



## Open Research Data Pilot and project Data Management Plan

v02

### Deliverable D11.2

30 Sep 2020

*Iliyana Demirova, Pensoft Publishers*

*Margarita Grudova, Pensoft Publishers*

*Teodor Metodiev, Pensoft Publishers*

*Mark Brown, RHUL*

**PoshBee**

**Pan-european assessment, monitoring, and mitigation  
of stressors on the health of bees**



---

**Prepared under contract from the European Commission**

Grant agreement No. 773921

EU Horizon 2020 Research and Innovation action

Project acronym: **PoshBee**  
 Project full title: **Pan-european assessment, monitoring, and mitigation of stressors on the health of bees**  
 Start of the project: June 2018  
 Duration: 60 months  
 Project coordinator: Professor Mark Brown  
 Royal Holloway, University of London  
[www.poshbee.eu](http://www.poshbee.eu)

Deliverable title: Open Research Data Pilot and project Data Management  
 Deliverable n°: D11.2  
 Nature of the deliverable: Report  
 Dissemination level: Public

WP responsible: WP11  
 Lead beneficiary: RHUL

Citation: Demirova, I., Grudova, M. & Metodiev, T., Brown, M. J. F. (2020). *Open Research Data Pilot and project Data Management*. Deliverable D11.2 EU Horizon 2020 PoshBee Project, Grant agreement No. 773921.

Due date of deliverable: Month n° 36

Actual submission date: Month n° 25

Deliverable status:

---

Version	Status	Date	Author(s)
2.0	Revised	30 September 2020	RHUL, PENSOFT

---

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

Preface .....	5
Summary .....	5
Intended audience .....	6
Structure .....	6
1. Data summary .....	7
1.1. Purpose and type of data and relation to the objectives of the project .....	7
1.2. Types and formats of collected data.....	8
1.3. Data use .....	9
2. Applicable legislation, guidelines and principles .....	10
3. FAIR Data.....	10
3.1. Findable.....	10
3.2. Accessible.....	11
Deposition in open repositories.....	11
Publishing as data papers: .....	11
3.3. Interoperable .....	12
3.4. Reusable.....	12
4. Allocation of resources and responsibilities .....	13
5. Data security .....	13
5.1. Personal data .....	14
6. Ethics.....	15
7. Conclusions .....	15

## Preface

This document is the second version of PoshBee's data management plan (DMP) and Open Research Data Pilot (ORDP). The document describes the PoshBee consortium's data management life cycle for the data collected, generated and/or processed by the project, as well as outlining the project's adherence to the ORDP and open data publication policies and plans. Because PoshBee participates in the Open Research Data Pilot (ORDP), the project is obliged to develop a second version of the DMP. This DMP is based on the template for FAIR data management in Horizon 2020 (H2020), supplied by the European Commission (EC) (European Commission 2019). The DMP is a document that will be internally updated whenever significant changes in the project require revisions. This second version of the DMP has been developed in close collaboration with all consortium partners who provided information on their planned data collection, generation and processing as well as relevant institutional policies via surveys in June and September 2020.

## Summary

PoshBee's ORDP & DMP aims to address various aspects of data management and open data publication, including:

- What data will be collected, generated and/or processed by the project;
- Which standards will be applied;
- How and when will data will be shared/made open access;
- How data will be curated and preserved, including after the end of the project.

PoshBee has also opted in for the EC's ORDP, meaning that the project is committed to open science and FAIR and open data management. Article 29.3 of the grant agreement (GA) states:

Regarding the digital research data generated in the action ('data'), the beneficiaries must:

(a) deposit in a research data repository and take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate — free of charge for any user — the following:

- (i) the data, including associated metadata, needed to validate the results presented in scientific publications as soon as possible;
- (ii) other data, including associated metadata, as specified and within the deadlines laid down in the 'data management plan' (see Annex 1);

(b) provide information — via the repository — about tools and instruments at the disposal of the beneficiaries and necessary for validating the results (and — where possible — provide the tools and instruments themselves).

This document will clarify the ways in which PoshBee will adhere to the ORDP and how FAIR principles to data handling and its publication will be secured.

### **Intended audience**

The target readers of this document, besides the members of the PoshBee consortium involved in the production and preservation of (open) data, are all interested third parties that want to learn about the results of the PoshBee project in terms of (open) data.

### **Structure**

**Section 1** follows the structure of the Horizon 2020 DMP template and thus provides an overview of all data (i.e. research data, personal data and other data) PoshBee will collect, generate and process. The section includes information on the specific purposes of the various data and shows their relation to the objectives of the project. Furthermore, information on types, origins, formats, expected sizes and utilities of the data is provided.

**Section 2** expounds applicable legislation, guidelines and principles that data management in PoshBee will follow, namely the General Data Protection Regulation (GDPR) (European Union 2016), the H2020 Programme Guidelines on FAIR Data Management in Horizon 2020 (European Commission 2016), the H2020 Programme Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020 (European Commission 2017).

**Section 3** elaborates how PoshBee will make research data findable, accessible, interoperable and reusable (FAIR) as part of its participation in the ORDP.

**Section 4** outlines the allocation of resources, responsibilities and compliance.

**Section 5** explains how data security will be ensured and how personal data will be protected by PoshBee.

## 1. Data summary

This section explains the purpose of PoshBee collected data and its relation to the project's work programme. Furthermore, it outlines the different types of data generated and preferred formats. The section will also outline plans for data re-use.

### 1.1. Purpose and type of data and relation to the objectives of the project

To accomplish its objectives PoshBee will collect data, with the following specific purposes, mapped in relation to its work plan:

- **WP1 A site network for assessing exposure of bees to chemical, nutritional, and pathogen stressors:** Field survey data and field data from the collection of a range of samples (bee and bee product data, landscape data etc.), which will feed into the work of WPs 2 and 9; farmer surveys will collect information on the use of agrochemicals on the farms where primary field work was conducted, giving precise indications of the dosage and use of chemicals in the fields.
- **WP2 Measuring chemical exposure, pathogens and aspects of nutrition in honey bees, bumble bees and solitary bees:** data on the exposure of bees to various stressors, which will feed into a database on bee exposure to stressors that will be used to validate a new monitoring tool in WP2;
- **WP3 Toxicokinetics, toxicodynamics and interactions among agrochemicals:** data will be collected to determine the dose-response relationship of 3 classes of agrochemicals in bees; document the toxicokinetics of these chemicals in bees; evaluate synergies among the chemicals in bees; data will be used to determine how chemicals alone, in mixtures, and in combination with pathogens and nutrition, affect bee health.
- **WP4 Development of novel wild bee species for risk assessment:** Based on data collected PoshBee will suggest a new wild bee species for the use of risk assessment.
- **WP5 Effects of agrochemical-nutrition interactions on bee health in the laboratory:** Data will be collected to determine how chemicals alone, in mixtures, and in combination with nutrition, affect bee health.
- **WP6 Effects of agrochemical-pathogen interactions on bee health in the laboratory:** Data will be collected to determine how chemicals alone, in mixtures, and in combination with pathogens, affect bee health.
- **WP7 Effects of chemicals and their interactions with other stressors on bees tested in semi-field and field experiments:** Data will be collected from the semi-field experiments to determine how chemicals and their interactions with other stressors (nutrition, pathogens) affect the health of bees.

- **WP8 Systems and agent-based modelling approaches to assess the synergistic effects of multiple stressors on bee health:** Data will be used to measure hazard: drawing on the expertise of a diverse range of actors, PoshBee will quantify the exposure of honey bees, bumble bees, and solitary bees to chemicals within major agricultural cropping systems across Europe. Estimate health effects: taking a trans-disciplinary approach, we will integrate laboratory, semi-field, field, and landscape studies to provide a holistic understanding of how chemicals, their mixtures, and their interactions with pathogens and nutrition drive health in honey bees, bumble bees, and solitary bees. Develop a bee health model: within PoshBee we will develop the first mechanistically-underpinned holistic model of bee health.
- **WP9 OMICS of agrochemical responses in bees:** Data will be collected to develop a blood-based tool, “MALDI-Beotyping”, for the monitoring and assessment of bee health and exposure to stressors, and in order to develop a ‘health card’ for bees.
- **WP10 Knowledge Exchange and Impact Strategy:** Data collection in the form of grower and beekeeper surveys will be used for the provision of validated tools for the monitoring and assessment of bee health and exposure to stressors. The survey will identify barriers and incentive to the use of the tools developed by the project. Data will be used to inform policy and regulators.

## 1.2. Types and formats of collected data

For the purposes and objectives specified above and given the complexity of the different PoshBee experiments: field, semi-field and laboratory, various types of data will be collected during the project:

Type:	Use and rationale:	Format:
Observation data (Lab, field, semi-field)	Data on bee survival in experimental paradigms; data on bee pathogens in experimental paradigms (field and semi-field); survival, behaviour, physiology; field data on pesticide and pathogen load, and nutritional status with respect to different measures of bee health in combination with landscape and climatic data; observational data such as counts of mortality, assessments of bee behaviour, measurements of crop growth. Automatic recording of ambient conditions by weather station; chemical data (protein, lipid), wing shape data, fat body data.	.xlsx; .docx; .pdf; GIS shape files; .RData; .shp, .tiff; .jpg
Experimental data (Lab, semi-field, field)	From experiments where bumblebees, honeybees & solitary bees are exposed to agrochemicals, pathogens, and nutritional stress, and outcomes are measured; concentrations of pesticide residues in beebread/pollen store samples; proteomics data derived by the analysis of the sampled bees’ blood; data generated from the analysis of the received	.xlsx; .docx



	samples (lab, semi-field and field) will be collected in the form of mass spectra, processed by proteomics-dedicated bio informatics tools to obtain list of proteins and of modulated proteins under different stress conditions.	
Survey data	Collected from participating farmers only (WP1); collected from a snowball sample of growers and beekeepers (WP10)	.xlsx; Qualtrics;
Literature review	From horizon scanning (WP10)	.docx; .xlsx
Field survey data	Bees, flowers, land use, bee products from managed hives (WP1)	GIS shape files; text data files; .RData; .shp; .tiff
Interviews	Collected from participating farmers	.docx, Qualtrics
Photographic records	Photographic records of honeybee brood, and subsequent analyses by image recognition systems.	.tiff, .jpg
Historical data	Historical data on chemical applications from farm records.	.docs, .xlsx, .pdf
Protocols	Developed in WP1, WP9; videos of most important protocols for training purposes.	.docs, .mp4

Due to the complexity of field and semi-field data collection and the different data types collected in field, semi-field and lab experiments, it is hard to predefine the exact size of the collected datasets. More space consuming data types are the images collected on the field and in experiments and the .mp4 files for the video protocols. When there is a difficulty in storing data internally (within the respective institution), the website of PoshBee offers a login protected file repository that can be used for this purpose, and also for the purpose of sharing data within the consortium.

### 1.3. Data use

Beyond the project, the research data collected and generated by PoshBee will be useful for further research on bee health and the effects of nutrition, pathogens and agrochemicals on honey bees, solitary bees and bumblebees. As the recommended data format for research data and teaching and campaign materials will be openly licenced under the Open Data Commons Attribution License (ODC-By) or equivalents, interested researchers will be able to use PoshBee's outputs. To enhance dissemination an open science collection of all PoshBee Data Papers will be published in the Research Ideas and Outcomes Journal.

All data resulting from PoshBee will be made openly available except for the following identified cases during the survey conducted for the purposes of this DMP:

- Data for mapping will not be made openly available but will be used to create landscape models

- Farmer interviews and georeferenced data for landscape structure and habitat composition cannot be shared since they can be linked to data about field management and thus are not in agreement with the GDPR.

## 2. Applicable legislation, guidelines and principles

PoshBee will adhere to the following legislation, guidelines and principles in the management of its data collection, generation, processing, sharing and preservation:

- European Parliament and the Council, Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, repealing Directive 95/46/EC (General Data Protection Regulation), 27 April 2016
- European Commission Directorate-General for Research and Innovation, H2020 Programme Guidelines on FAIR Data Management in Horizon 2020, Version 3.0, 26 July 2016
- European Commission, H2020 Programme Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020, Version 3.2, 21 March 2017

## 3. FAIR Data

Participating in the ORDP, PoshBee is committed to publish all scientific articles the project will produce as well as the underlying research data in open access mode, unless legal requirements impose binding restrictions. In particular, PoshBee will apply the FAIR guiding principles for scientific data management according to which research data shall be findable (F), accessible (A), interoperable (I) and reusable (R).

### 3.1. Findable

To comply with the principles of open data, partners will deposit data in suitable data repositories or publishing data papers.

Project datasets will follow unified name conventions, which will be agreed upon prior to depositing data. An example of a name convention is [PoshBee\_dataset\_name\_version\_creation date].

The following list contains a recommendation of the minimum characteristics of how metadata should be described:

- Author(s)
- Year

- Dataset title
- Data repository or archive
- Global persistent identifier
- Version or subset and/or access date
- Language
- Metadata language
- Licence of use
- Date of metadata creation
- Hierarchy level
- Character encoding
- Format version

Since mostly numerical data will be created search keywords will not be requested as a compulsory element of metadata, however this field will be highly recommended for other data types.

### 3.2. Accessible

#### Deposition in open repositories

A data repository is a digital archive collecting and displaying datasets and their metadata. PoshBee will deposit research data resulting from project research and underlying scientific publications in online repositories, in order to cluster data together as well as to preserve and archive it.

Repositories used by the PoshBee consortium are:

**Dryad:** One possibility is publishing datasets in the non-profit repository for data underlying the international scientific and medical literature, namely Dryad.

**Zenodo:** An alternative to Dryad is Zenodo – a research data repository launched in 2013 by the EU-funded OpenAIRE project and CERN to provide a place for researchers to deposit datasets of up to 50 GB in any subject area.

**Other repositories** will be possibly considered in cases where high level journals request open data deposition in specifically selected resource.

**PoshBee Library and institutional servers** will be used for internal storage and sharing of raw data before it is ready for publication.

**Publishing as data papers:**

A data paper is a scholarly journal publication whose primary purpose is to describe a dataset or a group of datasets, rather than report a research investigation. As such, it contains facts about data, rather than hypotheses and arguments in support of those hypotheses based upon data, as found in a conventional research article.

To ensure that all data published by PoshBee is featured in a single, easily findable, and accessible collection, all datasets published by the project in various repositories will be described within the Research Ideas and Outcomes (RIO) Journal in the form of data papers, which will provide a single, citable point of access to all datasets made openly available by PoshBee.

### 3.3. Interoperable

PoshBee will ensure interoperability of its research data by using common file formats specified in the data summary (e.g. docx, xlsx, pdf, .jpg, .mp4, .shp). This facilitates recombination of the data with other datasets from different origins. Although the data formats mostly originate in or are primarily used with proprietary software, they are also accessible with open source software.

As PoshBee partners have different disciplinary backgrounds and as the project addresses various academic disciplines, the principle of using extant data and metadata vocabularies, standards and methodologies will be applicable for each field.

### 3.4. Reusable

All research data needed to validate results presented in scientific publications produced by PoshBee will be made openly accessible, unless restrictions apply due to data protection policies (also see information in section 1.3). Hence, all data that are potentially traceable to an individual are exempt from open access for both legal and ethical reasons.

As PoshBee will follow the gold route to open access, where possible, and the green route otherwise, all scientific publications emanating from the project will be published in open access mode immediately by either the respective journal or the author on their institutional or personal website. In parallel, the research data underlying publications will be made openly accessible as soon as possible. Whenever possible, research data as well as the associated metadata produced by PoshBee will be deposited in certified open access repositories suitable for the particular data (e.g. Zenodo) (see section 3.1). Data will be licenced using the Open Data Commons Attribution License ODC-By. Authors should explicitly inform the Project Coordinator and Project Management Group, if they want to publish data associated with a journal article under a license that is different from the **Open Data Commons Attribution License (ODC-By)**.

Other open data licences:

**Creative Commons CC0** (also cited as “CC-Zero” or “CC-zero”) and the **Open Data Commons Public Domain Dedication and License** (ODC-PDDL). According to the CC0 license, “the person who associated a work with this deed has dedicated the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and neighbouring rights, to the extent allowed by law. You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission.” Publication of data under a non-attribution waiver such as CC0 avoids potential problems of “attribution stacking” when data from several sources are aggregated for re-use, particularly if this re-use is undertaken automatically. In such cases, while there is no legal requirement to provide attribution to the data creators, the norms of academic citation best practice for fair use still apply, and those who re-use the data should reference the data source, as they would reference other research articles.

**The Attribution-ShareAlike Open Data Commons Open Database License (OdbL)** is NOT recommended for use for POSHBEE data, although it may be used as an exception in particular cases. The OdbL license assumes that "If one publicly uses any adapted version of the database, or works produced from an adapted database, he or she must also offer that adapted database under the OdbL."

#### 4. Allocation of resources and responsibilities

Opening the research data is associated with various types of costs. They can be grouped into two main categories: 1) article processing charges (APC) for publishing data in scholarly journals; 2) fees for depositing data in global data repositories. While many data repositories are free of charge to authors to upload their data (e.g. Zenodo, GBIF), some more generic data repositories such as Dryad charge users 120 USD for publishing their data, unless exemptions apply. Where costs have been allocated in the Grant Agreement to pay for open access, this will facilitate payment of APCs or data deposition. Where these costs have not been allocated, beneficiaries will either use green open access, free data repositories, or fund these costs at no cost to the project.

Each beneficiary has identified to the PoshBee consortium persons responsible for their obligations under this Data Management Plan. Ultimately, each beneficiary will be responsible to the PoshBee Coordinator, which holds ultimate responsibility for Data Management across the project.

#### 5. Data security

In general, data collected from the website, ICP and LimeSurvey is stored on Pensoft's servers, located in Sofia, Bulgaria, EU. This includes document files, images and audio, and any data stored by partners via the projects repository. The data storage infrastructure enables a reliable backup process. Data is duplicated on a second server to allow retrieval of datasets in case of data loss, caused by hardware or software failure, data corruption or a human-caused event, such as a malicious attack (virus or malware), or accidental deletion of data. Backup copies allow data to be restored from an earlier point in time. To meet compliance requirements, data is securely encrypted. Data backup occurs on a daily basis to ensure that no information is irreversibly lost or deleted. If consortium partners process and/or store data at other repositories, e.g. institutional ones, they will ensure that equivalent data security measures are implemented.

### 5.1. Personal data

PoshBee will take all necessary measures to adequately safeguard data security and ensure the protection of personal datasets. All personal data collection and processing will adhere to the GDPR as well as other applicable legislation and shall be based on informed consent/assent. Individuals can withdraw their consent to data processing at any time and without having to suffer any negative consequences. They will be informed of their rights and be provided with contact information of responsible data managers and DPOs (data protection officer) before data are collected.

With regard to personal data controlled by PoshBee, individuals (data subjects) have the following rights:

- Right to be informed: Individuals shall be informed in intelligible language how their personal data will be used.
- Right of access: Individuals shall be enabled to access their personal data and attendant supplementary information.
- Right to rectification: Individuals can request to rectify or amend their personal data if it is inaccurate or incomplete.
- Right to erasure: Individuals can request to have their personal data deleted or removed.
- Right to restrict processing: Individuals can request to restrict future processing of their personal data or object to it entirely, i.e. they can withdraw or alter their consent to data processing. This does not affect the lawfulness of the prior data processing.
- Right to data portability: Individuals can request data controllers to provide them with their data in a commonly used, machine readable format.
- Right to lodge a complaint with a supervisory authority.
- Data will be collected on a need to know basis only. No excess data will be generated.

- Data sets will not be merged, so as to avoid any unintended disclosure of personal information.

The consortium member responsible for data collection/generation and storage acts as controller of the respective data and is responsible for compliance with all pertinent legislation.

That said, processing of personal data will be limited to the extent necessary to accomplish a project objective. Personal data will only be used for the purposes it was collected.

## **6. Ethics**

PoshBee is committed to and will comply with Article 34 of the GA, stating that all beneficiaries must carry out the action in compliance with ethical principles and applicable international, EU and national law. Further work on ethical issues available in D13.1 – 13.4. No further particular ethical considerations are applicable to the project.

## **7. Conclusions**

This document will be shared and communicated within the PoshBee consortium, to ensure that all beneficiaries are enabled to comply with the DMP throughout the lifetime of PoshBee, and beyond. Where necessary, as the field of open access evolves, we will update our DMP and guidance