

on their local computers, which if data files contain personal data will be securely stored in compliance with guidelines set out in B-GOOD's WP10 Ethics deliverables..

B: For B-GOOD Datasets: We will use standardized metadata (see above) to enforce the user to store essential metadata. These will include date, subject, file type (csv, docx, etc.), version. We will enable users to describe the adjustments made between versions.

Data in this project will be created in different ways. Some data will be automatically recorded by means of sensors e.g. colony weight, temperature. Other datasets like those containing the results of e.g. pathogen analyses, genotyping, interviews and surveys will be created manually in spreadsheets or statistical analysis software amongst others (see table 1 and 2).

Not only the files should be named properly also the samples should be encoded in a standard manner. The coding of the samples will be described in detail in the sampling protocols (WP1) and should be maintained during the project.

Will search keywords be provided that optimize possibilities for re-use?

The following general keywords (e.g. honeybee, health status index, beekeeping management, IT-application, data flow, database platform, multi-actor approach) will be used to optimize the re-use and findability of the data. Depending on the data set more specific keywords may be added such as the names of pathogens, genotype, methodologies used, etc.

What is your approach for clear versioning?

Each file which is changed should be saved with a different version number. Draft versions will be numbered v0.1, 0.2, etc.

The first version which is approved by all co-authors will be numbered v1.0. Adding data to this file can be logged by changing this number to 1.1, 1.2 etc. All changes made to the original file should be documented in the corresponding metadata.

_____

What metadata will be created?

The metadata for the different project's datasets and files will be based on the DublinCore standard. During the project and depending on the generated data different fields may be added to this standard.

2.2. FAIR data: Making data openly accessible

Which data produced and/or used in the project will be made openly available as the default? If some data is kept closed provide a rationale for doing so. How will the data be made accessible?

Data will be made "as open as possible, as closed as necessary" during the project. The project intends to make all the data open in accordance to article 29.3 of the Grant Agreement. However, if any data would need to be closed or embargoed because of any legitimate circumstance foreseen by the ORD pilot, the DMP will be updated accordingly.

The generated data will be made openly available as early as possible during the project to increase use of data. This includes publishing in 'green' or 'gold' open access journals. The data sharing between the project partners will be done by the web portal for datasets and using the ICP on the project website b-good-projects.eu for project files.

The B-GOOD data web portal will allow for both closed and open sharing of datasets during and after the B-GOOD project. Usage (uploading and downloading) will be enabled for the bee research community at large. Project files will be ensured on the ICP from the website or shared with the public.

All B-GOOD tools and services will be compliant with the General Data Protection Regulation 2016/679 of the EU law on data protection and privacy.

What methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?

Most data collected and generated during the project will be stored using open formats (e.g. csv, txt), or at least, formats that are commonly used (e.g. xlsx, docx). These data will be accessible through a web browser. No other software is needed to find and access the data. To open specific (common) file formats, specific software may be required, e.g. SPSS in case of quantitative survey data.

Where will the data and associated metadata, documentation, and code be deposited? Have you explored appropriate arrangements with the identified repository?

During the project, project files will be stored on the ICP while B-GOOD datasets will be stored on the web portal. After publishing the results like described in the DoA of the project in Open Access journals, the underlying datasets will be encouraged to be made available through the OpenAIRE sharing web platform. The most appropriate repository will be determined at that time. This will enable automatic data extraction from the OpenAIRE platform, thus ensuring accessibility through a standard platform for Open Data access.
More technical details on the implementation of the data portal will be described later.

If there are restrictions on use, how will access be provided?

The datasets in the web portal will have different access restrictions. Authors of datasets will have the options to give explicit access to individual users or expose their datasets to the whole B-GOOD consortium members. Some data will be publicly accessible while others may

_____

be on request or completely restricted until publication. Also the project files on the B-GOOD website will have various access rights. There are two levels foreseen: access to B-GOOD partners only (on the ICP part of the B-Good project website) and access to the public (via the public B-GOOD project website). The data sets will be password protected and will allow user-specific rights and/or restrictions.

2.3. FAIR data: Making data interoperable

Are the data produced in the project interoperable? What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

The data will be made interoperable by publishing data in common machine-readable open formats (e.g. csv, txt) which can be assessed through open software. The used standards and methodologies regarding metadata amongst others are described in 2.1 and 2.2.

Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability? In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

Standard vocabularies will be used as much as possible for all datasets to ensure inter-disciplinary interoperability and re-use. If project-specific ontologies and vocabularies must be used appropriate mapping to more commonly used ontologies will be provided if possible.

2.4. FAIR data: Increase data re-use (through clarifying licenses)

How will the data be licensed to permit the widest re-use possible?

The project aims at a maximal degree of data re-use using appropriate open data licenses. Depending on the context different licenses will be considered. One of the most widely used open data licenses is the creative comments license but the country-specific requirements may necessitate the need for other licenses.

When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed.

The generated data will be made available as early as possible during the project to increase use of data. This includes publishing in 'green' or 'gold' open access journals with linkages to OpenAIRE, with accompanied data deposition in a public repository.

Are the data produced and/or used in the project usable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.

All data published in 'green' or 'gold' open access journals will be usable by third parties. The usability from other data will be described in the Data Management Plan at a later stage.

How long is it intended that the data remains re-usable?

At the end of year two, an Exploitation Plan will be developed (Deliverable 7.3) to describe how the project outputs will be exploited during and after the project. In task 6.3 a data platform will be developed for use during the program with the aim of continued use for many years after the end of the B-GOOD research program. The involvement of the EU BEE Partnership and BEEP will be further investigated and described in a later stage of the project. This will be updated in the DMP accordingly.

Are data quality assurance processes described?

The owner/creators of the dataset will be responsible for the quality of the data sets. Where possible an automatic quality control will be performed which can be done by the validation of

_____

the data entry (e.g. only numeric characters) which is already integrated in the BEEP digital beekeeping logbook.

3. Allocation of resources

What are the costs for making data FAIR in your project? How will these costs be covered?

There will be no extra project costs associated with the process of publishing data as data storage is described as Task 6.3 in the project. However, researchers need to spend time on proper data management, back up the data, creating documentation and metadata of each file etc.

The costs allocated to grant the open access are eligible and included in the project budget as foreseen in the Horizon2020 rules.

Who will be responsible for data management in your project?

Due to the different work packages in the project, the responsibility for the data management is divided between the partner institutions. As the project has a multi-disciplinary approach, different strategies will be needed to manage the specific work package data. Each work package leader will be responsible for the data management in his/her allocated work package. The overall coordination of the data management will be undertaken by the project coordinator in assistance with the Data Support of the coordinator and the Data Officer.

What are the costs and potential value of long term preservation?

Deliverable 7.3 will develop an Exploitation plan in month 24 of the project, which will probably describe the costs and potential value of long term preservation of the output. This will be updated in DMP accordingly.

4. Data security

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

We will adhere to the Open Web Application Security Project (OWASP) guidelines.

Data collection and storage in this project will be done on different levels.

Different partners will generate data. Each partner is responsible for the robust storage and backup of the active data during the experiments. We will encourage the different partners to back-up their datasets regularly and to use the institution's network drive when available next to their computer hard drive and/or external storage device. All data files will be identified by a version number and creation date. Logging the different versions and notes detailing differences between versions has to be documented in e.g. a spreadsheet.

All datasets generated and up-loaded on the platform (WP6), will be stored and backed up. In WP6 the BEEP team will use a highly available and redundant datastore, using the features of Amazon S3 (https://aws.amazon.com/s3/features/?nc=sn&loc=2).

_____

5. Ethical aspects

Are there any ethical or legal issues that can have an impact on data sharing?

- Yes.

Personal data will be collected from human participants (adult healthy volunteers, namely stakeholders and beekeepers) involved in the social sciences studies in B-GOOD's WP4 and WP8. These data will be collected by means of personal interviews, participatory workshops and surveys. The data will be shared for analysis purposes among the B-GOOD partners involved in the concerned work packages. All relevant ethical or legal issues in this respect have been addressed in detail with the B-GOOD WP10 Ethics deliverables, notably "D10.1 H - Requirement no. 1" (on studies involving human participants) and "D10.2 POPD - Requirement no. 2" (on the protection of personal data). Amongst others, these deliverable describe the informed consent procedures including consent for data protection and sharing, handling of the data including storage and transfers, protection of the identity of the study participants, pseudonymisation techniques, security measures, and reference to ethics committees for study approvals.

6. Other issues

Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?

The underlying data will in many cases be governed by national, regional, or institutional rules.

The host institution from the coordinator, UGENT, has different institutional policies.

- The Research Data Management (RDM) policy: https://www.ugent.be/en/research/datamanagement/rdm-policy.pdf/at_download/file
- Information security: https://www.ugent.be/en/facilities/ict/information-security
- Processing personal information and other confidential data will be performed following the UGent Code of Conduct for the Processing of Personal Data and Confidential Information

_____

GDPR section: Collection and processing of data

1. Are you collecting or processing personal data?

- Yes

One of the key objectives of B-GOOD is to map the business environment and to study the socio-economics of beekeeping. As a result, personal data will be collected from human participants involved in socio-economic research. The socio-economic studies and related data collection are covered in WP4 and WP8 of the project.

The proposed studies and chosen multi-actor approach require the involvement of human participants and the collection of primary personal data from actors directly involved in beekeeping. They participate as individuals in their role as stakeholders and beekeepers and as adult healthy volunteers involved in social sciences research.

2. Are you collecting or processing primary personal data and/or secondary personal data?

- Primary personal data

B-GOOD will only collect personal data from the human participants enrolled in the socio-economic studies within WP4 and WP8 that are strictly necessary to achieve the objectives of the research. These objectives concern the performance of social science studies and the implementation of socio-economic statistical analyses such as production and economic efficiency analyses.

The collected personal data will include, depending on the target population: socio-demographic characteristics such as age (years), gender, education, training (necessary for the profiling of aggregated segments) and urban-rural living environment (necessary for linking with environmental and ecological characteristics), as well as attitudinal (attitudes, beliefs, perceptions, opinions and views) and behavioural (management decisions, decision-making processes) characteristics, which will all exclusively relate to beekeeping and its context. All collected socio-economic data are cross-sectional data collected at one point in time.

Sensitive personal information relating e.g. to health, ethnicity, sexual lifestyle, political opinion, religious or philosophical conviction falls beyond the scope of B-GOOD and will not be probed for.

No confidential data, such as information on private businesses, sensitive business practices, finances or income will be collected from stakeholders. In the case of beekeepers, and in line with the objective of performing production and economic efficiency analyses in task 4.2, data relating to business, costs and revenues will be collected. In case these data are seen as and indicated to be confidential by the beekeeper but nevertheless voluntarily provided, they will be strictly treated as data given in confidence or data agreed to be kept confidential between the researchers and the beekeeper. These data will be kept secret and out of the public domain. Any reporting based on these data will be done in aggregated and non-identifiable form only. This will be explained to the beekeepers as part of the informed consent.

- Secondary personal data

B-GOOD will make use of the secondary data sets such as weather service data, LoRa message data, national geo-reference data, FADN data, and statistical information on the general characteristics of the beekeeping sector (as indicated in Table 1 of this DMP). From these, only the FADN data may eventually contain personal data.

_____

The interest of B-GOOD pertains to FADN data relating to "843 – Apiculture" and "SE251 – Other livestock and products (incl. honey)". The use of FADN data is conditional upon a formal agreement between the "EC – Directorate-General for Agriculture and Rural Development – E.3 Economic analysis of EU agriculture (DG AGRI-E.3)" and DG RTD and the involved B-GOOD partner as Beneficiaries. This agreement will stipulate the terms and conditions for use of the data. This will include, amongst others:

- The guarantee that the data will not be used for any other purpose;
- Observation of the rules on non-disclosure of the data and the secrecy of statistics;
- Not to publish a result when it is based on less than 15 observations (holdings);
- The specification of the only authorized persons to handle the FADN farm level data;
- The commitment to erase all FADN data once the project is completed;
- The commitment to take all necessary actions to protect the data.

In case information on individual farms (i.e. commercial beekeepers in our case) is obtained for the specific research purposes of B-GOOD, it will be ensured that data records are pseudonymised and protected in line with the procedures described in the Ethics Deliverable D10.2 on the protection of personal data (WP10), and with the terms and conditions for its use as agreed with DG AGRI-E.3.

3. If you are processing secondary personal data, will you inform the persons whose personal data are being processed or have they already been informed?

No

4. If no, explain why it is impossible or why it would take a disproportionate effort to inform the persons whose personal data are being processed.

If applicable, data records will be pseudonymised and de-identified when received from FADN. The agreement between the data provider and the beneficiary will stipulate amongst others that no results can be published when based on less than 15 observations, herewith safeguarding the rights and freedoms of individual data subjects.

GDPR section: Categories of personal data & data subjects

5. Are you collecting/processing any of the following special categories of data?

- ☐ Data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union memberships
- ☐ Genetic data
- ☐ Biometric data for the purpose of uniquely identifying a natural person, such as audio recordings of voices, fingerprints, facial images, iris scans
- ☐ Data concerning health (physical and mental health), an individual's sex life, or an individual's sexual orientation
- ☐ Data relating to criminal convictions and offences
- ☑ None of the above

6. Which other categories of personal data are you collecting/processing?

- Identification data (names, titles, addresses, phone numbers, passport numbers, IP addresses, cookies, electronic location data (GPS, mobile phone)…)

_____

- Personal characteristics (age, gender, date of birth, marital status, nationality…)
- Education and training
- Audio recordings

7. Whose personal data are you collecting/processing?

☐ Children below the age of 13

☐ Children between the age of 13 and 16

☐ Children over the age of 16

☐ Other vulnerable persons (e.g. pregnant women, elderly persons, people with mental disorders, asylum seekers, disabled persons, ethnic minorities, sick people or patients)

☑ Others: stakeholders involved in the beekeeping sector and hobbyist and professional beekeepers

8. Will your research be seriously hampered if the persons whose personal data are being collected/processed exercise their right to access, to rectification, to restriction of processing, to be forgotten, to data portability and/or to object?

- No

9. If yes, please justify the need to deviate from one or more of the rights mentioned in question 8. A justification is required for each deviation.

GDPR section: Purpose(s) of the processing

10. What is/are the purpose(s) of the personal data processing?

One of the purposes of B-GOOD is to map the business environment and to study the socio-economics of beekeeping. As a result, personal data will be collected from human participants involved in socio-economic research. The socio-economic studies and related data collection are covered in WP4 and WP8 of the project. The proposed studies and chosen multi-actor approach require the involvement of human participants and the collection of primary personal data from actors directly involved in beekeeping. They participate as individuals in their role as stakeholders and beekeepers and as adult healthy volunteers involved in social sciences research. The collected personal data will include, depending on the target population: socio-demographic characteristics such as age (years), gender, education, training (for the purpose of profiling of aggregated segments) and urban-rural living environment (for the purpose of linking with environmental and ecological characteristics), as well as attitudinal (attitudes, beliefs, perceptions, opinions and views) and behavioural (management decisions, decision-making processes) characteristics. These will all exclusively relate to beekeeping and its context, and its collection and analysis is necessary for the purpose of providing policy recommendations and guidelines fostering healthy and sustainable beekeeping in the EU.

11. What is the legal ground for the processing? If the data are being processed for multiple purposes, you must describe the legal ground for each purpose.

The legal ground refers to one of the key objectives of B-GOOD, which is to map the business environment and to study the socio-economics of beekeeping. As a result, personal data will be collected from human participants involved in socio-economic research.

_____

GDPR section: GDPR responsibility

12. Is there another university, research institute or partner involved in the research (besides Ghent University)?

- Yes.

The main partners involved in the collection, analysis and storage of personal data within B-GOOD are partners UGENT, AU, UCOI and WR.

13. If yes, please specify who determines the purposes ('why') and the means ('how') of the research.

This is determined together with a researcher/university/institution outside UGENT: UGENT is a joint controller

UGENT is the overall project coordinator and leader of WP4 dealing with socio-economic studies where most of the personal data will be collected. UGENT will act as the data controller of most of the B-GOOD personal data. Aarhus University is the leader of WP8 dealing with the multi-actor approach within B-GOOD. AU will perform participatory workshops in collaboration with UGent and other project partners. In cases where AU is the principal investigator, AU will abide GDPR and AU internal guidelines that are equivalent to those at UGent. Ethics approval will be obtained in both sites when applicable.

As a general guideline, the generic code of conduct for the processing of personal data within Ghent University, the institution of the B-GOOD coordinator and the B-GOOD WP4-leader, will be shared with the consortium partners and recommended to be adhered to.

GDPR section: Data transfers & categories of recipients

14. Are you disclosing/sharing/transferring personal data beyond your project team, either with recipients in UGent or UZ Gent, or with external recipients?

- No

15. If yes, to or with which categories of recipients are the personal data being disclosed/shared/transferred?

16. If yes, where are the personal data being disclosed/shared/transferred to?

17. What is/are the purpose(s) of the data transfer?

18. What is the legal ground for the data transfer? If there will be multiple data transfers, you need to indicate the legal ground for each data transfer.

GDPR section: Retention period

19. What is the envisaged retention period for the different categories of personal data? Please motivate.

In line with the UGENT General Code of Conduct for the Processing of Personal Data and Confidential Information, Article 5, §7, users shall ensure that the storage period/retention period of personal data and confidential information is determined in accordance with all relevant legal provisions and applicable agreements. In addition, the storage period/retention

_____

period should be limited to what is necessary and in accordance with the original purposes. Exceptions for longer retention can only be made within the context of the legislation or regulations established for this purpose (e.g. for scientific or historical research or statistical purposes, for archiving in the public interest, or for further research or control mechanisms for scientific integrity). After the retention period has expired, the data have to be completely and securely deleted, in accordance with the guidelines in the information security policy of UGENT. As a general guideline for scientific integrity purpose, a retention period of 10 years will be envisaged.

GDPR section: Risk analysis

20. To analyse the possible risks associated with the processing of personal data, please tick the boxes that apply to this research.

- The data are transferred beyond the borders of the EU or the EEA, or to a country not listed on the 'white list' (see question 16)

Personal data from stakeholders and beekeepers will be exchanged between the EU and Switzerland (UBERN). No personal data within B-GOOD will be exchanged with any other non-EU country than Switzerland. On the basis of article 45 of Regulation (EU) 2016/679, the European Commission has adopted an Adequacy Decision (the so-called 'white list') recognising Switzerland as providing adequate protection.

21. Does the research constitute a probable high-risk processing? If you ticked two or more boxes in question 20, the answer is 'yes'.

- No

22. If yes, explain why the processing is necessary to achieve the purposes of the research.

23. If yes, provide more details for any of the risks that you have ticked in question 20.

GDPR section: Security measures

24. What technical and organisational security measures are in place to protect personal data?

The principle of transparency will be adhered to during all collection of personal data within B-GOOD. This implies that data subjects/research participants will be properly informed and asked to provide informed consent. This will be done in a concise, transparent, intelligible and easily accessible format using clear and plain language, and using study-specific informed consent forms. Data subjects will be informed about the purpose of the research, what is expected from them, eventual benefits and risks involved, and steps taken to safeguard their anonymity and confidentiality. They will be informed that their participation is voluntary and that refusal to participate will involve no penalty or loss of benefits. They will also be informed about the plans with their contribution in terms of analysis, reporting, future archiving and sharing of their data.

The identity and contact details of the principal researcher and the Data Protection Officer (DPO) of the institution responsible for the data collection will be communicated to the data subjects. Information about the rights of the data subjects and how they can exercise these will be provided, including the right to revise their personal data, request their data to be removed, and to lodge a complaint with the supervisory authority.

_____

Adequate security measures will be implemented to prevent unauthorised access to personal data or the equipment used for processing.

B-GOOD research activities involving human participants and personal data have been designed following a risk-based approach. Taking into account appropriate technical and organisational measures to safeguard the rights and freedoms of the study participants (as described in Section 4) and to protect the data, the activities have been defined as low-risk.

Based on general principles within UGent that will be shared with all other B-GOOD partners, all researchers involved in B-GOOD are asked to commit to work on trusted networks, to use trustworthy devices, and to protect their institutional accounts and associated login data in line with their institutional guidelines and data protection policies.

Data will be stored on safe places only, i.e. on password-protected central disk spaces/storage (such as personal disk space and shares at institutional level), which are backed-up automatically and protected by institutional security systems.

The use of external cloud services to store personal data or confidential information will be avoided, unless this data or information is encrypted. Personal data in paper or any other physical form (e.g. non-digital audio- or video-recordings) will be kept in a secured area of a locked filing cabinet.

Access to personal data will be strictly limited to the researchers who are directly involved in the analysis of the concerned data. The principal researcher will ensure this data is password-protected and communicate passwords in person or in encrypted format (not in written). Passwords will be changed at regular intervals. Access conditions will be arranged and agreed upon during project meetings and decisions included in the minutes of the meeting. In case of encrypted data, files containing the encrypted data and the encryption keys will be kept and/or sent separately from the data.

B-GOOD will make use of the encryption functionalities of SPSS (in case of survey data collected from stakeholders and beekeepers, which will also be analysed using SPSS) and of Microsoft Office (in case of the interview and workshop transcripts as text files). In case the specific analysis software used does not provide encryption possibilities, 7-zip compression software.

All data will be pseudonymised, stored in a de-identified format, kept securely and shared for study purposes and in dissemination activities only in pseudonymised or aggregated form. Transcriptions of interviews and workshops, as well as data records from surveys, will not include the name(s) of the interviewees. Instead, unique personal identifiers will be attributed to participants and used in the transcripts or with the quantitative data records. Personal identifiers and transcripts will be stored and secured separately.

As a general guideline, the generic code of conduct for the processing of personal data within Ghent University, the institution of the B-GOOD coordinator and the B-GOOD WP4-leader, will be shared with the consortium partners and recommended to be adhered to.

25. If you answered 'yes' to question 21 (which means that your research constitutes a probable high-risk processing), indicate whether the security measures taken are sufficient to protect the rights and freedoms of the people whose data you collect/process.

Not applicable.

_____

26. If you have motivated the need to deviate from one or more of the rights of the persons whose personal data you are collecting/processing in question 8 and 9, please describe which safeguards are put in place to protect their rights and freedoms.

Not applicable.