

Data Health Components No.1

Deliverable D1.3

27 May 2021

Coby van Dooremalen¹, Zeynep Ulgezen¹, Dirk Jan Valkenburg¹, Xiaodong Duan², Ugoline Godeau³, Pim van Gennip⁴, Marten Schoonman⁴

¹WR, ²AU, ³INRAe, ⁴BEEP

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 Health Components No.1

Deliverable n°: D1.3

Nature of the deliverable: Report

Dissemination level: Public

WP responsible: WP1 Lead beneficiary: WR

Citation: van Dooremalen, C., Ulgezen, Z.N., Valkenburg, D.J., Duan,

X., Godeau, U., van Gennip, P., & Schoonman, M. (2021). Data Health Components No.1. Deliverable D1.3 EU Horizon 2020 B-GOOD Project, Grant agreement No.

817622.

Due date of deliverable: Month n° 24 Actual submission date: Month n° 24

Deliverable status:

Version	Status	Date	Author(s)					
0.1	Draft	12 May 2021	Dooremalen, Ulgezen, Valkenburg, Duan, Godeau, Gennip, Schoonman					
			WR, AU, INRAe, BEEP					
0.2	Review	25 May 2021	Schäfer, Beaurepaire, Dezmirean					
			FLI, UBERN, UCLUJ					
0.3	Review	27 May 2021	Dooremalen					
			WR					
1.0	Final	27 May 2021	Dooremalen, Ulgezen, Valkenburg, Duan, Godeau, Gennip, Schoonman					
			WR, AU, INRAe, BEEP					

The content of this deliverable does not necessarily reflect the official opinions of the European Commission or other institutions of the European Union.

Table of contents

Pre	face	3
	mmary	
	Introduction	
	Data collection method	
	Data storage and digital visualisation	
4.	Data collected (meta data)	8
5.	Raw data processing	10
6.	Data portal	10
7.	Acknowledgements	12
R	References	13

Preface

In this report we present the deliverable **D1.3**: **Data Health Components No.1** by Month 24 of the project B-GOOD. This report contains the meta data of all raw data that were collected within Tier 1 in the first field year of the project (year 2020). The data consist of different health components obtained in a classic, automated or semi-automated way, or by lab analyses, covering different biogeographic regions of the EU. The raw data underlying data for this report, are being checked, double-checked, and cleaned whenever needed. The raw data will be uploaded in the data portal (**Milestone MS36 Website Portal Phase 1**) for further analyses within the B-GOOD project (e.g. in WP5 to compose the Health Status Index) and will become public in full alignment with the project's B-GOOD publication and data sharing policy.

Summary

In this report we describe

- 1) The set-up of the data collection infrastructure (i.e. the experimental set up) and how the data are collected.
- 2) How the data are stored, what they look like, and how they can be downloaded
- 3) A meta-data overview of all raw data collected (e.g. variables, type of measurement, frequency of measurement
- 4) The data cleaning process that we will follow to make the raw 'unfiltered' data ready for uploading to the data portal and further analyses
- 5) The data portal

1. Introduction

B-GOOD stands for 'Giving Beekeeping Guidance by cOmputatiOnal-assisted Decision making'. The overall aim of the 4-year B-GOOD project is to pave the way towards healthy and sustainable beekeeping within the EU. A key to healthy beekeeping is the Health Status Index (HSI), inspired by EFSA's Healthy-B toolbox. Honey bee health can be assessed by 'indicators' associated with colony attributes (e.g. amount of brood, disease load) and colony outputs (e.g. pollination service, honey harvest), and 'factors' associated with external drivers (e.g. resource providing units), which will be extended with a fourth set of components related to the honey bee gene pool (e.g. local adaption, subspecies, ecotype).

We will make this enhanced HSI operational by working towards an automated or semiautomated data flow from various sources into a common EU-wide bee health and management data platform, and by testing and validating (or discarding) each component thoroughly, eventually leading to the objective selection of the most promising and relevant components.

In the B-GOOD project, WP1 is responsible to facilitate and standardize large scale data collection on honey bee health indicators and genepool charateristics across the EU. This **Deliverable D1.3: Data Health Components No.1**, is the first out of three deliverables that show what kind of raw data is collected in WP1, and how it is processed and stored before further use in the B-GOOD project (by e.g. WP5 to perform context dependent risk assessment for bee colony management and to give guidance in decision making).

2. Data collection method

Data collection on colony (health) status related variables occurs according to a 3-Tiered process (Figure 1) that spans three bee seasons, where Tier 1 (blue) involves Pilot A and B at apiaries of research institutes of B-GOOD and runs for three years, Tier 2 (green) involves Field study A at EU apiaries of selected beekeepers guided by a local B-GOOD partner and runs for two years, and Tier 3 (yellow) involving Field study B at EU apiaries of randomly selected beekeepers with coordination only, running for one year. With

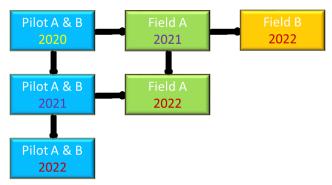


Figure 1. Data collection on colony (health) status related variables occurs according a 3-Tiered process that spans 3 bee seasons.

progression of the Tiers more variables will come into place as the dimension of the testing shifts from a restricted local (institutional) EU country coverage level, via an EU country coverage level on a North South axis, to a pan-European level (with among other things, a larger variation in subspecies, ecotypes, gene pool, environment, beekeeping management and business model). For more information on the installation of Tier 1, see Milestone MS1 Mini-Apiaries and Milestone MS2 Pilot & Field Study No.1.

This Deliverable report (D1.3) contains the meta data of all raw data that were collected within Tier 1 in the first field year of the project (year 2020). Within Tier 1 we tested the full set of health components suggested by the <u>EFSA's Healthy-B toolbox</u> (all indicators with high relevance, high technical feasibility, and high priority). The data consist of different health components obtained in a classic, automated or semi-automated way, or by lab analyses, covering different biogeographic regions of the EU.

In Pilot A, there are eight apiaries installed, each with 8 honeybee colonies (hence they are called mini- apiaries). Each mini-apiary is kept by one of the eight partner institutes, across eight different countries, at specific locations (including WR, Figure 2). Each B-GOOD mini-apiary started the experiment with eight presumably healthy colonies (i.e. not sick and performing within normal parameters related to their purpose, e.g. honey harvest). These colonies within the apiaries are the main bodies of research within Pilot A of Tier 1 in WP1 to operationalize the EFSA's Healthy-B toolbox and to validate the B-GOOD (semi)automated monitoring tools that are technologically sufficiently ready for use (In the Grant Agreement described as technological readiness level >6).



Figure 2: B-GOOD mini-apiary locations for partners UGENT (1), WR (2), INRA (5), MLU (6), UCLUJJ (7), UCOI (9), TNTU (10), and UBERN (11)

Pilot B involves one apiary at partner WR. It contains 25 colonies of honeybees, exposed to stressors in groups of five colonies (including five control colonies). Experimental stress exposures that were applied in 2020 were, induced change in: i) reproductive status, ii) natural parasite load, iii) exposure to the neonicotinoid acetamiprid, and iv) reduced pollen resource

availability. Monitoring and data collection methods in Pilot Study B are similar to those of Pilot Study A.

For data collection, a BEEP Base Sensor System was installed under each honey bee colony (89 systems in total) for automated measurements. The BEEP bases include sensors for weight of the hive, in-hive temperature and in-hive sound. A LoRa gateway was installed for wireless data transmission, as well as a local weather system. Furthermore, field observations and sampling for lab analyses on diseases were performed synchronously according to the retained (and described) procedures at all apiaries, as described in **Deliverable Report D1.1.** and **Deliverable Report D.1.6.**.

For collecting health components data of the all the participating B-GOOD apiaries, we used scientific protocols and manuals that were adapted and optimised for the purpose of the B-GOOD project, including support manuals for using the BEEP digital logbook app and BEEP base (see **Deliverable Report D1.1** and **Deliverable Report D.1.6**. on Protocols No1 and No2, respectively). Field observations of colonies consists of seven protocols in 2020, parameters were measured over time through the year, and were entered in the BEEP digital logbook app. All data, including automated data, field observations and management actions, were collected through the BEEP digital logbook app. Direct communication with all partners was made available using Microsoft Teams, with meetings on a monthly base to discuss project progression, tackle user problems, and gain feedback.

3. Data storage and digital visualisation

The BEEP digital logbook app (https://beep.nl/home-english) and the BEEP base measurement system are used to collect inspection data and sensor information respectively. When data is entered in the logbook and saved it is immediately stored in the BEEP database. The BEEP base collects sensor information which is transmitted to a time-series database wirelessly. Both information sources are available for download, cleaning and subsequent analysis. Access to the data is organized via the B-GOOD publication and data sharing policy.

BEEP provides a Research facility to users of the BEEP digital logbook app. This feature allows to share data in the B-GOOD research and provide consent to do so. This data is accessible to pre-defined B-GOOD researchers to download inspection and sensor data. Via an administration module of the BEEP platform, data records can be consulted and downloaded. Furthermore, a live data connection can be set up to automatically fetch data at a regular interval via an Application Programming Interface (API).

The current activity in the different mini apiaries on the BEEP digital logbook app is monitored in the research portal provided by partner BEEP (in WP6). Figure 3, shows some print screens on how the data are collected and visualized for the end-users (e.g. research institutes or beekeepers).

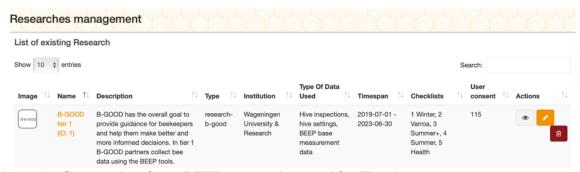


Figure 3a: Screenshot from BEEP research portal for Tier 1

								₽ Relo	ad cons	sent dat	a table												
	,	2020-12-10	2020-12-09	2020-12-08	2020-12-07	000-12-06	020-12-05	2020-12-04	2020-12-03	1020-12-07	2020-12-01	000-1-25	2020-1-28	000-1-28	020-1-21	000-1-20	020-1-25	000-1-24	020-1-23	000-1-22	020-1-21	000-1-20	2020-11
Users (8)	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
Apiaries (13)	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13
- Hives (93)	93	93	93	93	93	93	93	93	93	93	93	93	93	93	93	93	93	93	93	93	93	93	93
Inspections (168)			18				16			8		8			10		8		8			18	
Devices (74)	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74	74
Measurements (201951)	5175	5186	5068	5231	5193	5197	5211	5216	5157	5016	5154	5113	5106	5070	5102	5059	5037	5105	5092	5008	5086	5130	512
Weather data (0)																							

Figure 3b. Screenshot of the data overview of all collaborators in Pilot A in the Research Portal. It shows the data collection and past activities in Tier 1. The amount of participants, hives, measurements (sensor) and inspections (field measurements) are displayed. The screenshot shows the period 19 November – 10 December 2020, but the total amount of inspections from 2020, and inspections per day are visualized

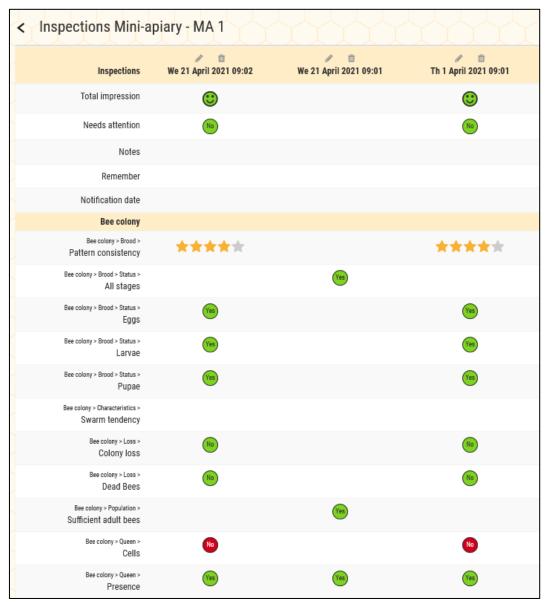


Figure 3c. Screenshot an inspection sheet, e.g. the inspections of Mini Apiary- colony No. MA1 of the Apiary at WR (Netherlands)

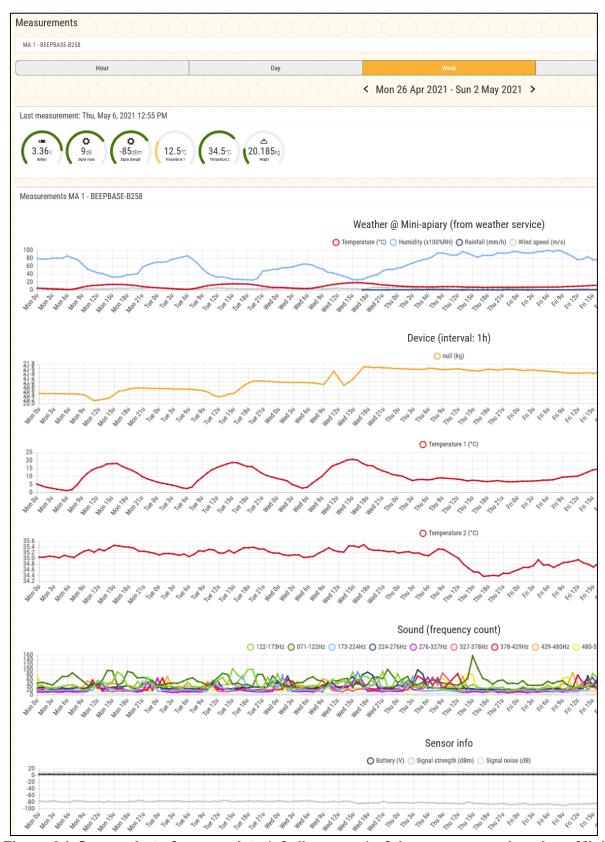


Figure 3d. Screenshot of sensor data (of all sensors) of the same example colony Mini Apiary- colony No. MA1 of the Apiary at WR (Netherlands)

4. Data collected (meta data)

In this chapter we show a meta-data overview of all raw data collected in the first field year of the B-GOOD project (Table 1), which only involves Tier 1 data at this moment in time (see Figure 1). Table 1. shows all variables that we logged in our B-GOOD colonies. These variables involve sensor data, annotation data, experimental observations (logged via inspections), laboratory analyses, management actions, weather data, and their frequency. For management actions, due to the infinite amount of options available, we only show what actions have been registered over 2020. The raw data underlying data for this report, are being checked, double-checked, and cleaned where needed, and will be uploaded in the data portal for further analyses within the B-GOOD project (e.g. in WP5 to compose the Health Status Index).

Table 1. All variables and frequency of measurements that we logged in our B-GOOD

colonies for Tier 1 during the Field year 2020

Variable	Category	Data / Units	Frequency of measurements				
Weight	Automated data	kg	15 mins				
Ambient temperature	Automated data	°C (Celsius degrees)	15 mins				
In-hive temperature	Automated data	°C (Celsius degrees)	15 mins				
Sound	Automated data	Frequency count (122- 583hz)	15 mins				
Battery	Automated data	Volt	15 mins				
Signal strength (data transmission)	Automated data	dBm	15 mins				
Signal noise (data transmission)	Automated data	dB	15 mins				
Sufficient adult bees	Data annotation	Yes/no	7 - 30 days				
Brood in all stages	Data annotation	Yes/no	7 - 30 days				
Presence of queen	Data annotation	Yes/no	7 - 30 days				
Suitable space	Data annotation	Yes/no	7 - 30 days				
Absence of stressors	Data annotation	Yes/no	7 - 30 days				
Sufficient nutrition	Data annotation	Yes/no	7 - 30 days				
General impression	Experimental observation	Good, average, bad (smileys)	7 - 30 days				
Eggs	Experimental observation	Estimated number of cells	3 times a year (spring, summer, fall)				
Larvae	Experimental observation	Estimated number of cells	3 times a year (spring, summer, fall)				
Bees	Experimental observation	Estimated number of cells	3 times a year (spring, summer, fall)				
Pollen	Experimental observation	Estimated number of cells	3 times a year (spring, summer, fall)				
Sealed honey	Experimental observation	Estimated number of cells	3 times a year (spring, summer, fall)				
Pupae (capped brood)	Experimental observation	Estimated number of cells	3 times a year (spring, summer, fall)				
Drone brood	Experimental observation	Estimated number of cells	3 times a year (spring, summer, fall)				
Atypical behaviour	Experimental observation	Yes/no	3 times a year (spring, summer, fall)				
Colony loss	Experimental observation	Yes/no	Every 30 days				

Variable	Category	Data / Units	Frequency of measurements				
Dead bees	Experimental observation	Yes/no	Every 30 days				
Varroa natural fall	Experimental observation	mites/day	Once a week				
Clinical signs of disease	Experimental observation	Categorized by type	Every 30 days (beekeeping season)				
Presence of eggs	Experimental observation	Yes/no	Every 30 days (beekeeping season)				
Presence of larvae	Experimental observation	Yes/no	Every 30 days (beekeeping season)				
Presence of pupae	Experimental observation	Yes/no	Every 30 days (beekeeping season)				
Queen presence	Experimental observation	Yes/no	Every 30 days (beekeeping season)				
Top photo analysis	Experimental observation	Estimated number of bees	Every 30 days				
Varroa destructor	Lab analyses	Mites/100 bees	3 times a year (spring, summer, fall)				
Deformed wing virus - A	Lab analyses	PCR data	3 times a year (spring, summer, fall)				
Deformed wing virus - B	Lab analyses	PCR data	3 times a year (spring, summer, fall)				
Acute bee paralysis virus	Lab analyses	PCR data	2 times a year (spring, fall)				
Chronic bee paralysis virus	Lab analyses	PCR data	2 times a year (spring, fall)				
American Foulbrood	Lab analyses	PCR data	Once a year (fall)				
European Foulbrood	Lab analyses	PCR data	Once a year (fall)				
Nosema ceranae	Lab analyses	PCR data	2 times a year (spring, summer)				
Nosema apis	Lab analyses	PCR data	2 times a year (spring, summer)				
Sacbrood virus	Lab analyses	PCR data	3 times a year (spring, summer, fall)				
Black queen cell virus	Lab analyses	PCR data	3 times a year (spring, summer, fall)				
Foundationless frame	Management actions	Yes/no	When necessary				
Comb replaced	Management actions	# of combs replaced	3 times a year (spring, summer, fall)				
Nutrition/sugar feeding	Management actions	Weight/volume	When necessary				
Swarming control	Management actions	Yes/no/#	When necessary				
Swarming prevention	Management actions	Method	When necessary				
Queen introduction	Management actions	Reason and method	When necessary				
Queen marking	Management actions	Colour	When necessary				
Colony split	Management actions	Yes/no	When necessary				
Colony united	Management actions	Yes/no	When necessary				
Colony feeding	Management actions	Volume/weight	When necessary				

Variable	Category	Data / Units	Frequency of measurements
Honey harvest	Management actions	Weight/volume	When necessary
Varroa treatment	Management actions	Method	When necessary
Temperature	Weather (from weather service)	°C (Celsius degrees)	15 mins
Wind speed	Weather (from weather service)	m/s	15 mins
Humidity	Weather (from weather service)	x100%RH	15 mins
Rainfall	Weather (from weather service)	mm/h	15 mins

5. Raw data processing

In order to make the raw data ready for analyses, it needs further processing. The data cleaning process to make the raw 'unfiltered' data ready for uploading to the data portal and further analyses, involves to following steps:

- 1. The research data in BEEP database, including automatic data, weather data, inspection data, is downloaded using BEEP API by using a Python script. Several data files are generated after the downloading as below.
 - a. Raw data in json files.
 - b. Parsed data in csv files.
- 2. The downloaded data is processed by the same Python script in order to find missing data gap and abnormal changes in the automatic data. After this step, a report is generated for each apiary.
- 3. If the abnormal changes in the automatic data cannot be explained by inspection activities or missing data gap, the apiary manager will be contacted for further investigations. The abnormal changes will be fixed based on the inspection note and feedback from the apiary managers.
- 4. For inspection data, a check of the respect of the planned schedule is carried out (count Varroa, top photo, colony evaluation ...). If any inspections appear to be missing, the apiary manager is contacted to determine if the inspection has not taken place or if the data has not been entered.
- 5. After these steps, the data from each apiary will consist of two csv files for a calendar year containing automatic data and inspection data respectively. In addition to these two files, each apiary has one csv file with weather data for a calendar year.

6. Data portal

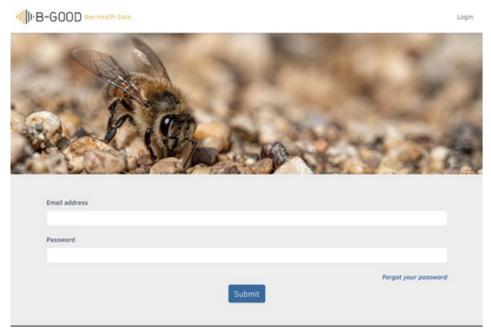
The processed data will be uploaded onto the data portal, that is provided for by partner BEEP (**Milestone MS36 Website Portal Phase 1**, April 2021). The first phase of this portal has very recently been finalized and is now ready to have the first dataset being uploaded in full alignment with the B-GOOD publication and data sharing policy.

The EU-wide bee health data web portal ('data portal' in short) is used to store and share B-GOOD datasets. Honey bee colony health data is the main subject. The main purpose of the portal is to store raw and pre-processed data generated in the B-GOOD project. This data is used for research purposes. Datasets can be uploaded and retrieved depending on the access rights. Access can also be requested. A dataset can consist of multiple files and metadata is

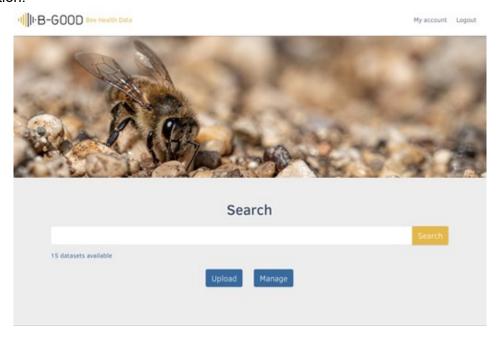
used to describe the datasets. A large variation of file types is supported. Where and when possible, datasets are shared openly. An account is required to be able to use the portal.

The B-GOOD data portal is available via https://beehealthdata.org/. BEEP Foundation is responsible for the hosting of the database and website, the maintenance thereof and providing support via the BEEP helpdesk (**Deliverable D1.2 Helpdesk**).

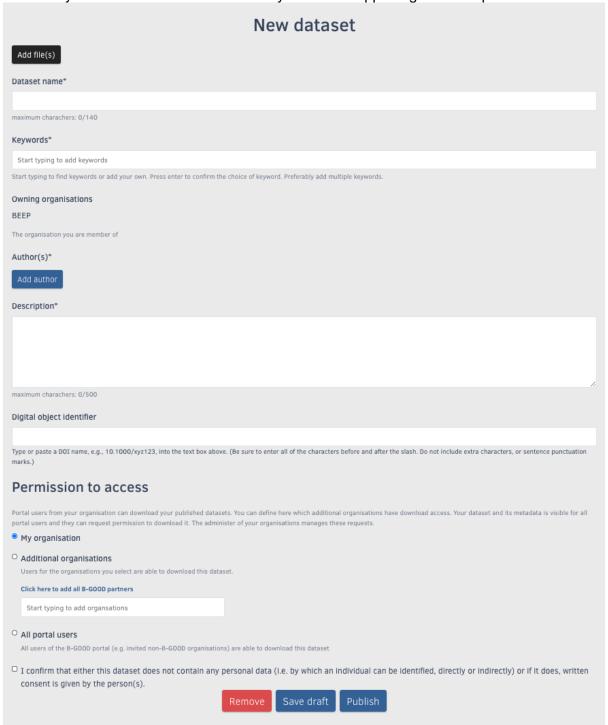
On the following pages, we present some screenshots of the portal related to uploading a dataset (also presented in **Milestone MS36 Website Portal Phase 1**). Login screen



Homepage with Search feature and an indication of number of datasets available. Also, from here a new dataset can be uploaded and a user can manage his/her datasets and account information.



New dataset upload page. A dataset can consist of multiple files and versions of those files. Mandatory fields are indicated with an '*' symbol and supporting texts are provided.



7. Acknowledgements

We thank all mini-apiary partners (UGENT, INRA, MLU, UCLUJ, UCOI, TNTU, UBERN, WR) for their continuous effort they put managing their apiaries and doing all the measurements. We thank BEEP, FLI, SCIEN, AU, INRAe, for their constructive collaboration within WP1 on data collection and processing. We thank all reviewers for giving feedback on the protocols, for writing and for reviewing this deliverable.

8. References

BEEP digital logbook app https://beep.nl/home-english

B-GOOD data portal https://beehealthdata.org/

EFSA AHAW Panel (EFSA Panel on Animal Health and Welfare), 2016. Scientific opinion on assessing the health status of managed honeybee colonies (HEALTHY-B): a toolbox to facilitate harmonised data collection. EFSA Journal 2016;14(10):4578, 241 pp. doi:10.2903/j.efsa.2016.4578

Schoonman M. (2021). **Website Portal Phase 1. Milestone M36** EU Horizon 2020 B-GOOD Project, Grant agreement No. 817622

Schoonman, M., De Smet, L. & de Graaf, D.C. (2020). Publication and Data Sharing Policy. EU Horizon 2020 BGOOD Project, Grant agreement No. 817622.

Schoonman, M.J.L., Van Delden, A.N., Van Gennip, P., Van Dooremalen, C., Ülgezen Z. & de Graaf, D.C. (2019). Establishment of a professional **helpdesk** to assist beekeepers and researchers that experiencing difficulties during the execution of the field studies. **Deliverable D1.2** EU Horizon 2020 B-GOOD Project, Grant agreement No. 817622

van Dooremalen, C. & Ülgezen, Z.N. (2021). **Protocols No.2. Deliverable D1.6** EU Horizon 2020 B-GOOD Project, Grant agreement No. 817622.

van Dooremalen, C., Tom, J. & Ülgezen, Z.N. (2020). *Mini-apiaries*. *Milestone* **MS1** EU Horizon 2020 B-GOOD Project, Grant agreement No. 817622

van Dooremalen, C., Troost, J. & Ülgezen, Z.N. (2020). *Pilot & Field Study No.1.* Milestone MS2 EU Horizon 2020 B-GOOD Project, Grant agreement No. 817622.

van Dooremalen, C., Ülgezen, Z.N., Schoonman, M., van Gennip, P. & de Graaf D.C. (2020). *Protocols No.1*. **Deliverable D1.1** EU Horizon 2020 B-GOOD Project, Grant agreement No. 817622.