



# PLAN-B

Tackling light and noise pollution

## REVISED DATA MANAGEMENT PLAN

DELIVERABLE 7.3

Project information	
Project name	PLAN-B - The Path Towards Addressing Adverse Impacts of Light and Noise Pollution on Terrestrial Biodiversity and Ecosystems
Grant Agreement No.	101135308
Start date	01/01/2024
Duration	48 months
Coordinator	prof. Geert Van Hoorick, Ghent University
Contact	Yana Yakushina
 <b>Funded by the European Union</b>	
Deliverable details	
Description:	D7.3 – Revised Data Management Plan (DMP)
Version	0.1
Dissemination level	public
Due date	M24 (31/12/2025)
Lead:	Ghent University
Author(s):	Yana Yakushina (UGent)
Contributors:	Paweł Tysiąc (Gdansk Tech), Catherine L. Barnett (UKCEH), Reinhard Klenke (MLU), Dick Botteldooren (UGent), Timothy Van Renterghem (UGent), Daniel Lisbona (IBE), Sergio Pascual Ramires (UCM), Raúl Pérez Parras (UCM), Graeme Sherriff (USAL), Penelope Vounatsu (Swiss TPH)
Reviewers:	Geert Van Hoorick (UGent), Els Vermeulen (UGent) and Michael D. Wood (USAL)

### Legal Disclaimer

The information in this document is provided "as is," with no guarantee of suitability for any specific purpose. Consortium members are not liable for any damages resulting from the use of these materials, except as required by law. This report includes a disclaimer and copyright notice. Users must obtain permission from copyright holders to reuse third-party material. Users are responsible for any infringement claims. This report was produced under European Commission contract number 101135308. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

## Executive summary

The PLAN-B Data Management Plan (DMP) provides information relating to the management and processing of data that are produced and collected throughout the project, together with the measures for making these data Findable, Accessible, Interoperable and Reusable (FAIR). It facilitates PLAN-B project management, consideration of Ethical requirements and the development of the project's Plan for Dissemination and Exploitation, including Communication Measures (DEC Plan). The PLAN-B DMP is a living document. As a part of the ongoing update, **this is the revised version of the PLAN-B DMP**, submitted at M24 (**31/12/2025**). The subsequent updates will be provided by M48 at the end of the project.

This revised DMP **builds upon the work carried out during the first 24 months of the project**. In this updated version, the DMP has been revised to align with the project's current progress, reflecting the evolution of individual tasks, the generation of new datasets, and the refinement of data management procedures. It provides an updated overview of data collection, processing, storage, and sharing practices to ensure coherence with the project's operational and ethical framework. Additionally, measures to ensure compliance with the GDPR and other relevant regulations were refined, where necessary, and implemented.

The DMP includes information on:

- the handling of research data during and after the end of the project;
- what data will be collected, processed and/or generated;
- which methodology and standards will be applied;
- whether data will be shared/made open access and;
- how data will be curated and preserved (including after the end of the project).

This deliverable (**Deliverable 7.3 of the PLAN-B project**) is based on the Horizon Europe generic template for DMPs.

## Table of contents

Executive summary.....	3
List of Tables .....	5
List of Abbreviations .....	5
1. Introduction .....	6
2. Data Summary .....	6
Data Definition .....	7
Purpose of data .....	8
Types of data .....	8
Data origin .....	8
Data formats .....	8
Data size .....	9
Data utility .....	9
3. FAIR DATA .....	10
Making Data Findable .....	10
Making Data Openly Accessible .....	10
Making Data Interoperable .....	13
Increase Data Re-use (through clarifying licences) .....	13
4. Data security .....	14
5. Ethical aspects.....	14
6. Data Management Plan per Work Package .....	15
WP 1: Mechanism of light and noise pollution impacts.....	15
WP2: Modelling spatio-temporal impacts at regional scale .....	16
WP3: European situational analysis .....	22
WP4: Societal drivers of light and noise pollution .....	23
WP5: Specific measures to assess, prevent and mitigate negative impacts on terrestrial biodiversity .....	24
WP6: PLAN-B measures to raise Awareness and Maximise Impact.....	25
WP7: Coordination and management.....	30
7. Data Management Risks and Mitigation Solutions .....	31

## List of Figures

Figure 1: Overview of data storage options available through UGent to support PLAN-B activities. .... 30

## List of Tables

Table 1: Secondary data sources planned to be used on a European scale in development of the PLAN-B spatio-temporal database .....	17
Table 2: Networking and Collaboration Contacts Database.....	27
Table 3: Communication and Dissemination Metrics Dataset .....	27
Table 4: FAIR data. Making WP6 data findable, including provisions for metadata. ....	28
Table 5: Making WP6 data accessible.....	28
Table 6: Making WP6 data interoperable.....	29
Table 7: Increase WP6 data re-use. ....	29
Table 8: Data Management Risks and Mitigation Solutions.....	31
Table 9: Asset Valuation .....	32
Table 10: Threat Actions.....	33
Table 11: Types of Loss .....	34

## List of Abbreviations

DoA	Description of the Action – annex 1 to the GA
DOI	Digital object identifiers
EC	European Commission
EU	European Union
GA	Grant Agreement
GDPR	General Data Protection Regulation
HEU	Horizon Europe
M [1-48]	Month (number starting from January 2024)
OA	Open Access
OECD	Organisation for Economic Co-operation and Development
DECP, DEC Plan	Plan for Dissemination and Exploitation, including Communication Measures
DMP	Data Management Plan
SCo	Scientific Coordinator
T	Task
WP	Work Package
WPL	Work Package Leader
LP	Light Pollution
NP	Noise Pollution
LNP	Light and noise pollution
TBES	Terrestrial biodiversity and ecosystem services

## 1. Introduction

This report provides **PLAN-B's revised Data Management Plan (DMP)**. Overall, the DMP outlines **how the data collected and generated will be managed** during and after PLAN-B action. It describes the standards and methodology for data collection and generation to be followed, and whether and how these data will be shared.

In this updated version, the DMP has been **revised to align with the project's current progress**, reflecting the evolution of individual tasks, the generation of new datasets, and the refinement of data management procedures. It provides an updated overview of data collection, processing, storage, and sharing practices to ensure coherence with the project's operational and ethical framework. Additionally, measures to ensure compliance with the GDPR and other relevant regulations were refined, where necessary, and implemented.

As indicated in the "Guidelines on FAIR Data Management in Horizon 2020" published by the Directorate-General for Research & Innovation (EUROPEAN COMMISSION), Data Management Plans (DMPs) are a key element of good data management. As part of making research data findable, accessible, interoperable and reusable (FAIR), a DMP should include information on:

- the handling of research data during and after the end of the project;
- what data will be collected, processed and/or generated;
- which methodology and standards will be applied;
- whether data will be shared/made open access and;
- how data will be curated and preserved (including after the end of the project).

The **DMP presented** here, the second version of the DMP (previously, Deliverable 7.2) - **Deliverable 7.3 of the PLAN-B project**, is the position at month 24 (**M24**) of the project. The next formal update is planned for M48. However, the DMP is also being updated on a regular basis throughout the project whenever significant changes arise, such as (but not limited to):

- new data;
- changes in consortium policies (e.g. new innovation potential, the decision to file for a patent);
- changes in consortium composition and external factors (e.g. new consortium members joining or consortium members leaving).

## 2. Data Summary

**Data Management** is an **integral part of the research process** for our project, and PLAN-B's data/research outputs are **managed** in line with the **FAIR principles** (Findable, Accessible, Interoperable, Reusable). All partners, based on the work they will be performing in the project, have contributed to the development of this Data Management Plan (DMP) and provided specific inputs into this Deliverable in the context of the data that they will be generating, processing and/or collecting throughout the project, according to the Work Package (WP).

PLAN-B is operating a 'Data summary spreadsheet' to **track all data/research outputs** and to ensure that these **comply with the requirements** and approaches specified within this DMP. The 'Data summary spreadsheet' is a **living internal document** which will be updated on an ongoing basis as the implementation of the project progresses. This DMP has been developed drawing on information from the 'Data summary spreadsheet', and the evolution of the spreadsheet during the project will provide a basis for future iterations of this DMP.

PLAN-B project data are separated into **four bundles of data**:

- the project management data, collected for the project coordination and management purposes;
- the input data used during the project life within the work packages;
- the research data produced during the project life within the work packages;
- the results data.

## Data Definition

A very **broad definition of data** can be found in a joint document by the United Nations Statistical Commission and the Economic Commission for Europe, produced for the Conference of European Statisticians Statistical Standards and Studies held in Geneva in 2000: *"Data is the physical representation of information in a manner suitable for communication, interpretation, or processing by human beings or by automatic means."*

On 20 January 2021, the **OECD Council** adopted a revised Council Recommendation on Access to Research Data from Public Funding, which **provided a definition** of "*research data*" - "factual records (such as numerical scores, textual records, images, and sounds) resulting from research that is partially or fully funded by public funds, used as primary sources for scientific research, and that are commonly accepted in the scientific community as necessary to validate research findings. This term does not cover laboratory notebooks, preliminary analyses, or drafts of scientific papers, plans for future research, peer reviews, personal communications with colleagues, or physical objects, (e.g., laboratory samples, strains of bacteria, or test animals)." <sup>1</sup> The Recommendations also cover "other research-relevant digital objects from public funding" - "metadata, algorithms, workflows, models, and software." <sup>2</sup>

Together with that, **OECD Recommendations defined "*research data management*"** as "*the part of the research process that deals with organisation and handling of research data, including data management planning, structured storing, description, curation, preservation and provision of*

---

<sup>1</sup> OECD, "Recommendation of the Council concerning Access to Research Data from Public Funding", Adopted on: 14/12/2006, Amended on: 20/01/2021, available online:

<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0347>

<sup>2</sup> Ibid.

*metadata and complementary algorithms, code, software, and workflows, and compliance with internal, national and international privacy legislation."*

This **DMP adheres to the definitions provided** by the OECD and will comply with all binding and non-binding requirements for research data management.

## Purpose of data

The purpose of the data collection/generation in the PLAN-B project is the proper implementation of the project activities and documenting them.

## Types of data

There are different types of data, that will be generated during the project:

1. **Primary data** is data that is collected by a researcher from first-hand sources, using methods like surveys, interviews, or pilot studies (experiments). It is collected with the research project in mind, directly from primary sources. The term is used in contrast with the term secondary data.
2. **Secondary data** is data that has already been collected through primary sources and made readily available for researchers to use for their own research. It is a type of data that has already been collected in the past.
3. **Metadata** is *"a set of data that describes and gives information about other data"*.<sup>3</sup>

## Data origin

**Primary data** will be **collected** by researchers **through various PLAN-B activities**, including pilot study activities, surveys, and citizen science. **Secondary data** will be collected **from sources** such as peer-reviewed journals, patents, technical reports, books, and **data sets compiled through previous studies/monitoring activities or provided by institutions like EEA, JRC, etc.**

The PLAN-B **consortium will re-use data**, which were published **as Open Access** (e.g. GBIF, data from EEA), or collected or produced in other national or EU-projects, such as records of marked individuals from bird- and/or bat species, distribution of sites of renewable energy infrastructures, or parts of computer code, which will be used either directly, changed, updated or extended during PLAN-B and used to inform other models.

## Data formats

Data will be collected in **various formats depending on its type**. However, where technically possible, data will be stored in non-proprietary common and open formats like ASCII, XML (ODF, OSF, CSV, etc.),

---

<sup>3</sup> Oxford Dictionary. Available online:

<https://www.oxfordreference.com/display/10.1093/acref/9780199571444.001.0001/acref-9780199571444-e-2097>



BibTex, GPKG (Maps), etc., and all PLAN-B written deliverables will be compiled and made available as PDF files. Large spatial datasets will be stored in standard geospatial formats such as GeoTIFF (.tif) for raster data and GeoPackage (.gpkg) or other OGC-compliant formats for vector data.

## Data size

Data will have different sizes. **Large data sets**, e.g. for satellite data, maps, model results or image and sound data, are stored in suitable, publicly accessible repositories like <https://zenodo.org>, <https://www.gbif.org>, <https://www.pangaea.de/>. **Pure Text documents** will be kept below 5MB wherever possible to facilitate sharing. For **texts with illustrations** (figures, pictures, pictograms), compression algorithms will be used to reduce file size as far as possible whilst maintaining readability and usability of the output. **Some PLAN-B datasets** (high-resolution satellite imagery, ELF image series, 3D lighting models (e.g. Gaussian splatting representations), and European-scale LP/NP maps) are expected to **reach sizes up to tens or probably also hundreds of gigabytes**.

## Data utility

Data will be **useful to complete the project's activities**. Furthermore, public data will have **broad utility beyond PLAN-B**, including:

- Supporting strategic decisions in nature conservation as well as urban and landscape planning.
- Defining thresholds, underpinning legal and policy frameworks.
- Informing industries, policy- and decision-makers and the general public about the environmental terrestrial impact of light and noise pollution, including spatial and/or temporal trends.
- Informing industries, policy- and decision-makers and the general public about adequate solutions for preventing and mitigating light and noise impacts on terrestrial biodiversity and ecosystem services.
- Researchers to use the data in new research.
- Creating overviews and statistics.

### 3. FAIR DATA

#### Making Data Findable

The PLAN-B project **uses** the online collaboration platform **SharePoint**. SharePoint is a web-based collaborative platform that integrates natively with Microsoft 365. SharePoint is used as a **document management and storage system**. PLAN-B's SharePoint site is available as a repository and collaboration instrument for all working documents, minutes, reports and datasets. For more information, see "[D7.1 Project Management Handbook](#)". As data products and associated **documentation are finalised**, they will be made **publicly findable** through the use of appropriate **repositories**. Data products uploaded to repositories and made publicly available will be documented with corresponding metadata providing appropriate **keywords** and linked via unique **digital object identifiers** (doi). Where appropriate, we will also utilise data papers published under Open Access policies in corresponding journals (e.g. Biodiversity Data Journal (<https://bdj.pensoft.net/>), NATURE Scientific Data (<https://www.nature.com/sdata/>) or Data (<https://www.mdpi.com/journal/data>) and provide them with relevant metadata to maximise the findability and utility of our data products. Additionally, together with AquaPLAN, the PLAN-B project has created LNP [Zenodo](#) and [Wikidata](#) repositories, which can be used free of charge and provide long-term storage for more than 10 years.

#### Making Data Openly Accessible

Horizon Europe's open access policy pursues that the information generated by the projects participating in that programme is made publicly available. But, in line with EC guidelines on Data Management, *"As an exception, the beneficiaries do not have to ensure open access to specific parts of their research data if the achievement of the action's main objective, as described in Annex I, would be jeopardised by making those specific parts of the research data openly accessible. In this case, the data management plan must contain the reasons for not giving access."*

**PLAN-B project will not put into the open domain the data** linked to exploitable results if they were **provided by third parties under restriction for limited use** only for particular aims in PLAN-B, compromise its commercialisation prospects or have inadequate protection or constitute a commercial or other secret; **the rest of the data and derived products produced by PLAN-B will be open access**, made findable as described above, and promoted via the project website, social media and repositories. **Where datasets cannot be made openly available due to such restrictions, PLAN-B** will, wherever possible, **provide derived or aggregated secondary data products** that can be safely reused under less restrictive conditions. The associated repository records will clearly describe the relationship between the original and secondary data and specify any remaining access constraints or steps required to obtain the original data from the primary provider.

For **WP2 light pollution mapping activities**, the **datasets originate from sources with different licensing restrictions**. The SDGSAT-1 data are distributed under a non-commercial licence, which

explicitly prohibits open redistribution. **PLAN-B** is therefore **not authorised to make the raw data publicly available**, as doing so would violate the data provider's terms. Access to the original SDGSAT-1 products may be granted on request for strictly non-commercial scientific use and only under the conditions established by CBAS and the data providers.

In addition, part of the processing workflow relies on **Google Earth Engine (GEE)**. According to Google's terms of service, both the platform and its outputs are licensed **exclusively for research and educational purposes** and cannot be used or disseminated for operational or commercial applications. Consequently, **PLAN-B cannot release raw or intermediate datasets generated via GEE in an unrestricted form.**

In contrast, astronaut imagery from **the International Space Station (ISS)** can be released as **fully open access**, provided that no data product infringes the licensing conditions associated with GEE. All ISS images and mosaics produced within **PLAN-B will therefore be openly accessible through project repositories.**

To comply with third-party restrictions while still ensuring transparency, **PLAN-B will make available only derived or aggregated datasets that can be safely shared.** These **secondary products will be openly accessible for non-commercial research** use and will include metadata explaining their provenance and how they relate to the original restricted data. They will be deposited in trusted repositories for the long-term access and linked via unique **digital object identifiers (doi)**, as already mentioned above.

Where **third-party licensing or commercial constraints prevent PLAN-B from making original datasets fully open** (e.g. SDGSAT-1 or data linked to exploitable results), **access will be provided through controlled, request-based procedures:**

#### **During the project:**

- Original third-party datasets (e.g. SDGSAT-1 products) will not be redistributed by PLAN-B. Instead, the **metadata and documentation will provide clear instructions** on how to **obtain access directly** from the primary data providers, including a link to the relevant request portals and applicable licence conditions.
- Where **PLAN-B is allowed to share data** under restricted conditions (e.g. bilateral agreements), **access will be granted upon submission of a data access request** to the project coordinator or data manager. Requests will need to explicitly acknowledge and accept the relevant licence terms (non-commercial use, no further redistribution, citation requirements, etc.).

#### **After the end of the project:**

- The **same procedures will apply.** Repository records and project documentation will continue to point to the primary providers for original SDGSAT-1 and other restricted datasets, and to the relevant contact point within the consortium (or successor institution) for any PLAN-B-managed restricted data.

- All **derived or aggregated secondary datasets** that can be safely shared will **remain openly accessible** via trusted repositories (with DOIs), **under appropriate non-commercial licences** where required. The associated metadata will clearly distinguish between (i) openly available secondary products and (ii) original restricted datasets that must be requested from third parties.

In all cases, the repository entries will state the **licence, usage restrictions, and access procedure**, including any forms, agreements or contacts needed to request access.

For **datasets that can only be shared under specific conditions** (e.g. non-commercial, research-only, no redistribution), PLAN-B will use the **following mechanisms**:

- Users will be asked to submit requests using a **recognisable institutional email address** and to provide their **name, affiliation, and role** (e.g. academic, student, public authority, NGO).

**Metadata for all datasets and secondary products will be made openly available** and, by default, will be released under a **public domain dedication (CC0)**, in line with the Horizon Europe Grant Agreement.

This applies **even where the underlying data cannot be openly shared** (e.g. SDGSAT1). In such cases:

- the metadata will remain openly accessible;
- the records will clearly specify the **access constraints**, licence terms, and the procedure for obtaining the data (e.g. contacting the relevant PLAN-B partner, including links/forms where applicable).

**Metadata records will provide key information of relevance for the dataset, such as:**

- dataset title, description and keywords;
- spatial and temporal coverage;
- data formats and structure;
- provenance and processing steps (including references to source datasets and tools);
- licensing and usage conditions;
- **access information** (primary provider, access URLs, contact details, or access request procedures);
- associated DOIs and links to related publications.

All **openly accessible datasets and secondary products** will be deposited in **trusted repositories** (institutional, disciplinary, or generalist) that commit to long-term preservation. PLAN-B aims to ensure availability and findability for **at least 10 years after the end of the project**, subject to repository policies and infrastructure.

**Software**, which is used within PLAN-B, will be **either commercially available, freely available open source or produced by PLAN-B**. **Algorithms and code developed** in PLAN-B, such as those written in

R, Python, SQL or other standard programming languages, will be **made publicly available** either as supplemental material of scientific papers or via the repositories.

## Making Data Interoperable

PLAN-B research **data will use open standard formats** wherever possible to **maximise interoperability**. Where this is not possible, we will seek to use proprietary formats that are compliant with available and widely used software applications. This will facilitate re-combination and interoperability between researchers, institutions, organisations and countries. The associated vocabulary and methodologies will be communicated using plain language to allow interoperability at the interdisciplinary level.

## Increase Data Re-use (through clarifying licences)

**PLAN-B Deliverable 6.1 and Deliverable 6.2 will set up the Plan for Dissemination and Exploitation, including Communication Measures (DEC Plan) of the results.** The DEC Plan will address project results and identify the key exploitable results.

With regard to increasing the re-use of the data, the PLAN-B project will follow the guidelines provided under the “Horizon Europe Programme Guide” published by the European Commission.<sup>4</sup>

Given the nature of the PLAN-B consortium, the DMP will **facilitate interoperable access** and use of data between researchers, organisations and countries by **standardising data management criteria** to facilitate the combination and analysis of numerous data sets. Wherever possible, data and associated **outputs will be published as “Gold Open Access”**. Where this is **not possible, a “Green Open Access”** route will be followed, recognising any embargo periods that may need to be applied. The preservation of data beyond the lifetime of the PLAN-B project will ensure that the **data generated is usable by third parties for wider sector** applications beyond its original purpose. However, a situation may arise where access to PLAN-B data will be curtailed in the Consortium Agreement and the DMP. Restrictions in data sharing may be due to IP issues or commercial sensitivity.

To **maximise future re-use, key PLAN-B databases**, such as the European-scale light and noise maps, will be **accompanied by clear and simple data schemas** so that external users can load and combine them with their own data using standard tools. Where the underlying raw inputs cannot be shared, PLAN-B **will provide reusable derived products**, such as LP and NP indicator maps and aggregated tables, that can still be used in future research and policy assessments at local, national and European scales.

---

<sup>4</sup> European Commission (2024). Horizon Europe Programme Guide. Available online: [https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide\\_horizon\\_en.pdf](https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide_horizon_en.pdf)

The PLAN-B project will **promote data re-use by providing clear metadata and licensing information for all datasets**.

For light pollution mapping products, the **original SDGSAT-1 data cannot be redistributed**, but all **derived datasets generated within PLAN-B** (e.g. colour-classified maps, trend analyses, and ALAN exposure indices) **will be published under CC-BY-NC 4.0 licences**. This ensures wide reusability for non-commercial scientific applications while protecting third-party rights.

**Metadata will explicitly reference** the relationship between original and derived datasets and include access links or instructions to request SDGSAT-1 data directly from CBAS. The **open datasets will remain accessible beyond the lifetime of PLAN-B**, allowing future integration into European-scale assessments of artificial light and biodiversity impacts.

## 4. Data security

**Each partner** involved in generating or collecting data for the project purpose **has the responsibility for secure data storage, data recovery and secure transfer of data**, whether they are sensitive or not. The partners have the responsibility of identifying or appointing their Data Protection Officer within their organisation.

The project coordinator can support partners, whether they need specific information about common procedures and general regulations.

PLAN-B **project does not implement activities or results raising security issues**, and it **does not include “EU-classified information”** as background or results.

## 5. Ethical aspects

The **only ethical issue** to be considered for the **project data is the General Data Protection Regulation (GDPR)**.

It will be considered when getting personal data for **newsletter** subscriptions, **citizen science** activities, **surveys** and whenever personal data is needed. In these cases, a tick box will be set to get the access rights permission, and **data will be treated only for the purpose informed to the person**, ensuring the right to:

- the protection of their personal data.
- access to data collected about them and the right to have it rectified or removed.

For **the survey of the general public** that is conducted as part of **WP4**, the **personal information** (email address) will be collected **only for the purpose of our prize draw and providing further information to the participant**. In both these cases, these questions are optional. All **other data** collected as part of the survey are **completely anonymous**. This personal information will be separated from the survey

responses before analysis. No participants will be identifiable in any outputs containing data and analysis relating to the survey. The survey is **conducted on JISC Online Survey**, a service created for and widely used by UK Universities.

## 6. Data Management Plan per Work Package

PLAN-B is a complex interdisciplinary multi-actor project, involving a variety of activities stemming from different fields and areas. The section below provides more specific details of data management concerning a particular Work Package (WP).

### WP1: Mechanism of light and noise pollution impacts

**WP1 provides crucial underpinning knowledge for other PLAN-B WPs**, advancing understanding of the mechanisms and extent of LNP impacts on terrestrial biodiversity and ecosystem services (TBES). Following a systematic literature review, **bibliographic information** will be collated into an overarching **BibTeX file (.bib)**, which will subsequently be **queried for publications related to specific organism groups** (e.g. mammals, birds, insects, etc.) **to create organism-specific BibTeX files (.bib)**. **Secondary data** obtained following the literature review will **be collated into a knowledge base (.csv)**. **The knowledge base** will also contain **descriptive metadata**. All these files and the subsequent meta-analysis will form the major data outputs from WP1.

The **knowledge base will collate published and unpublished** ('grey' literature data) on the impacts of LNP on TBES. For transparency of data provenance, **summarised bibliographic information** as well as **research observation/measurement data** will be **contained within the knowledge base**. Where originating data sources are not currently accessible, such as unpublished data, the PLAN-B team will approach the originators to **make those data accessible** and provide guidance to facilitate this.

**Data within the knowledge base** will be **compiled** through a **Systematic Review**, which will be conducted by a subgroup within the PLAN-B consortium. This knowledge base subgroup will also be responsible for **inputting compiled data into the knowledge base**.

During development and testing, the knowledge base will be **hosted on the UGent PLAN-B's SharePoint site** in the dedicated WP1 folder. Data entry testing will be conducted by members of the WP1 subgroup using a subset of publications. This will ensure that all members of the subgroup are able to develop a **common understanding and approach** to extracting and inputting data into the knowledge base.

For data population, **exact copies of the knowledge base** will be created, and these will also be **hosted on the UKCEH PLAN-B's SharePoint site**. Upon completion of data entry, these **copies** will be **combined** into the **final version of the knowledge base**. This file will be **hosted and accessed via the UGENT PLAN-B SharePoint site** (WP1 folder) for the remainder of the PLAN-B project and used for meta-analysis. Hosting in this way will ensure full version control of the knowledge base, and that edit



rights can be restricted to just those within the PLAN-B knowledge base subgroup. It will also ensure continuous, secure backup of the knowledge base as it develops.

**Knowledge base data quality** will be assured through procedures developed within WP1. These procedures will include:

- Each **member** of the knowledge base data entry **subgroup** has agreed to **take responsibility for data entry** related to a specific organism group (e.g. birds, insects, etc.), so the individual responsible for inputting each data row is identifiable.
- **Quality control of the final knowledge base will be conducted.** A proportion (e.g. 5%) of randomly selected data entries will be **manually checked** by a member of the PLAN-B team who was not responsible for the initial input of the data.

Upon completion of the **final version of the knowledge base**, it will be made **available** via an **appropriate open-access data repository** (such as <https://eidc.ac.uk/>). Supporting **metadata** and a **full data description** will also be **provided**. The **bibliographic information** collated following the systematic literature review will also be **published** via an appropriate **open-access source** (e.g. [Zenodo](#) and [WikiData](#)).

**Other WP1 deliverables** will be predominantly text-based and shared **openly** using **standard formats** (e.g. ODT, PDF, etc.). During development, these documents will be **hosted on the PLAN-B SharePoint site** (WP1 folder) to ensure version control and secure backup.

## WP2: Modelling spatio-temporal impacts at regional scale

Due to the complexity of this WP, the data management is discussed under the following headings:

- Spatio-temporal data;
- Data related to light pollution modelling and mapping;
- Data related to noise pollution modelling and mapping.

These data sets will all be **large spatial data sets** and are expected to **exceed 10GB**. It will therefore be necessary to host on secure local servers of participating organisations during the PLAN-B project. Each of the following sections provides a description of the large spatial dataset that will be developed, an outline of the data specification and details of plans for storage and preservation of the data.

- **Spatio-Temporal Data**

A **spatio-temporal database on terrestrial biodiversity indices, conservation status and contextual factors during 2000 – 2025 will be developed**. The key secondary data sources to be used are indicated in the construction of the **PLAN-B spatio-temporal database**, as shown in *Table 1*. These data will be used to **model spatiotemporal impacts** of light, noise and other drivers **on terrestrial biodiversity and conservation status** over the last 25 years and **predict changes in Europe till 2030**.



The temporal resolution will be yearly, and the spatial resolution will be at specific point locations across Europe, defined by latitude and longitude.

## Data Description

The **database** will compile (1) **existing indices on biodiversity and conservation** status, including species habitat index (SHI) and species protection index (SPI); (2) **new biodiversity indices sensitive to LP & NP** (based on taxa/species identified in WP1) that will be developed as part of the project, and (3) **data on confounders such climatic and anthropogenic factors**. The new biodiversity and conservation status indices developed within PLAN-B will be model-based and use data extracted from the Global Biodiversity Facility (GBIF) and the IUCN Red List of Threatened Species databases. Climatic data such as air temperature and extreme precipitation will be extracted from the ERA5 climate reanalysis dataset produced by the Copernicus Climate Change Service. Air pollution data will be extracted from Bayesian geostatistical model-based surfaces of PM<sub>2.5</sub>, PM<sub>10</sub> and NO<sub>2</sub> available at Siss TPH at 1km<sup>2</sup> spatial resolution across Europe and annual intervals.

## Data Specifications

- **Digital/Non-Digital:** The data is digital.
- **Data Type:** Observational and model-based data.
- **Digital Format:** The data will be stored in .tif file format.
- **Digital Data Volume:** The data volume is expected to exceed 10 GB.

## Storage and Preservation

- **Storage During Project:** Secure, backed-up servers at Swiss TPH, WP2 folder of SharePoint hosted by UGent, external hard disks, backed-up servers at the Department of Information Technology at UGent.
- **Deposition Repository:** Upon completion of the project, the data will be deposited in [Zenodo](#), a reputable repository for preserving and sharing research data.
- **Persistent Identifier:** As the data have not yet been generated, a persistent identifier (such as a DOI or accession number) has not yet been assigned, but it will be in due course. Either via [Zenodo](#) or the publication. This will be obtained upon deposition in the repository.

*Table 1: Secondary data sources planned to be used on a European scale in development of the PLAN-B spatio-temporal database*

<b>Description:</b> <a href="#">Biodiversity</a> <b>Source:</b> Global Biodiversity Information (GBIF) network; <b>Time period:</b> 1992-2024; <b>Spatial Resolution:</b> Points (Presence data); <b>Temporal Resolution:</b> monthly
<b>Description:</b> <a href="#">Conservation status</a> <b>Source:</b> IUCN Red List of Threatened species; <b>Time period:</b> 1992-2024; <b>Spatial Resolution:</b> Points (Presence data); <b>Temporal Resolution:</b> annual
<b>Description:</b> <a href="#">Species Habitat Index (SHI)</a> , <a href="#">Species Protection Index (SPI)</a> , <a href="#">Species Status Information Index (SSII)</a> <b>Source:</b> GeoBon; <b>Time period:</b> 2001-2020; <b>Spatial Resolution:</b> 1km x 1km; <b>Temporal Resolution:</b> annual

<b>Description:</b> <a href="#">Air temperature, Total precipitation</a> <b>Source:</b> ERA5-Land Monthly Averaged – ECMWF Climate Reanalysis; <b>Time period:</b> 1992-2024; <b>Spatial Resolution:</b> Points; <b>Temporal Resolution:</b> monthly
<b>Description:</b> <a href="#">Extreme precipitation</a> <b>Source:</b> ERA5-Land Daily – ECMWF Climate Reanalysis; <b>Time period:</b> 1992-2024; <b>Spatial Resolution:</b> 11.132 x 11.132 km; <b>Temporal Resolution:</b> daily
<b>Description:</b> <a href="#">Palmer drought severity index (PDSI)</a> <b>Source:</b> TerraClimate; <b>Time period:</b> 1992-2024; <b>Spatial Resolution:</b> 4.638 x 4.638 km; <b>Temporal Resolution:</b> monthly
<b>Description:</b> <a href="#">Land cover</a> <b>Source:</b> Corine Land Cover, Copernicus Land Monitoring Service; <b>Time period:</b> 2006, 2012, 2018, 2022; <b>Spatial Resolution:</b> 100 x 100 m; <b>Temporal Resolution:</b> annual
<b>Description:</b> <a href="#">Air pollution</a> <b>Source:</b> Bayesian geostatistical model-based estimates; <b>Time period:</b> 2006-2022; <b>Spatial Resolution:</b> 1 x 1 km; <b>Temporal Resolution:</b> annual
<b>Description:</b> <a href="#">Soil pollution</a> <b>Source:</b> LUCAS database; <b>Time period:</b> 2009, 2012, 2015, 2018, 2022; <b>Spatial Resolution:</b> Points; <b>Temporal Resolution:</b> annual
<b>Description:</b> <a href="#">Population density</a> <b>Source:</b> Global Human Settlements layers; <b>Time period:</b> 1990, 2000, 2015, 2025, 2030; <b>Spatial Resolution:</b> 1 x 1 km; <b>Temporal Resolution:</b> 5-year-steps
<b>Description:</b> <a href="#">Specially Protected Areas (SPAs) for birds, adopted Sites of Community Importance (SCIs), and Special Areas of Conservation (SACs) for habitats and other species</a> <b>Source:</b> Natura2000; <b>Time period:</b> 1992-2021; <b>Spatial Resolution:</b> 1:100'000; <b>Temporal Resolution:</b> annual
<b>Description:</b> <a href="#">Macroalgae, Angiosperms, Fungi, Reptiles, Amphibia, Invertebrates, Fish, Plants, Mammals, Birds, Macrophytes, Phytoplankton, Plant community</a> <b>Source:</b> EASIN data; <b>Time period:</b> 2014-2021; <b>Spatial Resolution:</b> Points (Presence data); <b>Temporal Resolution:</b> annual
<b>Description:</b> <a href="#">Habitat Conservation</a> <b>Source:</b> Wild Bird Indicator; <b>Time period:</b> 2018-2022; <b>Spatial Resolution:</b> Points (Presence data); <b>Temporal Resolution:</b> annual
<b>Description:</b> <a href="#">Habitat Connectivity</a> <b>Source:</b> wildlife corridor network, Switzerland; <b>Time period:</b> 2021; <b>Spatial Resolution:</b> Polygon Data; <b>Temporal Resolution:</b> annual

### 6.1.1. Data related to light pollution modelling and mapping.

The “Atlas of Colour Light Pollution Map of Europe” aims to create a **comprehensive set of maps depicting Europe at night using satellite and astronaut imagery**. This work combines **SDGSAT-1 satellite data** with images acquired from the **International Space Station (ISS)** to analyse artificial light emissions across the continent. The lead beneficiary, Universidad Complutense de Madrid (UCM), is responsible for generating these new observational datasets, which will be managed according to the following principles.

#### Data Description

- **Title:** Atlas of Colour Light Pollution Map of Europe
- **Dataset Name:** Collection of maps from Europe at Night from satellite data.
- **Data Origin:** SDGSAT-1 satellite imagery and ISS astronaut photography.
- **Purpose:** To quantify, classify and analyse artificial light emissions across Europe for environmental assessment and research on ALAN impacts.
- **New/Reused:** All datasets generated within PLAN-B are new, although they are based on external image sources (SDGSAT-1, ISS)

## Data Specifications

- **Digital/Non-Digital:** Digital observational data.
- **Data Type:** Digital observational data.
- **Digital Format:** GeoTIFF (.tif) and associated metadata files.
- **Digital Data Volume:** Expected to exceed 10 GB during the project duration.

## Storage and Preservation

- **Storage During Project:** During the project, **all datasets** will be **stored** on secure hard drives and institutional network servers managed **by UCM**, ensuring both **accessibility** and **redundancy**. Storage arrangements will also **respect the licensing restrictions** of the original data sources: **SDGSAT-1 raw imagery will be kept under controlled access** according to its **non-commercial licence**, and intermediate **datasets** generated using **Google Earth Engine** will be stored in **compliance with GEE's research-only terms of use**.
- **Deposition Repository:** Upon completion of the project, **only those datasets that can legally be shared**, primarily the **ISS** imagery and the derived or aggregated **secondary products** generated within PLAN-B, **will be deposited in [Zenodo](#)** or an equivalent **open repository**. These openly shareable datasets will include light pollution indicator maps, classified emission layers, and other derived products that do not violate SDGSAT-1 or GEE licensing conditions. **Raw SDGSAT-1 data** and any **GEE-restricted outputs will not be deposited**, but **metadata and documentation will reference the procedure** for requesting access from the original provider.
- **Persistent Identifier:** For datasets that can be **openly deposited**, a persistent **identifier** (such as a DOI) will be assigned upon upload to [Zenodo](#) or through associated peer-reviewed publications. Restricted datasets that cannot be deposited openly will be documented through metadata records but will not receive a DOI unless a compliant repository mechanism is available.

## Documentation and Accessibility

- **Documentation:** Comprehensive documentation will be provided to ensure that all **datasets** are **understandable** and **reusable** by other researchers. This will include detailed **metadata**, descriptions of the processing **methodology**, and any **relevant contextual information** required to interpret the outputs correctly. The documentation will also clarify the relationship between the original image sources (SDGSAT-1, ISS, or Google Earth Engine outputs) and the derived datasets generated within PLAN-B, ensuring full transparency regarding data provenance and processing steps.
- **Post-Project Access:** **Access** to the data after the project's conclusion will **depend on the licensing conditions** of the original sources. **SDGSAT-1** imagery is subject to a restricted **non-commercial licence**, meaning that PLAN-B cannot redistribute the raw data openly; **interested**

users will need to request access through a formal agreement and obtain approval from CBAS for non-commercial scientific use. Outputs generated using **Google Earth Engine** may **only be reused for research or educational purposes**, in line with the platform's terms of service. In contrast, **ISS imagery** and the derived products created within PLAN-B may be made **fully open access**, provided that no component conflicts with GEE restrictions. To support broader reuse while respecting all third-party constraints, **PLAN-B will also make available derived and aggregated secondary datasets** that can be shared openly under non-commercial licences.

## GDPR and Sensitivity

- **GDPR Sensitive:** The datasets contain no personal data and are not GDPR-sensitive.
- **Data Security and Ethics:** Ethical and security considerations relate **exclusively to compliance with third-party licences** and the **responsible reuse** of satellite imagery. All procedures will follow Horizon Europe guidelines on research data management.

### 6.1.2. Data related to noise pollution modelling and mapping.

The "*European natural areas noise exposure maps*" aim to create a **comprehensive collection of maps predicting biodiversity-relevant noise indicators for all European natural areas**. These maps are the results of numerical simulations and are based on a large number of open data sources for both the identification of the relevant sound sources and propagation-relevant environmental features. The lead beneficiary – Ghent University (**UGent**) is **responsible for generating** these maps, which will be **stored and managed** according to the following DMP.

#### Noise source data

Relevant identified environmental noise sources are road traffic, railway traffic, air traffic, mining and wind turbines. All identification and localisation are based on **openly available data sources** and consequently **do not entail specific data management concerns**. The road and rail network will be derived from open street maps (OSM).

Based on statistical data linked to road or railway type, intensities, composition and speeds will be assigned to such infrastructure. Air traffic routes rely on web scraping for a representative period. Wind turbines are localized by OSM in combination with satellite imagery. Active mining zones will be derived from comparing satellite images at different moments. The relevance of industrial and harbour noise for exposure in natural areas in Europe is still unclear.

#### Sound propagation relevant environmental data

Relevant input data for sound propagation modelling is land use, terrain elevation and **meteorological data**. **Open data sources are used here as well**.

- **Land use:** Different types of natural grounds based on Google's "Dynamic World" (water surfaces, trees, grass, crops, shrubs, bare, snow/ice etc.), building footprints based on "Microsoft's Planetary Computer datasets", and other ground types than natural soils based on OSM.
- **Terrain elevation:** using "Digital Elevation Model - SRTM 30 m", "ASTER Global Digital Elevation Map" and "Sentinel 1 - DEM (Synthetic Aperture radar Data)". Where applicable, local DEM databases of countries with finer spatial resolution will be used.
- **Meteorological data:** using "ECMWF Climate Reanalysis" and "ECV Inventory" (ERA5) as long term averages.

### Data Description

- **Title:** Noise pollution maps
- **Dataset Name:** Natural\_areas\_noisemaps\_Europe
- **Data Origin:** Numerical simulations using typical source power spectra and a propagation model dedicated to the specific features in natural areas.
- **Purpose:** Exposure assessment with relation to environmental noise, for different noise indicators and biodiversity relevant receiver heights, to start analysis on its impact on biodiversity.
- **New/Reused:** The data generated for this project is entirely new.

### Data Specifications

- **Digital/Non-Digital:** The data is digital.
- **Data Type:** Simulation results.
- **Digital Format:** The data will be stored in .tif file format and .txt files or similar ASCII formats.
- **Digital Data Volume:** The data volume is expected to be <100 GB.

### Storage and Preservation

- **Storage During Project:** WP folder of SharePoint hosted by UGent, external hard disks, Department of Information Technology servers.
- **Deposition Repository:** Upon completion of the project, the data will be deposited in Zenodo, a reputable repository for preserving and sharing research data.
- **Persistent Identifier:** As the data has not yet been generated, a persistent identifier (such as DOI or accession number) has not been assigned. This will be obtained upon deposition in the repository.

## Documentation and Accessibility

- **Documentation:** Comprehensive **documentation will be provided to ensure that the data is understandable and usable by other researchers**. This will include metadata, methodology, and any relevant contextual information.
- **Post-Project Access:** The noise **maps** produced will be **accessible** after the project has ended. Access to input data for the noise modelling efforts will depend on restrictions of the open data sources.

## GDPR and Sensitivity

- **GDPR Sensitive:** The data is not sensitive in terms of GDPR. It does not contain personal data and is primarily focused on environmental observations.
- **Data Security and Ethics:** Ethical considerations and data security measures will be strictly followed to maintain the integrity and confidentiality of the data throughout the project.

## WP3: European situational analysis

The **purpose of collection of data in WP3** is (1) to compile a sufficient and consistent **basis for the description of spatial and temporal trends** in the increase of intensity and expansion of LNP with regard to biodiversity, (2) to **perform special statistics** about influence and increasing impact of LNP on occurrences and activities of selected species and ecosystem services, and (3) to **derive new products and indicators for monitoring and decision making** in the protection of biodiversity, landscape, as well as ecosystem services. Data collected will be of **different types**, ranging from simple tables to large geographical maps of high resolution. The **size** will range perhaps from a **few kilobytes up to gigabytes**.

WP3 will collect **data on LNP** itself, **drivers of LNP** (industrial development, spread of traffic and energy infrastructure, settlements), **species, habitats and landscape** features or **corresponding indicators**. It is not planned to collect sensitive personal data, which would need special protection.

The data will be **stored internally** on storage devices with regular backup and limited access to secure the property rights of third parties. New data, databases, maps, programming code or other products generated or derived by special analyses and technical methods within **PLAN-B will be described sufficiently with corresponding metadata** (see paragraph “Making Data Findable” above, for geospatial data, this will be based on the EU INSPIRE directive), made **identifiable and accessible via digital object identifiers**, declared as **open access** based on appropriate licenses in required machine-readable form and uploaded to special repositories like [Zenodo](https://zenodo.org/), [pangea.de](https://pangea.de/), [iData](https://data.dryad.org/) or [datadryad.org](https://datadryad.org/). Links and corresponding disclaimers were set to providers of original sources from third parties to take property rights into account and allow backtracking. If possible, **widespread, well-documented and open formats and structures will be used to store and provide data and results**, such as ASCII, XML,



GPKG, BibTex, ODF (ODT, ODS), PDF, etc. Input **data**, which **cannot be made freely accessible** because of third-party property rights and other policies, will be **stored and kept accessible internally for documentation** according to the legal rules valid for the Martin Luther University Halle-Wittenberg (MLU), as long as it is allowed by the contracts with the providing property owners.

Costs of storage and backup in repositories at MLU will be covered by the general overheads of the institution, costs of publication fees in long-term repositories with public access has to be paid. The German Research Foundation (DFG) gives financial support for Open-Access publication of results to institutions. The **repositories** [www.zenodo.org](http://www.zenodo.org), [pangea.org](http://pangea.org) and [Wikidata](https://www.wikidata.org) can be **used free of charge** and provide **long-term storage** for more than 10 years.

## WP4: Societal drivers of light and noise pollution

The **purpose of the WP4 data collection** is: (1) to **identify a regulatory framework** for LNP; (2) to **understand human behaviour**, impacting LNP, and (3) to **suggest the best social and regulatory solutions** for mitigating LNP impacts on TBES.

The DMP for WP4 aims to comprehensively capture and analyse **various datasets**. Primarily this data will **include qualitative data**, such as gathering existing research findings (research papers, research reports, newspapers, blogs, etc.), **legislative documents** such as laws, regulations, and standards (legal databases, case law databases, governmental websites, such as [EUR-Lex: EU law](#), the [UN Official Document System \(ODS\)](#), and other legal depositories) related to LNP. This data will be mainly **expressed in a textual or visual form in the usual formats** (.doc/.docx, .rtf, .pdf, .png, .jpeg/.jpg).

Additionally, **surveys** will be conducted to acquire **new primary data**. These will involve direct engagement with the public through data collection methods including online surveys, interviews and focus groups, some in person and some online. **Participants will be informed** about the nature of the research, their role in it, and the ways in which the data they provide will be stored and analysed. They will be provided with information about **how they can withdraw** from the study, the **timeframe** for doing this, and what would **happen to their data** after withdrawal. In reporting analysis from data collected through these mechanisms, participants will be **fully anonymised**. **Identifiers** (e.g. participant numbers) will be used in order **to categorise data** (e.g. by age group or location), but all reasonable care will be taken to ensure that the reader will not be able to identify the participants.

**Ethical approval for the online survey** of members of the public **was granted** by the University of Salford (**USAL**) on 5<sup>th</sup> August 2025 (Application reference 7362). Participants are provided with a **Participant Information Sheet** and asked to state that they understand the nature of the project, how their data will be used, and that they are 18 or over. This information is provided **in 8 EU languages**: Dutch, English, French, German, Italian, Polish, Portuguese, and Spanish.

The **literature review draws only upon publicly available data** (academic publications, websites and reports). No primary data is collected or analysed in this activity.

The data is **stored** in a common, secured project workspace – **SharePoint** (WP4 folder). Each partner involved is also responsible for storing the data in the secure storage provided by their organisation.

**Sensitive personal data** deserves special attention, and PLAN-B project implementation will **ensure** that data processing and management will **respect the GDPR provisions**. An **informed consent** form will ensure that the human participants voluntarily participate in this research activity within the project. When dealing with personal data, informed consent for data sharing and long-term preservation is included in surveys and explained to the participants. **Participants always need to consent to the management of the data produced.**

**All the data and Deliverables of WP4**, excluding personal data and other data protected under the current EU legislation, will be provided under **open access** and will be published on PLAN-B's website and [Zenodo](#) repository.

## WP5: Specific measures to assess, prevent and mitigate negative impacts on terrestrial biodiversity

The **purpose** of data collection and generation in **WP5** is to **identify** and **evaluate existing measures for the prevention and mitigation** of LNP to avoid significant disturbance to TBES. The project will **generate both qualitative and quantitative data**. **Qualitative data** will be gathered through **interviews** and **focus groups** with stakeholders, including citizens, environmental regulators, and practitioners. All the ethical approvals required will be obtained to comply with the GDPR. **Quantitative data** will be collected through **monitoring LP and NP levels** using specialised equipment such as luminance meters and noise measurement devices. In addition, quantitative measures of LNP will be **derived from spatial models**, which will be integrated with in situ measurements to enable spatially explicit assessment of mitigation measures.

The project will **reuse existing data** from **previous studies and reports** on LNP, leveraging data from international, national, and local guidelines, policies, and legal instruments, as well as other data. The origin of the data will **include both primary data** collected through project activities **and secondary data** sourced from literature reviews and existing databases. This **secondary data will also include spatial data products** from other PLAN-B WPs, and external environmental and infrastructural GIS layers, such as land cover. In addition, **local light environments at selected sites will be characterised using the Environmental Light Field (ELF) method**, which uses calibrated wide-angle digital images to quantify biologically relevant properties of the light field (e.g. radiance, spectral composition and contrast). **ELF data will be linked to model-based layers** in the WP5 **database** to provide ground-truth and context for mitigation analyses. The **expected size** of the data is not precisely known at this stage, but it will include **multiple datasets across different formats**, including .docx, .pdf, GeoTIFF (.tif) for raster outputs, GeoPackage (.gpkg) **and other GIS formats for vector data, CSV and XLSX (Excel) files** for tabular inventories of measures and scenarios, and **Python (.py, .ipynb) scripts** associated with the processing and modelling workflows.



Stakeholders such as environmental regulators, urban planners, conservationists, and researchers will find this **data useful for developing and implementing mitigation strategies and policies**. The data will benefit communities engaged in citizen science projects, contributing to enhanced public awareness and participation in environmental protection efforts.

**All data** produced and used in the project **will be discoverable with metadata**. Standard identification mechanisms such as Digital Object Identifiers (DOIs) will be employed to ensure data is **identifiable and locatable**. Naming conventions will follow a **structured format**, including project name, work package number, data type, and date. **Keywords** will be provided to **optimise re-use**, including terms such as "light pollution," "noise pollution," "biodiversity," and "mitigation measures."

Version numbers will be assigned to each dataset to track changes and updates. **Metadata will include details** such as creation date, data format, and a brief description of the dataset. Where applicable, existing metadata standards in environmental sciences will be adopted. In cases **where no standards exist**, a **custom metadata schema will be developed**, including essential fields for data description, context, and usage guidelines. Version control for processing workflows will be maintained using open repositories, ensuring that the generation of database layers from spatial layers and literature inventories is fully reproducible and transparently documented.

As a default, **data produced and used** in the project **will be made openly available**. **Exceptions** will be made for **datasets that cannot be shared due to legal, ethical, or contractual restrictions**. Open data will be deposited in certified repositories that support open access, such as institutional **repositories** provided by partner universities. Data will be **accessible via standard methods and software tools commonly used** in the field, with **necessary documentation included**.

**Documentation** will **cover software requirements**, usage **instructions**, and data **formats**. Access to data will be facilitated through **clear licensing and machine-readable licenses**, ensuring **transparency** in usage conditions. In cases of **restricted access**, a data access committee will oversee the provision of access and maintain records of data requests and approvals.

**Data** will be **licensed** under the **Creative Commons Attribution 4.0 International (CC BY 4.0) license to permit wide re-use**. Data will be made available for re-use immediately after validation and quality assurance processes are completed. Embargo periods may apply to protect intellectual property or pending publications, with clear timelines and reasons provided.

## WP6: PLAN-B measures to raise Awareness and Maximise Impact

The **purpose** of the **WP6 data** collection is to **communicate and disseminate PLAN-B's objectives, achievements and clustering activities** to its community, **encourage participation** in its activities and **be part of the CoPs** (Communities of Practice) in the project's pilot areas.

These data will only be **used by the PLAN-B consortium** for the purposes of delivering the PLAN-B project and will be managed in accordance with the data security and ethical considerations described in Sections 4 and 5 of this DMP.

This data will be collected and stored in accordance with current personal data protection regulations (GDPR), ensuring its anonymization where possible and the implementation of appropriate security measures to protect users' privacy.

This data includes:

- **User data:** Contact data (name, institution, email address), collected through registration forms and Google Forms.
- **Web Interaction data:** Information about user behaviour on the website, such as pages visited, time spent, captured through web analytics tools.
- **App collected data:** Measurements collected by users and participants in the citizen science activities proposed by the project through the app.

The **project's mobile app** will collect specific user data, including geolocation points and other usage data. To **comply with the GDPR**, the following measures will be implemented:

- **Informed Consent:** Before collecting any data, **explicit consent** will be obtained from users. The app will provide a clear description of the data to be collected, the purpose of the collection, and how it will be used.
- **Data Minimisation:** Only **data strictly necessary** for the project's purposes **will be collected**. Geolocation data will be captured only when essential for specific functionalities of the app.
- **Anonymisation and Pseudonymization:** Personal data, including geolocation points, will be **anonymised** or **pseudonymised** to protect users' identities. This will be achieved through techniques that unlink the data from direct identifiers.
- **Data Security:** Robust **security measures** will be implemented to **protect the collected data**. This includes encryption of data both in transit and at rest, user authentication, and access control to the data.
- **Access and Rectification:** Users will have the right to access their data, correct any errors, and request its deletion. The app will include functionalities to facilitate these rights.
- **Data Retention:** The collected data will be **stored only for the duration necessary to fulfil the project's objectives**. Once the project is completed, the data will be securely deleted.

These measures ensure that the processing of data collected by the mobile app complies with GDPR regulations, protecting users' privacy and rights.

Both web and app (including the citizen science app) interaction data will be analysed to assess the achievement of the project's communication objectives.

**PLAN-B's web hosting** is provided by the company OVH (<https://www.ovhcloud.com>), which guarantees the project's sovereignty with respect to the data hosted there.

The **SecNumCloud security visa**, obtained by OVHcloud at the beginning of 2021, offers customers of certified cloud services the guarantee that they benefit from solutions whose level of security and trust has been **verified by the French ANSSI** (National Agency for the Security of Information Systems).

*Table 2: Networking and Collaboration Contacts Database.*

Dataset name	Networking and Collaboration Contacts Database
Related WP and task	WP6 T6.1
Brief description of the data	Database compiling contact and descriptive information of European, national and regional projects, organisations, and professionals involved in light and noise pollution mitigation and related fields. Includes details such as project title, organisation, contact person (professional email), thematic focus, and type of collaboration. The dataset supports coordination and exchange of good practices between Horizon Europe projects.
Types of data	Text (contact information, project metadata), Categorical (collaboration status).
Form of data	Excel or CSV files stored on Ibercivis internal servers or PLAN-B SharePoint.
Format of data	.csv; .xlsx; .pdf (summaries).
Purpose of the data collection/generation, and data and results exploitation	To establish and manage networking and collaboration with relevant projects and stakeholders, avoiding duplication and enhancing project synergies within the Horizon Europe programme.
Size of the data	Approx. 2-5 MB total.
Origin of the data	Contacts provided by project partners and identified through open databases (e.g. Horizon Europe).
Which partner will use the data?	IBE (leader) and all partners for collaboration and dissemination coordination.
Data utility outside of the project	Aggregated information (not personal data) may be useful for future EU networking activities and reporting.
Re-use of existing data	Reuse of publicly available project contact data from Horizon Europe databases.

*Table 3: Communication and Dissemination Metrics Dataset*

Dataset name	Communication and Dissemination Metrics Dataset
Related WP and task	WP6, T6.2, T6.3
Brief description of the data	Dataset including aggregated communication metrics related to project dissemination activities, such as website analytics, newsletter subscribers, social media interactions, press releases, media coverage, and attendance at events or workshops. Dataset includes qualitative feedback where available.

Types of data	Numerical (counts, metrics), Text (titles, descriptions), Categorical (channels, audiences).
Form of data	Excel or CSV datasets; internal dashboards; PDF or PPT summary reports.
Format of data	.csv; .xlsx; .pdf; .pptx
Purpose of the data collection/generation, and data and results exploitation	To monitor and evaluate the impact and effectiveness of dissemination and communication activities within PLAN-B.
Size of the data	Approx. 10-20 MB total.
Origin of the data	Collected from website analytics tools, social media platforms, email campaign statistics, and event registration forms.
Which partner will use the data?	IBE (leader) and project coordination teams for monitoring KPIs and reporting to the EU.
Data utility outside of the project	Aggregated data and communication outcomes may be used by other EU projects as reference for good practices.
Re-use of existing data	No reuse; all data collected during the PLAN-B project duration.

*Table 4: FAIR data. Making WP6 data findable, including provisions for metadata.*

Which naming conventions will be used?	<YYYYMMDD>_<work package>_<description>_<type>_<organisation>_<version>
Will data be identified by a persistent identifier?	DOI for public datasets (aggregated data) on <a href="#">Zenodo</a> .
What metadata will be created?	Metadata standards: Dublin Core for descriptive data; internal indexing for contacts.
Will search keywords be provided?	Search keywords and version control will be applied following project conventions.
Clear versioning approach	The version will be identified in the name of each file. We will try to use <a href="#">Zenodo's</a> format in order for the data to be traceable and easily indexed.

*Table 5: Making WP6 data accessible.*

Will this dataset be made openly available and how?	Individual contact data will remain confidential and not publicly accessible.
Repository link(s)	Aggregated, anonymised datasets and communication reports will be made publicly available via <a href="#">Zenodo</a> (Light and noise Pollution Community) and the project website after validation.

Is an embargo period foreseen?	An embargo may apply until related results are published.
Protocols for access	The data will be accessible through a free and standardized access protocol.
Are there limitations and restrictions to access the data?	Individual citizen data anonymised.
Identification	User registration on the Lost at Night app for data upload.
Committee for accessibility	No specific committee foreseen.
Will metadata be made openly available?	Metadata will be openly available under CC0; data summaries under CC BY 4.0.
How long will the data remain available?	Data and metadata to remain permanently available.
What methods or software tools are needed to access the data?	Data accessible via standard software (Microsoft, GIS, R, Python).

*Table 6: Making WP6 data interoperable.*

What data and metadata vocabulary, standards, formats or methodologies will you follow to make your data interoperable?	Aggregated metrics follow standard indicators (visits, reach, engagement).
---	--

*Table 7: Increase WP6 data re-use.*

How will you provide documentation needed to validate data analysis and facilitate data re-use?	Documentation (readme files, data dictionaries, methodological notes) will accompany public datasets.
How will the data be licensed to permit the widest re-use possible?	All non-confidential data will be made freely available via the <a href="#">Zenodo</a> repository and institutional repository to allow the widest re-use possible. They will be licensed under Creative Commons Attribution 4.0 (CC BY 4.0).
When will the data be made available for re-use (expected date)?	Communication metrics and summaries will remain permanently available for reuse by EU institutions and projects.
Will the data produced in the project be usable by third parties, in particular after the end of the project?	Communication metrics and summaries will remain permanently available for reuse by EU institutions and projects.
How long is it intended that the data remains re-usable?	All open data will remain permanently available and usable.

Describe all relevant data quality assurance processes.	Quality assurance ensured by Ibercivis and collaborating partners.
---	--

## WP7: Coordination and management

For the **purposes** of project **coordination and management**, a **set of data will be produced, such as documents, emails, and chat conversations**. These data will be generated and collected only for management purposes. No diffusion of such data is envisioned outside the consortium.

The project coordinator - **UGent** - is **responsible for the secure storage** of the project management data. There are various storage options at Ghent University, as outlined in *Figure 1*. These include **data storage for researchers on network drives ('Shares')** and via **O365 services**, and **storage on the HPC** (High Performance Computing Infrastructure).

Figure 1: Overview of data storage options available through UGent to support PLAN-B activities.

	Microsoft 365		SharePoint	Shares		HPC (only showing temp storage)	
Name	OneDrive FB	SharePoint Online - Teams	UGent SharePoint	H:/	Project Share (act & bulk)	HOME/DATA/S CRATCH	VO_DATA/VO_SCRATCH
Location	Cloud	Cloud	UGent	UGent	UGent	UGent	UGent
Size	3TB per user	25 TB per site/Team	10 GB per site	5 GB	GB-TB range	25 GB	GB-TB range
Backup mechanism	Versioning	Versioning	Versioning	snapshots	snapshots	none	none
Sync to local	✓	✓					
Confidential	✓*	✓*	✓	✓	✓	✓	✓
Intended use	Personal	Sharing	Sharing	Personal	Sharing	Personal	Sharing
Sharing external users?	✓	✓	✓				✓
Suitable for TB datasets					✓		✓
Data in use	✓	✓	✓	✓	✓	✓	✓
Data at rest					✓		

An appropriate shared workspace has been created by the Project Coordinator, through a **SharePoint workspace**. This workspace is used as a shared repository to facilitate the accessibility and exchange of the whole PLAN-B documents. SharePoint is a web-based collaborative platform seamlessly integrated with Microsoft 365, for which UGent holds a full professional license. The access right to documents and folders can be regulated at the single-item level. The Ghent University SharePoint instance is set up 'on premise', meaning that the data are stored within Ghent University.

The European Union's General Data Protection Regulation (**GDPR**) and other EU data protection legislation set an important bar globally for data protection rights, information security, and compliance. Microsoft claims to comply with these regulations.<sup>5</sup>

**SharePoint is complemented by a backup process** including all the data stored (emails, documents, calendar and contacts).

The **project mailing lists are stored and archived on a dedicated server at UGent** headquarters. Backup of the servers is made on a regular basis. The mailing list archives are solely accessible by the mailing list subscribers.

The **chatting tool based on the Microsoft Teams** chatting service has been provided for all the consortium members. The tool ensures secure data storage outside Ghent University. Therefore, no confidential or sensitive data is being shared through this service.

The information on how PLAN-B handles project management is provided within the [Project Management Handbook \(Deliverable 7.1\)](#).

## 7. Data Management Risks and Mitigation Solutions

Table 8: Data Management Risks and Mitigation Solutions.

Component	Description	Extended Threats (Low Frequency/Low Risk)
<b>Threat Event Frequency</b>	How often threats are expected to occur.	<b>Infrequent Threats:</b> Some threats may occur rarely due to robust security measures, low attractiveness to attackers, or infrequent access to sensitive data. For example, highly specialised attacks requiring significant resources and expertise may occur less frequently.
<b>Contact Frequency</b>	Frequency of interaction between threat agents and data assets.	<b>Rare Interactions:</b> Low-contact frequency reduces exposure to threats. For instance, data that is accessed infrequently or stored in highly secure environments has fewer opportunities for threat agents to interact with it, lowering the risk of compromise.
<b>Probability of Action</b>	Likelihood that a threat agent will take action against an asset.	<b>Low Probability of Action:</b> Certain threat agents may have little motivation to target specific data due to its low value, high risk of detection, or strong legal deterrents. For example, data with minimal financial or strategic value is less likely to be targeted by sophisticated attackers.

<sup>5</sup> <https://www.microsoft.com/licensing/docs/view/Microsoft-Products-and-Services-Data-Protection-Addendum-DPA>

<b>Vulnerability</b>	Susceptibility of data assets to threats.	<b>Low Vulnerability:</b> Data assets with strong security measures, regular updates, and limited access points are less susceptible to threats. For example, systems with up-to-date patches and robust encryption face lower risk of exploitation.
<b>Threat Capability</b>	Skill and resources available to threat agents.	<b>Limited Threat Capability:</b> Threat agents with low skill levels or limited resources are less capable of successfully attacking well-protected data. For instance, amateur hackers or low-level insider threats may not possess the necessary tools or knowledge to bypass advanced security systems.
<b>Loss Event Frequency</b>	How often loss events are expected to occur.	<b>Rare Loss Events:</b> Well-protected environments and robust incident response plans reduce the likelihood of loss events. For instance, comprehensive backup strategies and disaster recovery plans can minimize the impact and frequency of data loss events.
<b>Primary Loss Magnitude</b>	Direct losses resulting from an event.	<b>Minimal Direct Losses:</b> Some data breaches may have limited financial impact due to the nature of the data. For example, losing non-critical, anonymized data results in minimal direct losses compared to highly sensitive personal information or intellectual property.
<b>Secondary Loss Event Frequency</b>	How often do secondary loss events occur due to the primary event?	<b>Infrequent Secondary Events:</b> Primary events that are well-contained and quickly addressed are less likely to lead to secondary loss events. For example, immediate detection and remediation of a minor data breach can prevent further exploitation or cascading failures.
<b>Secondary Loss Magnitude</b>	Indirect losses, such as reputational damage or legal costs.	<b>Minor Secondary Losses:</b> If the data compromised is not sensitive or critical, the secondary losses such as reputational damage or legal costs may be minimal. For instance, an incident involving non-sensitive internal reports is less likely to attract legal penalties or media attention compared to a breach of customer personal data.

Table 9: Asset Valuation

Component	Description	Extended Threats (Low Frequency/Low Risk)
<b>Criticality</b>	Effect on the organization's productivity.	<b>Low Productivity Impact:</b> Data that is not essential to daily operations has a low impact on productivity if



		compromised. For example, archived data or non-critical research data can be lost without significantly affecting current project workflows.
<b>Cost</b>	Bare cost of the asset and cost of replacing a compromised asset.	<b>Low Replacement Cost:</b> The cost to replace or restore some data can be minimal. For example, data that is regularly backed up and easily restored has low associated costs if compromised.
<b>Sensitivity</b>	Cost associated with the disclosure of the information, including embarrassment, competitive advantage loss, legal/regulatory costs, and other general losses.	<b>Low Sensitivity:</b> Data that does not contain sensitive or personally identifiable information (PII) poses a low risk if disclosed. For example, general project management information or public data sets have limited sensitivity and thus lower associated risks if exposed.

Table 10: Threat Actions

Component	Description	Extended Threats (Low Frequency/Low Risk)
<b>Access</b>	Unauthorized reading of data.	<b>Limited Unauthorized Access:</b> Strict access controls and regular monitoring reduce the risk of unauthorized access. For instance, using strong authentication methods and role-based access control (RBAC) limits unauthorized access attempts and their success.
<b>Misuse</b>	Improper use of data.	<b>Minimal Data Misuse:</b> Clear policies and regular audits minimize improper use of data. For example, comprehensive user training and regular audits ensure that data is used as intended, reducing the risk of misuse.
<b>Disclose</b>	Unauthorized sharing of data.	<b>Low Disclosure Risk:</b> Data that is not highly confidential or sensitive has a lower risk if disclosed. For example, general project data or already publicly available information poses minimal risk if inadvertently disclosed.
<b>Modify</b>	Unauthorized changes to data.	<b>Low Modification Risk:</b> Robust integrity controls and regular audits ensure data accuracy. For instance, implementing checksums and regular data integrity audits can quickly detect and rectify unauthorized changes, minimizing the risk.
<b>Deny Access</b>	Preventing legitimate access to data.	<b>Low Denial Risk:</b> Redundant systems and robust access controls ensure continued availability. For example, having multiple backup systems and disaster recovery plans ensures that data remains accessible even during an attack, reducing the risk of denial of access.

Table 11: Types of Loss

Component	Description	Extended Threats (Low Frequency/Low Risk)
<b>Productivity</b>	Reduction in the organisation's ability to effectively produce goods or services.	<b>Low Productivity Impact:</b> Data that is not essential to daily operations results in minimal productivity loss if compromised. For example, archived records or supplementary project data can be lost without significantly impacting ongoing activities.
<b>Response</b>	Resources spent addressing data breaches or losses.	<b>Low Response Costs:</b> Effective incident response plans and minor breaches result in minimal resource expenditure. For example, minor incidents that are quickly detected and resolved with automated tools have low associated response costs.
<b>Replacement</b>	Costs to restore lost or corrupted data.	<b>Low Replacement Cost:</b> Regular backups and simple data structures result in minimal restoration costs. For example, routine backups of non-critical data ensure that restoration efforts are straightforward and inexpensive.
<b>Fines and Judgments (F/J)</b>	Legal expenses resulting from data breaches.	<b>Minimal Legal Penalties:</b> Data with low sensitivity and no regulatory implications results in low legal costs. For example, breaches involving non-sensitive data that do not violate any laws or regulations have minimal associated legal expenses.
<b>Competitive Advantage (CA)</b>	Loss of research opportunities or intellectual property.	<b>Low Competitive Impact:</b> Data with limited strategic value results in minimal loss of competitive advantage. For example, generic research data or publicly available information poses little risk of competitive disadvantage if compromised.
<b>Reputation</b>	Damage to the project's and involved institutions' reputations.	<b>Minor Reputational Damage:</b> Incidents involving non-sensitive, non-public-facing data result in minimal reputational harm. For example, internal project data breaches that do not involve customer or partner information are less likely to affect public perception.