



## D.9.1 DATA MANAGEMENT PLAN Draft

Contributing Authors:

Jorrit H. Poelen, <https://orcid.org/0000-0003-3138-4118>

Mirella Miettinen, <https://orcid.org/0000-0003-3593-3938>

Laura Drivdal

Ugo Mendes Diniz

Nicholas James Balfour

Maria Clara Castellanos

Jeff Ollerton


Lorraine Balaine

Nicola Gallai

Jeroen van der Sluijs

**DISCLAIMER: This is a work in progress document and input is needed!!**

*(To be filled in and uploaded as deliverable in the Portal Grant Management System, at the due date foreseen in the system (and regularly updated).)*

 *The template is recommended but not mandatory. If you do not use it, please make however sure that you comply with the research data management requirements under Article 17 of the Grant Agreement.)*

PROJECT	
Project number:	101181930
Project acronym:	Butterfly
Project name:	Mainstreaming pollinator stewardship in view of cascading ecological, societal and economic impacts of pollinator decline.

DATA MANAGEMENT PLAN	
Date:	[dd/mm/yyyy]
Version:	[1th DMP version]

## Innhold

### 1 Introduction

2

2

2. Data	3
2.1 Data Summary	3
2.2 RE-use of data (see also table 2)	4
2.3 Purpose of data generation in relation to objectives	5
2.4 Size and origin of data	5
2.5 Usefulness of data outside the project	7
3. FAIR data management	8
3.1 Making data findable, including provisions for metadata	8
3.2 Making data accessible	9
3.3 Making data interoperable	10
3.4 Increase data re-use	10
4. Other research outputs	11
5. Allocation of resources	11
6. Data security	12
7. Ethics	12
References	14
Appendix A: General informed consent form	15

## 1 Introduction

This document describes the Data Management Plan (DMP) for project Butterfly. The DMP is a 'living' document that will be reviewed and updated over the course of the project. Additionally, we aim to archive versioned snapshots of this document, specifically in the months 24 and 48. As Butterfly is working closely with related projects, we aim to align, reuse, or take inspiration, from the DMPs of our sister project VALOR (Breeze et al 2025; Breeze and Zurgic, 2025), and projects that we closely collaborate with: SAFEGUARD (Zhang and Steffan-Dewenter 2022) and RestPoll (Wintermantel et.al 2024).

This document provides an overview of data products that will be generated (see "Data Index"). In addition, we describe how we aim to use community best practices (e.g., FAIR principles) to facilitate the re-use of this data within, and beyond, the scope of this project ("Data Re-use") . Finally, we discuss our ethical handling, and re-use of personal data (section "Data Ethics").

Butterfly is a transdisciplinary project, generating different types of data (both quantitative and qualitative) across the various Work Packages and tasks. In section 2 our data overview lists all generated, re-used and curated data products, an important first step to enable data review, referencing and cross-referencing.

In section 3 we describe how the project will adhere to FAIR principles – making the data Findable, Accessible, Inter-operable, and Re-usable. Attention will be given to what kind of metadata the outputs should contain and how and where they will be stored. Furthermore, Butterfly is designed to ensure accessible and re-usable data through: a) the EuroAPPA portal which aims to provide user-friendly access for all stakeholders to the most complete, taxonomically-harmonised and well-curated database of plant-pollinator interactions for Europe and three Overseas Territories/Outermost Regions; and b) the project's co-creation approach in the Living Labs, facilitating 'openness by design', meaning that data creation is a shared venture from the start

(Mačiulien, 2022).

Section 4 discusses the generation of research outputs other than data, and how we have considered the allocation of project resources, approaches to data security, and ethical considerations that arise from these outputs.

Section 5 is a consideration of how we allocate project resources to data management, including the time and financial costs of making data or other research outputs, how this will be funded, and where responsibilities lie.

In section 6 we summarise our approach to the security of data as well as non-data outputs, such as plans for storage, long-term archiving, and recovery of data, plus how sensitive data will be transferred.

Finally, in section 7, ethical issues around the treatment of personal data will be discussed. Project Butterfly is committed to making the data 'as open as possible, as limited as necessary'. Thus, attention will be given to how we deal with data that contains personal information, and how we ensure that all participants are giving informed consent to participate. A general informed consent sheet that can be adapted to each data collection activity in the different countries is included in the appendix.

**Suggestion for consortium:** Although interaction between disciplines is central in Butterfly, adherence to FAIR principles and ethics issues will be overseen by expertise in specific disciplines (**data stewardship**). Thus, two teams will be established as contact points:

- 1) Team for overseeing FAIR data management of ecological data, consisting of the EuroAPPA team (Jorrit, Claus, Jeff, Joseph, Sara, Cala and Nick Balfour), Lead of **T1.4**, and Noa Simon, Beelife, co-lead **WP8**,

Team for discussing and overseeing human participant data – attending to FAIR measures AND informed consent for participants, including: George Vrontos, UTH, **T2.2** lead - MoU for the LLs and **T8.4 + T8.5** co-lead, Nicola Gallai, ENEFSA, lead **T2.3** pop-survey and **T8.5**, Martin WP6 Georgious, CIHEAM, lead data repository **WP2** and educational part **T8.5** Mirella Miettinen, UEF, **WP4** May-Brith, UIA, **T6.3** leader and/or Katharina Schwarz and AyeAye UT, T6.3 lead on psychological experiments?. Laura Drivdal, UiB, **WP8** and **T9.3** - DMP,

transdisciplinary review

Suggestion: Meetings 4 times a year - decide on process and outline responsibilities

Butterfly Data Registry

dataset	summary	contributors	discipline	review dates/notes
(short name link to dataset page)	(single paragraph)	(contact info)		

## 2. Data

### 2.1 Data Summary

#### *What types and formats of data will the project generate or re-use?*

Project Butterfly is fundamentally interdisciplinary / transdisciplinary, where many different types of data will be compiled and generated across various task. This includes collecting, generating and reusing a) **ecological data** b) **biological data**, c) **human participant data**, and d) **economic & production data**. These three types of data carry different implications for data sharing and ethical approvals, and this will be outlined systematically throughout this document.

- a) **Ecological data:** the interactions between species of pollinators and the plants whose flowers they visit. This plant- pollinator interaction data will be geo-located using decimal coordinates and elevation, and time-stamped by the day/month/year in which it was collected. In addition, each interaction will be assigned a measure of data 'quality' *sensu* Ollerton et al. (2025).
- b) **Biological data:** this includes DNA sequence data from pollen samples collected during the Living Labs and Overseas Sites field work.

**Ecological and biological data will mostly be collected and managed through WP1.** The EuroAPPA “one-stop shop” web portal and database for information on plant-pollinator associations across Europe will include **both re-used data and new data**. T1.2.1. will review and reuse data from existing databases and mobilise data that is not yet indexed in order to provide a Europe-wide synthesis of information on plant-pollinator databases. In T1.2.2, in collaboration with the LL's (WP7), **field-based data** will be collected on pollinators, plant-pollinator interactions, pollinator services and pollinator dependencies. The Field Protocol (WP1) (Diniz et al, 2025), specifies procedures for sampling, sampling efforts required, and standards for labelling physical samples and digital sample data.

- c) **Human participant data, quantitative and qualitative:** Participants perceptions, values, belief and opinions, arising from surveys, interviews, workshops, field observations (of human behaviours?), engagement activities, stakeholder input, co-creation and psychological experiments

**Human participant data will be collected and managed in several WPs and tasks.**

**T2.3, together w. T6.3 will collect Citizen perception data** on a) willingness to pay for the preservation of pollinators a national scale through a large-scale **population-based survey** in at least six European partner countries, and b) and **qualitative workshops** with students specializing in agricultural education. Indeed, in addition to the national survey, we will specifically assess the willingness to pay of young people—considered as future key actors in rural territories. This targeted component will be implemented in France, Greece, and Ireland within agricultural training programs. Questionnaires administered to these students will be complemented by participatory workshops designed to elicit the multiple values they associate with pollinators. These workshops will also help us analyze their willingness to engage collectively in pollinator conservation efforts.

**T4.1 and T4.4 will collect data through** five sector-specific Delphi surveys. Quantitative and qualitative data will be generated through an one-round, online real-time questionnaire, targeting 15 experts per sector (food/micronutrients, pharmaceuticals, cosmetics, biomaterials and biomass energy). The Delphi surveys will be used to assess how vulnerable global supply chains in different sectors are to pollinator loss and how prepared actors in each sector are to the risks of pollinator loss. Commercial software (such as Surveylet) will be used in conducting the surveys, the software will be compared, and a suitable one will be selected based on calls for tender in early 2026. Data collected in Task 4.1 consists of name, email address, area of expertise, and country of the potential experts to be invited to complete the Delphi online questionnaire in each sector. The Delphi online survey can only be accessed via email, so the data collected is needed to complete

the surveys. The experts will be identified by utilising the existing networks of Butterfly partners, stakeholders in the Butterfly living labs, and literature review conducted in the first step of the Delphi survey procedure. The contacted experts can also be asked to nominate other experts to be included (snowballing)

Data generated in Task 4.4 consists of i) raw data sets from the five online questionnaires, ii) automatic reports delivered by the survey software (such as log reports for Milestone 11 or basic statistics and visualisation of the completed responses in the platform), iii) five reports (one per each sector, D4.3a-e) of the results of Delphi surveys, including identified resilience options for each sector and guidelines on how to reverse pollinator decline and mitigate future risks. Raw data sets contain both quantitative (such as % scale and 9-point Likert scale) and qualitative data (arguments). The anonymity of individual respondents will be ensured, and quantitative results will be presented in aggregated form. Qualitative data (arguments) can be quoted, but the identity of the author will not be used.

**In WP 6, several tasks involve gathering human participant data. T6.3, together with T2.3, data on general public's basic knowledge** of pollinators, the ecosystem services they provide and their importance for nature and humans, through a **pan-EU (incl. overseas) survey**.

**Attitude-behaviour data will be generated from experimental psychology set-ups:** the testing strategy will first involve controlled psychological experiments. The results of the controlled experiments will be used to develop tools that transfer knowledge about pollinators to everyday citizens (details in Table 1).

**Field observations and newly developed paradigms** testing cognitive and social motivating strategies, skill sets and capacities, for furthering pollinator stewardship, eco literacy, historical agency and slow hope among everyday citizens.

**T7.1** together with **T6.3**: Data will be generated from the BUTTERFLY LLs through structured dialogues between stakeholders and participants, field observations of individual practices regarding pollinators and social and environmental interventions, **citizen science participation, interviews, photography, video, document analysis**, and other material produced in the different LL processes. The analysis will provide insights into how stakeholders and individuals produce, use and share knowledge about pollinators.

**T7.3** Participatory scenario planning and co-creation (?)

#### **d) Economic and land/agricultural production data**

**T2.1 w. WP7 (LLs):** Data on agricultural activities, value chains and other economic related to pollinators and ecosystems. Existing digital data sets of eg predicted climate (Eyring et al 2016) and land use and cover (Hoffmann et al 2023) under each of the Shared Socio-economic Pathways (IPCC 2023) will be re-used. In addition, data will be collected by each LL-leader through surveys, remote sensing, land uses, pollinator stress, and more. **T2.1**, asks each LL to collect data related to farm structures, key practices, sustainability measures, and market access. (data collection finished end of dec 2025. For the general survey, data collected include location, years of operation, farm size, key agricultural activities, annual yield, practises such as organic, integrated pest management, water conservation etc), types of fertilisers and pesticides, Value chain etc.

## **2.2 RE-use of data (see also table 2)**

*Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.*

Ecological, economic, legal and policy data will be collected from existing sources and re-used, specifically in WP1, WP4 and WP6.

- **T1.2.1** ("synthesis and mobilization of data sources"), will index and disseminate information contained within **existing databases of biotic interactions** (plant-pollinator

networks). Specifically building on the Database of Pollinator Interactions – DoPI (Balfour et al. 2022, currently focused on the UK), and the Global Biotic Interactions platform – GloBI (Poelen et al. 2014). This is significant for producing a one-stop shop” (PS3) for plant-pollinator interactions. Further, data sources that are currently not indexed by the sources will be mobilized to provide a Europe-wide synthesis of information on plant-pollinator databases with a particular view to targeting lesser-known groups of pollinators.

- **T1.4.1:** Information from existing digital products on the distribution of managed plants and pollinators (such as the EU Crop Map and the Eurostat dataset on main livestock indicators)
- **T1.4.2: digital data sets of predicted climate** (Eyring et al. 2016) **and land use and cover** (Hoffmann et al. 2023) under each of the Shared Socio-economic Pathways (IPCC 2023) to produce estimates of plant-pollinator network structure across Europe under various scenarios of human development for 2050/2100
- **T1.4.3:** carry out a systematic review and meta-analysis of the state-of-the-art field experiments in which plant and/or pollinator diversity and/or abundance have been manipulated and an assessment made of the impact on plant-pollinator networks
- **T4.1:** Existing data that will be utilised in step 1 of a Delphi survey protocol includes scientific literature on potential vulnerabilities and tipping points in each supply chain to pollinator loss and possible response options (mitigation and adaptation). Literature search can be complemented by searches on relevant databases (may be different for each supply chain) and other desk research (such as company websites, previous EU or national projects etc.)
- **T6.1:** Collect and re-use existing data on human dimensions of pollinator decline from academic literature, including grey literature.
- 
- **T6.2:** In Task 6.2, openly accessible EU legal documents and case law (e.g. from EUR-Lex (<https://eur-lex.europa.eu/homepage.html>) or Court of Justice of the European Union ([https://curia.europa.eu/jcms/jcms/j\\_6/en/](https://curia.europa.eu/jcms/jcms/j_6/en/))), as well as national CAP Strategic Plans of selected countries (links available: [https://agriculture.ec.europa.eu/cap-my-country/cap-strategic-plans\\_en](https://agriculture.ec.europa.eu/cap-my-country/cap-strategic-plans_en)) will be collected and analysed. The Legal analysis done in **Task 6.2** will produce i) a report (D6.2) that includes the results of the analysis on how pollinator conservation has been considered in key pieces of EU legislation, EU's common agricultural policy (CAP), and national agricultural policy (e.g. in Norway), and how new Regulation (EU) 2024/1991 on nature restoration affects the provisions in these existing instruments; and ii) a report (D6.3) that includes the results of the assessment and recommendations on how the selected instruments developed in WP5 could be integrated into existing legal or regulatory instruments (such as CAP).
- 
- **T6.4** Historical analysis attempts to systematically recapture the complex nuances, the people, meanings, events, and ideas of the past that have influenced and shaped the present. It relies on a wide variety of sources, both primary & secondary including unpublished material. Primary sources can be found in public records & legal documents, minutes of meetings, corporate records, recordings, letters, diaries, journals, drawings, located in university archives, libraries or privately run collections such as local historical society. Secondary sources can be found in textbooks, encyclopaedias, journal articles, newspapers, biographies and other media. The data and insights from these analyses will then be built together in an overall coherent analysis and synthesis presenting the human, social and historical (past-present-future) aspect of pollinator loss and restoration on micro, meso and macro level of society.

## 2.3 Purpose of data generation in relation to objectives

*What is the purpose of the data generation or re-use and its relation to the objectives of the project?*

Butterfly has eight specific objectives (SO) presented in Table 1.1 of The DoA. Data generation is specifically relevant in relation to:

*SO1: Provide a holistic overview of actionable knowledge on animal pollination ecology and pollination services provided for wild and cultivated plants covering the European continent as well as EU overseas territories.*

- For this, it is essential to provide and analyse biological/ecological data on plant-pollinator interaction.

*SO3: To comprehensively model and quantify the macro- economic implications of pollinator decline, to model the country-specific economic butterfly effects of dependencies on pollinators, and to provide forward- looking analysis of policy options and scenarios.*

- It is vital to collect and analyse economic data for the modelling

*SO4: Understand how 5 key biomass supply chains (food/micronutrients, pharmaceuticals, cosmetics, biomaterials, biomass energy) depend on pollination and co-create pollinator restoration options that increase resilience of these supply chains. Promulgate resilience-thinking to businesses beyond BUTTERFLY stakeholders and to EU policymakers.*

- To assess potential vulnerabilities and tipping points in each supply chain to pollinator loss and possible response options (mitigation and adaptation), data from experts is generated through five sector-specific Delphi surveys.

*SO5: Develop, test and implement **transferable tools** that enable **systematic mainstreaming of proactive pollinator stewardship into key vulnerable sectors** through multi-actor co-creation approaches and LLs*

*And SO7: Establishing a test-system of multi-actor communities across sectors to **accelerate knowledge transfer** and serve as field study sites, **multi-actor co-creation of knowledge and solutions**, and forum for continuous discussion and networking.*

- The multi-actor dialogues and the co-creation approach implies the collection of feedback and data in workshops and seminars

## 2.4 Size and origin of data

*What is the expected size of the data that you intend to generate or re-use?*

The size of the biological/ecological data will be particularly extensive. Current assumption **for data volumes would be >10TB (?)**

*What is the origin/provenance of the data, either generated or re-used?*

Table 1: Types and origins of 'primary data' that will be collected

Data type	Place/sector of collection	WP
Field-based data on pollinators, plant-pollinator interactions, pollinator services and pollinator dependencies. Specifically, the baseline data will have the following origins:	All B-Sites, which include the Living Labs (LL) - Region of Murcia (ES), Zeeland (NL), Northern Jutland (DK), Ile-de-France (IDF) (FR), Southern Norway (NO), Milano region (IT) -, as well as non-LL sites, such as the	WP1 w. WP7



<ul style="list-style-type: none"> <li>- 10 pollen samples from every B-/S-pollinator species in each site. Interaction data between plants and pollinators will be obtained via pollen DNA metabarcoding;</li> <li>- 100 flower heads from each B-/S-/I-Plant species. Interaction data will be obtained via eDNA metabarcoding.</li> <li>- 120 minutes of Flower-Insect Timed Counts for each B-/S-/I-Plant species, providing visitation and plant fitness (dependency) data.</li> </ul>	RestPoll Sites and overseas territories (Greenland, Curaçao, and Martinique).	
<b>Qualitative workshops</b> with students specializing in agricultural education. 100 students targeted per country	France, Ireland and Greece with the support of the ENTER Network	WP2
<b>Survey on willingness to pay</b> at a national scale	at least six European partner countries (France, Greece, Ireland, Germany, Norway and Italy) ( <b>PS4</b> ).	WP2
<b>Data on agricultural activities, value chains and other economic related to pollinators</b> and ecosystems.	<b>Each LL-leader</b> collects accurate local data through surveys, remote sensing, land uses, pollinator stress, and more.	<b>WP2</b> w. WP7
Data generated in Task 4.4 consists of i) raw data sets from five online Delphi questionnaires, ii) automatic reports delivered by the survey software (such as log reports or basic statistics and visualisation of the completed responses in the platform). Raw data sets contain both quantitative (such as % scale and 9-point Likert scale) and qualitative data (arguments).	International - 15 most relevant experts in each sector: supply chains for food/micronutrients, pharmaceuticals, cosmetics, biomaterials and biomass energy	WP4
Co-creation workshops in WP7 ( <b>creating what kind of data?</b> )	Region of Murcia (ES), Zeeland (NL), Northern Jutland (DK), Ile-de-France (IDF) (FR), Southern Norway (NO), Milano region (IT),	WP7
Experimental data from controlled psychological experiments will be used to determine how attitudes toward environmental issues and pollinators, social parameters, and training influence how well people adopt and retain pollinator-friendly behaviors. Recruited participants (sample size tbd) will follow a computer-based program with the goal of testing strategies to facilitate pollinator	Students and employees of the participating universities (Trier University and TUM), citizens from the local municipalities, as well as participants found via online platforms such as Prolific.	WP6

<p>stewardship (e.g., managing a garden in a computer program with the aim of maximizing pollinator numbers and diversity).</p> <p>Data type:</p> <ul style="list-style-type: none"> <li>- Personal data (demographics: age, gender, living space, region of residence, etc.)</li> <li>- Experimental data (attitude evaluations, participants' ad hoc knowledge of pollinator-plant interactions, participants' choices within the experimental paradigm, choice outcome, response times, questionnaires, etc.)</li> <li>- To evaluate transfer of learned plant-pollinator interactions to real life, participants may document their efforts via photographs that will have all identifying information removed (blackened, location data and time stamp removed, etc.); photographs will not be shared with others, but will be encoded by several involved experimenters into a quantifiable measure.</li> </ul> <p>To protect the personal data of participants, all information collected from them will be first pseudonymized until data collection is complete and then anonymized. Only anonymized data will be shared between partners outside of participating partners (i.e., UT and TUM). All experimental procedures will be checked by a local ethics board for compliance with European, national, and local standards.</p>		
---	--	--

Table 2: Types and origins of 'publicly available data that will be re-used'

Data Type	Source and link	WP
Ecological baseline data on published plant-pollinator networks to be added to The Database of Pollinator Interactions (DoPI)	<a href="https://www.sussex.ac.uk/lifesci/ebe/dopi/">https://www.sussex.ac.uk/lifesci/ebe/dopi/</a>	WP1
Ecological baseline data from Global Biotic Interactions (GloBI)	<a href="https://www.globalbioticinteractions.org/">https://www.globalbioticinteractions.org/</a>	WP1
Data not currently indexed by existing databases on all pollinators but with special emphasis on <b>lesser-known pollinators</b> such as birds, bats, and other insects than bees, wasps, or syrphid flies	<p><b>Suggestions:</b></p> <p><b>Bats:</b> <a href="https://www.batbase.org/">https://www.batbase.org/</a></p> <p><b>Birds, bats, and many other pollination networks (including "lesser known" insects):</b></p>	WP1

	<a href="http://www.ecologia.ib.usp.br/iwdb/resources.html">http://www.ecologia.ib.usp.br/iwdb/resources.html</a>	
EU Crop Map and the Eurostat dataset on main livestock indicators	<a href="https://ec.europa.eu/eurostat/web/agriculture/database">https://ec.europa.eu/eurostat/web/agriculture/database</a>	
Digital data sets of predicted climate (Eyring et al. 2016)	<a href="https://gmd.copernicus.org/articles/9/1937/2016/">https://gmd.copernicus.org/articles/9/1937/2016/</a>	WP1
Digital data sets of land use and cover (Hoffmann et al., 2023)	<a href="https://essd.copernicus.org/articles/15/3819/2023/">https://essd.copernicus.org/articles/15/3819/2023/</a>	
Shared Socio-economic Pathways (IPCC 2023) Future Global Climate: Scenario-based Projections and Near-term Information	<a href="https://www.cambridge.org/core/books/climate-change-2021-the-physical-science-basis/future-global-climate-scenario-based-projections-and-near-term-information/309359EDDCFABB031C078AE20CEE04FD">https://www.cambridge.org/core/books/climate-change-2021-the-physical-science-basis/future-global-climate-scenario-based-projections-and-near-term-information/309359EDDCFABB031C078AE20CEE04FD</a>	W
Scientific literature on potential vulnerabilities and tipping points in each supply chain to pollinator loss and possible response options (mitigation and adaptation). Literature search can be complemented by searches on relevant databases (may be different for each supply chain) and other desk research (such as company websites, previous EU or national projects etc.)		WP4
Literature review on human dimensions		WP6, T6.1
EU legal documents and case law.  National CAP Strategic Plans of selected countries.	EUR-Lex ( <a href="https://eur-lex.europa.eu/homepage.html">https://eur-lex.europa.eu/homepage.html</a> ) Court of Justice of the European Union ( <a href="https://curia.europa.eu/jcms/jcms/j_6/en/">https://curia.europa.eu/jcms/jcms/j_6/en/</a> )  Links available here: <a href="https://agriculture.ec.europa.eu/cap-my-country/cap-strategic-plans_en">https://agriculture.ec.europa.eu/cap-my-country/cap-strategic-plans_en</a>	WP6, T6.2
<b>Data for a historical (past-present-future) meta-analysis</b> on human and social determinants and consequences of pollinator loss and restoration covering the <b>period 1850-2050</b> (	Primary sources in public records & legal documents, minutes of meetings, corporate records, recordings, letters, diaries, journals, drawings, located in university archives, libraries or privately run collections such as local historical society. Secondary sources found in textbooks, encyclopaedias, journal articles,	WP6, T6.4

	newspapers, biographies and other media	
--	---	--

## 2.5 Usefulness of data outside the project

*To whom might your data be useful ('data utility'), outside your project?*

BUTTERFLY researchers share new knowledge and data with relevant actors as early in the research process as possible to ensure beneficiaries, particularly the at-risk sectors, benefit from outcomes and learning as it emerges. Beyond BUTTERFLY, the consortium will actively share data and outputs with other initiatives (§1.2.2), including EU-funded projects. We acknowledge that effective knowledge exchange between initiatives can avoid duplicative or unnecessarily competing efforts and instead foster a collaborative culture of effective pollinator restoration research for impact.

Data will also be used to shape the WP5 Decision support tools, maps, and guidelines. A key feature of these tools is that they will inform stakeholders about the risks of pollinator loss for their businesses (data from WP2 and WP3) and assess the impact of the measures on pollinators (task 5.2, task 5.4), within the framework of a global conservation strategy (data from WP1).

Data within the EuroAPPA database will have long-term utility for ecologists, agriculturalists, policy makers, and others interested in the biodiversity of plant-pollinator interactions in Europe and some of the Overseas Territories/Outlying Regions.

## 3. FAIR data management

### 3.1 Making data findable, including provisions for metadata

*Will data be identified by a persistent identifier?*

*Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.*

*Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?*

*Will metadata be offered in such a way that it can be harvested and indexed?*

Every dataset will have a **persistent and unique identifier** throughout the entire project. Depositing datasets in Zenodo will automatically give them a Digital Object Identifier (DOI) for a record once you publish it <sup>1</sup>).

In order to increase the findability of the data, all generated data will be accompanied by **metadata**. According to our Grant Agreement, annex 5 (add ref), metadata of deposited publications must be open under a Creative Common Public Domain Dedication (CC 0) or equivalent, in line with the FAIR principles (in particular machine-actionable) and provide information at least about the following: publication (author(s), title, date of publication, publication venue); Horizon Europe funding; grant project name, acronym and number; licensing terms; persistent identifiers for the publication, the authors involved in the action and, if possible, for their organisations and the grant.

a) Metadata for ecological and biological data

<sup>1</sup> <https://help.zenodo.org/docs/deposit/about-records/>

The Field Protocol (WP1) (Diniz et al., 2025) specifies the standardised naming and labelling of the physical samples of pollinator specimens pollen loads and flower head (eDNA), as well as the digital data for samples and sampling events to enable data exchange and allow for syntheses. **Ecological data** sets archived by Butterfly will adhere to the Darwin Core standard vocabulary (Wieczorek et al. 2012) with a thorough description of the data generation process and its spatial, temporal, taxonomic, and thematic extent **also adhering to current metadata standards, such as the Ecological Metadata Language standard**. This ensures that the primary data sources are archived so that they are retrievable long after the project is completed (regardless of the status of EuroAPPA or its constituent databases), adhere to relevant standards to improve interoperability, and are visible to biodiversity data indexing services. In addition, we will improve the visibility of these datasets through the publication of the data through the Global Biodiversity Information Facility (GBIF) and the use of EuroAPPA as a web portal for all ecological deliverables, which will be integrated into the EU Pollinator Hub as well. Similarly, sequence data and their associated annotations generated from the genetic analysis conducted as part of task 1.2.2 will be deposited on an openly accessible sequence database (such as GenBank). Protocols on the archiving of data products generated in WP1 (both the intermediate data products generated as part of the mobilisation of grey literature in T1.2.1 and the new data generated as part of the field campaign in T1.2.2) in the open-access extension of the EuroAPPA repositories.

b) Metadata for human participant and other data

**TBA**... Publications will have bibliographic metadata attached. It will be in a standard format and include the terms "European Union (EU)" & "Horizon Europe"; the name of the action, acronym & grant number; publication date, length of the embargo period, if applicable; and a persistent identifier. The metadata will comply with anonymisation processes.

### 3.2 Making data accessible

*Will the data be deposited in a trusted repository?*

*Have you explored appropriate arrangements with the identified repository where your data will be deposited?*

*Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?*

*Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.*

*If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.*

*Will the data be accessible through a free and standardized access protocol?*

*If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?*

*How will the identity of the person accessing the data be ascertained?*

*Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?*

*Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?*

*How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?*

*Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)?*

It is an overarching aim of Butterfly to make data and results visible and freely accessible and to ensure long term data preservation. Access to research data should be '**as open as possible, but as closed as necessary**', and here there are some differences between the types of data.

#### **a) Ecological, biological and economic data**

All **ecological data** sets that Butterfly will assemble, such as the plant-pollinator network information garnered from the literature review and field campaigns undertaken as part of WP1, will be archived on at least two open-access repositories: Zenodo and EOSC and another platform that specializes in biodiversity data. The European Open Science Cloud (EOSC) enables the storage, sharing, processing and reuse of digital research outputs following FAIR practices. EOSC will support the long-term legacy of Butterfly-initiated research, through the preservation of unpublished project knowledge and data.

The EuroAPPA portal will provide open access to all ecological data gathered in WP1.

For Butterfly's source code of software for the APIs and R packages (WP1) **GitHub** will be used as a repository for the archival of source code related to digital deliverables.

Curated data from T2.1 on agricultural practices, agricultural systems, value chains, other economic sectors, and ecosystems, will culminate in a Data Repository (D2.1).

#### **b) Human participant data**

**Main repositories: Zenodo and EOSC – or disciplinary repository?**

For human participant data, the 'as closed as necessary' needs specific consideration. Specific measures will be taken to accommodate protection of privacy and GDPR. Anonymized and De-identified human participant data can be archived in EOSC and in repositories consortium partners' home countries (e.g. the Norwegian Agency for Shared Services in Education and Research) to ensure transparency and reproducibility of our social science and humanities research in WP2, 4, 6 and 7

**Quantitative** data from the large-scale surveys on citizens' perception and willingness to pay for preserving the pollinators and pollination services (6 countries, 1000 to 1500 respondents per country) will **be anonymous, and the data can be deposited in Zenodo**

**Qualitative data** is more challenging to anonymize. Thus, the exact extent of openness of the actual datasets can be amended. When in doubt, the consortium will refrain from publishing raw datasets and only report aggregate measures. Data that cannot be anonymised due to practical or technical reasons is excluded from publication to ensure sufficient protection of the fundamental rights and freedoms of the (potentially) affected data subjects. Data that can be curated and be made de-identifiable can be shared more broadly. Decisions will be made on a case-by-case basis to ensure that privacy, anonymity, and confidentiality are not breached by publication of datasets or any other type of publication. Consultation with the relevant Data Protection Offices can be sought during the lifetime of the project.

**However, the metadata of the data will be opened**

### **3.3 Making data interoperable**

*What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones?*

*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? Will you openly*



*publish the generated ontologies or vocabularies to allow reusing, refining or extending them?*

*Will your data include qualified references<sup>2</sup> to other data (e.g. other data from your project, or datasets from previous research)?*

File formats that are universal, cross platform, open source, with open standard will be applied, such as (xt, pdf, csv, tsv etc.).

**As the project progresses and data is identified and collected, further information on making data interoperable will be outlined in subsequent versions of the DMP.**

### **3.4 Increase data re-use**

*How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?*

*Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?*

*Will the data produced in the project be useable by third parties, in particular after the end of the project?*

*Will the provenance of the data be thoroughly documented using the appropriate standards?*

*Describe all relevant data quality assurance processes.*

**TBA!**

**In informed consent - letting us utilise data for other**

## **4. Other research outputs**

*Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects related to the allocation of resources, data security and ethical aspects.*

*In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).*

*Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.*

Deliverable D1.4 will create a Europe-wide interactive prediction maps of plant-pollinator networks to assess current and future trends in the spatial structure and function of plant-pollinator networks across Europe.

---

<sup>2</sup> A qualified reference is a cross-reference that explains its intent. For example, X is regulator of Y is a much more qualified reference than X is associated with Y, or X see also Y. The goal therefore is to create as many meaningful links as possible between (meta)data resources to enrich the contextual knowledge about the data. (Source: <https://www.go-fair.org/fair-principles/i3-metadata-include-qualified-references-metadata/>)

D 4.2 Toolbox for resilience thinking, which will be disseminated to businesses beyond BUTTERFLY stakeholders and to EU policymakers. Sector specific reports produced as deliverables of Delphi studies are aimed to be disseminated via sectoral trade associations, such as Cosmetics Europe, EFPIA, AnimalhealthEurope etc.

Deliverables 5.1 to 5.5 will establish 'pollination alert maps' and landscape and 'environmental mitigation tools' that consolidate project-generated results and expert knowledge to raise awareness about the pollinator crisis and facilitate communication and evaluation of diverse mitigation actions.

## 5. Allocation of resources

*What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.) ?*

*How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)*

*Who will be responsible for data management in your project?*

*How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?*

Managing data according to the FAIR brings two overarching types of costs:

- 1) Fees for depositing data in global data repositories. We have chosen to use Zenodo, which is free of charge for uploading data.
- 2) Article processing charges (APC) for publishing data in open access journals.

**Each beneficiary leading the respective Task** is responsible for preparing the datasets including metadata. All partners must create, manage, analyse, store and/or share data and/or datasets with respect to the applicable national and international legislation on data protection. Also, the quality control of these data falls under responsibility of the institution leading the respective Task. In the Consortium agreement, it is stated that the principal investigators and the Data Protection Officer of each beneficiary organization are considered responsible for the DMP actions

Data collectors have the ultimate responsibility of complying with the specifics of the Data Management Plan, as well as with the related GDPR policies and applicable local, government and international laws, regulations and guidelines.

## 6. Data security

*What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)?*

*Will the data be safely stored in trusted repositories for long term preservation and curation?*



Following the UoB storage guide<sup>3</sup>, and as mentioned in section 3.2, data will be stored in two trusted repositories. This will ensure data recovery if needed.

Data will be stored and processed on each partner's own harddrives.  
To share data, partners will use Nextcloud with password protection.

Complying with GDPR, no personal or identifying data will be stored with response data. Such personal/identifying data will be kept in a separate, password protected location with access only for authorised members of the project team. It will not be shared between partners.

No personal data that is considered sensitive, according to the EC commissions definition<sup>4</sup> will be collected in this project.

## 7. Ethics

*Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).*

*Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data?*

*Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?*

In Butterfly's 'Ethics self-assessment' (chapter 4 in the DOA)), the following ethical issues were identified: 1) human participation, 2) personal data collection of data subjects. An initial assessment was included in the GA and summarised below:

*The consortium will ensure that all necessary procedures are followed, particularly with regard to the signing, collation, and storing of all necessary Informed Consent Forms prior to the collection of any data. All involved stakeholders and citizens will be informed in detail about measures and the consortium will obtain free and fully informed consent*

*All necessary actions will be taken within the project management and by all beneficiaries to ensure compliance with applicable European and national regulations and professional codes of conduct relating to personal data protection. This will include in particular Directive 95/46/EC regarding data collection and processing, the General Data Protection Regulation (GDPR, 2016/679), and respective national requirements, ensuring legal and regulatory compliance. Ethics considerations will feed into research and data collection protocols used in the project. This will include the collecting and processing of personal data as well as surveys and interviews. For all identified issues, in line with the above standards, ethical approvals will be obtained from the relevant national data protection authorities and/or institutional boards.*

*In addition to relevant national data protection authorities, the university partners have separate institutional ethics boards or respective national research boards, which will ensure the correct implementation of all human participation and data protection procedures and protocols around social science research. In detail, this includes for Norway (UiB, UiA) the Norsk senter for forskningsdata (Sikt)*

<sup>3</sup> <https://www.uib.no/en/foremployees/153608/storage-guide>

<sup>4</sup> The following personal data is considered 'sensitive' and is subject to specific processing conditions:

- personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs;
- trade-union membership;
- genetic data, biometric data processed solely to identify a human being;
- health-related data;
- data concerning a person's sex life or sexual orientation.

[https://commission.europa.eu/law/law-topic/data-protection/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive\\_en](https://commission.europa.eu/law/law-topic/data-protection/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_en)

(Butterfly 2024, Description of the action (DoA) Part B, pp 34-35)

In order to follow up this summary, the following general measures will be taken:

- Each time participants will be invited to provide data, an informed consent sheet will be provided that specifies the purpose of the data collection, how personal data will be anonymised and stored, etc. In the Appendix in this document, a generic informed consent form is provided that can be used, adapted and translated by the project members who are carrying out the different data collection activities. The consent form also contains information on how data will be shared and preserved, and that participants can decide to withdraw at any point. For WP7, a Memorandum of Collaboration (MoC) will clarify in detail the responsibilities and rights of participants, access to information and results obtained, and processes for resolving issues arising among members. However, it does not replace informed consent. When specific activities are set up for human participant data collection, such as a focus group interview, individual interviews, more detailed consent form has to be provided.
- All personal data will be anonymised by the person collecting the data in each country. In WP7, this will be guaranteed through Anonymisation: Each LL assign a unique ID number to each participant – store codes in a secure file, replace all names with ID numbers – only the anonymised data file with ID number will be sent to WP 2 for analysis
- As mentioned in chapter 6, no personal data will be shared among partners or transferred from the country where it is collected. Only anonymized data will be shared with other project members for analysis.
- Measures will be taken to avoid questions that provide recognisable data. For example, where information about gender, age or place of birth is not necessary, these questions will be avoided to ensure that participants are not recognizable in the anonymised data.
- Where the data is intended to be open access, participants will be made aware of this fact, including the planned processes of anonymisation/pseudonymisation, and any potential risks to their identification.
- Data collection will be carried out in different countries and ethics boards or similar must be consulted in each country, as rules and procedures differ between countries.
- To minimise research fatigue among participants, re-use of existing data is encouraged, such as farm data for WP 2.

### Ethics and sample collection

As specified in **The FIELD PROTOCOL (WP1)**, p17, it is the responsibility of every B-Site leader to obtain the **collection permits** for their sites and to ensure that sample collection and exportation comply with the **Nagoya Protocol**.

HISTORY OF CHANGES		
VERSION	PUBLICATION DATE	CHANGE

1.0		Initial version (new MFF).
1.1		Reformatted to align with other deliverables templates.

## References

Breeze T.D. and Zugic N.J. (2025). *Outline of Issues raised by the External Independent Ethics Advisor*. VALOR project deliverable D9.1

Breeze T.D., Bartomeus I., Damiens F., Kantelhardt J., Kleijn D., Klein A.M., Meinilä J., Schaak H., Sihvonen H., Stoyanova C. and Zugic N.J. (2025). **VALOR Standard Operating Procedures**. VALOR project deliverable D8.1.

### Butterfly (2024b) **Description of the action, Part B**

Diniz, U.M., Myint, A.A., Rasmussen, C., Ollerton, J., Chipperfield, J., Fossøy, F., Leonhardt, S.D. Field protocol (WP1). Description and instructions for the standardization of the sampling design, data collection methods, and sample processing across B-Sites. **Butterfly project deliverable (?)**.

Drucker, D. P., Salim, J. A., Poelen, J., Soares, F. M., Gonzalez-Vaquero, R. A., Devoto, M., Ollerton, J., Kasina, M., Carvalho, L. G., Bergamo, P. J., Alves, D. A., Varassin, I., Tinoco, F. C., Rünzel, M., Robinson, D., Cardona-Duque, J., Idárraga, M., Agudelo-Zapata, M. C., Marentes Herrera, E., Taliga, C., Parr, C.S., Cox-Foster, D., Hill, E., Maués, M.M. Agostini, K. Rech, A.R., Saraiva, A. (2024). WorldFAIR (D10.3) Agricultural biodiversity FAIR data assessment rubrics (Version 1). Zenodo. <https://doi.org/10.5281/zenodo.10719265>

Mačiulienė, M. (2022). Beyond open access: conceptualizing open science for knowledge co-creation. *Frontiers in communication*, 7, 907745. [Frontiers | Beyond Open Access: Conceptualizing Open Science for Knowledge Co-creation](#)

Ollerton, J., Taliga, C., Salim, J.A., Poelen, J.H., & Drucker, D.P. (2025) Incorporating measures of data quality into plant-pollinator databases. *Journal of Pollination Ecology* 38: 151-160

Wintermantel, A., Thompson, A., Fornoff, F., Lenohart, S., Prucker, P., Basaran., Z., Kleoftodimos, G., Anderson, G., Rundoff, M., Schweiger, O., .... Breeze, T. (2024). *RestPoll Data Management Plan 1*. Available at <https://restpoll.eu/publications/?term=deliverables>

Zhang, J. and I. Steffan-Dewenter (2022). *Data Management Plan. Deliverable D8.3*

ALLEA (2023) **European Code of Conduct for Research Integrity** 2023 Revised Edition. <https://allea.org/portfolio-item/european-code-of-conduct-2023/>

## Appendix A: General informed consent form



# Butterfly

INSERT INSTITUTION NAME

## Information letter

### Collection of data for research

#### About Project Butterfly

The decline of pollinator populations poses a serious threat to ecosystems and food security, with cascading effects on biodiversity and economic stability. In this context, the EU-funded BUTTERFLY project will strengthen society's ability to anticipate and respond to these challenges. More specifically, it will establish geographically diverse multi-stakeholder communities to collaborate on proactive restoration solutions for pollinators. The project aims to collect and share key ecological information, model the economic consequences of pollinator loss, and assess the dependence of key supply chains on pollination. Through innovative tools and strategic alliances, BUTTERFLY will integrate pollinator management across sectors, ultimately informing EU policies and promoting resilience in vulnerable communities.

#### Consent form:

**Name and organization of data collector:** (to be filled in by research team).

**Name of the research participant:** \_\_\_\_\_

I, \_\_\_\_\_ (the research participant), have been informed that:

1. Data is being collected as part of the project Butterfly.
2. Data will be used for scientific analysis, publication and dissemination activities.
3. Data will be anonymized for publication/dissemination purposes.
4. Anonymised data will be analyzed by (insert task leader name).
5. Participation is voluntary.
6. Consent for participation in the project can be withdrawn by contacting the data collector, before (insert date), after which date the data will be anonymised.
7. (If applicable): The conversation will be voice recorded, for transcription, and will subsequently deleted.
8. Data will be used by the [specify partner institution] and information containing personal identification will not be exchanged.
- 9.

Signature: \_\_\_\_\_ (participant)

Signature: \_\_\_\_\_ (data collector)

Date \_\_\_\_\_

**Article 13 - EU GDPR: "Information to be provided where personal data are collected from the data subject"**

1. Where personal data relating to a data subject are collected from the data subject, the controller shall, at the time when personal data are obtained, provide the data subject with all of the following information:

- (a) the identity and the contact details of the controller and, where applicable, of the controller's representative;
- (b) the contact details of the data protection officer, where applicable; *Please contact Aarhus University at dpo@au.dk*
- (c) the purposes of the processing for which the personal data are intended as well as the legal basis for the processing;
- (d) where the processing is based on point (f) of [Article 6\(1\)](#), the legitimate interests pursued by the controller or by a third party;
- (e) the recipients or categories of recipients of the personal data, if any;
- (f) where applicable, the fact that the controller intends to transfer personal data to a third country or international organization and the existence or absence of an adequacy decision by the Commission, or in the case of transfers referred to in [Article 46](#) or [47](#), or the second subparagraph of [Article 49\(1\)](#), reference to the appropriate or suitable safeguards and the means by which to obtain a copy of them or where they have been made available.

2. In addition to the information referred to in paragraph 1, the controller shall, at the time when personal data are obtained, provide the data subject with the following further information necessary to ensure fair and transparent processing:

- (a) the period for which the personal data will be stored, or if that is not possible, the criteria used to determine that period;
- (b) the existence of the right to request from the controller access to and rectification or erasure of personal data or restriction of processing concerning the data subject or to object to processing as well as the right to data portability;
- (c) where the processing is based on point (a) of [Article 6\(1\)](#) or point (a) of [Article 9\(2\)](#), the existence of the right to withdraw consent at any time, without affecting the lawfulness of processing based on consent before its withdrawal;
- (d) the right to lodge a complaint with a supervisory authority;
- (e) whether the provision of personal data is a statutory or contractual requirement, or a requirement necessary to enter into a contract, as well as whether the data subject is obliged to provide the personal data and of the possible consequences of failure to provide such data;
- (f) the existence of automated decision-making, including profiling, referred to in [Article 22\(1\)](#) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.

3. Where the controller intends to further process the personal data for a purpose other than that for which the personal data were collected, the controller shall provide the data subject prior to that further processing with information on that other purpose and with any relevant further information as referred to in paragraph 2.

4. Paragraphs 1, 2 and 3 shall not apply where and insofar as the data subject already has the information.

Read more about the Butterfly project here: (insert web page)

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101181930

