



DATA MANAGEMENT PLAN

DELIVERABLE 7.2



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	01/10/2024
Duration	48 months
Coordinator	prof. Geert Van Hoorick, Ghent University
Contact	Yana Yakushina
Funded by the European Union	
Deliverable details	
Description:	D7.2 – Data Management Plan (DMP)
Version	0.1
Dissemination level	public
Due date	M6 (30/06/2024)
Lead:	Ghent University
Author(s):	Yana Yakushina (UGent), Michael D. Wood (USAL), Reinhard Klenke (MLU), Paweł Tysiąc (Gdansk Tech)
Contributors:	Dick Botteldooren (UGent), Timothy Van Renterghem (UGent), Shreyas Rangaswamy Chandrashekar (UGent), Alejandro Sanchez de Miguel (UCM), Daniel Lisbona (IBE), Graeme Sherriff (USAL), Catherine L. Barnett (UKCEH), Penelope Vounatsu (Swiss TPH)
Reviewers:	Geert Van Hoorick (UGent) and Els Vermeulen (UGent)

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 managing 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. It will be updated on an ongoing basis throughout the project and revised versions of the DMP will be provided by M24 and M48.

During the first six months of the project, the main elements of the Data Management Plan were collected, with the identification of the main datasets produced by the different project's tasks. The responsibilities for the managing and processing of each dataset were identified, and the measures to comply with GDPR and other rules were set up.

The DMP includes information on:

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

This deliverable (Deliverable 7.2 of the PLAN-B project) is based on the Horizon Europe generic template for DMPs.





Table of contents

Ex	xecutive summary	3
Lis	st of Tables	5
Lis	st 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	8
	Data utility	9
3.	FAIR DATA	9
	Making Data Findable	9
	Making Data Openly Accessible	9
	Making Data Interoperable	10
	Increase Data Re-use (through clarifying licences)	10
4.	Data security	11
5.	Ethical aspects	11
6.	Data Management Plan per Work Package	11
	WP 1: Mechanism of light and noise pollution impacts	11
	WP2: Modelling spatio-temporal impacts at regional scale	13
	WP3: European situational analysis	18
	WP4: Societal drivers of light and noise pollution	19
	WP5: Specific measures to assess, prevent and mitigate negative impacts on terrestrial biodiversity	20
	WP6: PLAN-B measures to raise Awareness and Maximise Impact	21
	WP7: Coordination and management	22
7.	Data Management Risks and Mitigation Solutions	24





List of Tables

Table 1: Secondary data sources planned to be used on a European scale in development of the PLAN-B spat	io-
temporal database	14
Table 2: Overview of data storage options available through UGent to support PLAN-B activities.	23

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)
ОА	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
Т	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 the PLAN-B's Data Management Plan (DMP). It 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.

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 first version of Deliverable 7.2 of the PLAN-B project, is the position at month 6 (M6) of the project. There are formal update points for the DMP in M24 and M48. However, it will also be updated on an ongoing 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 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. That '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 that '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)."¹ The Recommendations also cover "other research-relevant digital objects from public funding" - "metadata, algorithms, workflows, models, and software.²

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

This Data Management Plan will adhere to the definitions provided by the OECD and will comply with all binding and non-binding requirements for research data management.

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



¹ 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:

² Ibid.



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".3

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.

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, etc.), BibTex, GPKG (Maps) etc., and all PLAN-B written deliverables will be compiled and made available as PDF files.

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://zenodo.org, https://zenodo.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.

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



plan-b-project.eu

³ Oxford Dictionary. Available online:



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

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 and linked via unique digital object identifiers (doi). 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) to maximise the findability and utility to our data products.

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.

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, organizations 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 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.⁴

Given the nature of the PLAN-B consortium, the DMP will facilitate interoperable access and use of data between researchers, organizations 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.

⁴ 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



1



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 organization.

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 only treated 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.

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).

WP 1: 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 light and noise pollution impacts on terrestrial biodiversity and ecosystem services (TBES). Secondary data collation, in the form of a knowledge base, and subsequent meta-analysis form the major data products from WP1.





The knowledge base will collate published and unpublished ('grey' literature data) on the impacts of light and noise pollution on TBES. For transparency of data provenance, full bibliographic 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 encourage the originators to make those data accessible and provide guidance to facilitate this.

Data within the knowledge base will be compiled through Systematic Review, which will be conducted by a sub-group within the PLAN-B consortium. This knowledge base sub-group will also be responsible for input of compiled data into the knowledge base.

During development and testing (and for population with data from PLAN-B WP1 consortium members), a copy of the knowledge base will be hosted on the UGent PLAN-B's SharePoint site in the dedicated WP1 folder. For population with data from the Systematic Review undertaken within WP1, a second copy of the knowledge base will be used and this will be hosted on the UKCEH PLAN-B's SharePoint site. Upon completion of these WP1 activities the two knowledge bases will be combined. The combined output will then 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 sub-group. It will also ensure continuous secure back-up of the knowledge base as it develops.

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

- Using a unique ID for each member of the knowledge base sub-group so the individual responsible for inputting each data row is identifiable.
- Duplicate entries (during initial knowledge base population) of a sub-set of publications by different members of the knowledge base sub-group and subsequent comparison of the entered data. This will ensure that all members of the knowledge base sub-group are able to develop a common understanding of and approach to extracting data from other sources and inputting into the knowledge base.
- A quality control process for the final knowledge base in which a proportion (e.g. 10%) of randomly selected entries are checked by a member of the PLAN-B team who wasn't responsible for the initial input of the respective entry.

Upon completion of the knowledge base, the final version will be made available via an appropriate open access data repository (such as https://eidc.ac.uk/). Supporting metadata will be provided and it is likely that a 'data paper' will also be published to accompany the knowledge base.

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 construction of the PLAN-B spatio-temporal database is shown in Table 1. These data will be used to model spatiotemporal impacts of light, noise and other drives 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 PM2.5, PM10 and NO2 available at Siss TPH at 1km2 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, 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. But will be. 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

Description: <u>Biodiversity</u> <u>Source:</u> <u>Global Biodiversity Information (GBIF) network;</u> <u>Time period:</u> 1992-2024; <u>Spatial Resolution:</u> Points (Presence data); <u>Temporal Resolution:</u> monthly

Description: Conservation status Source: IUCN Red List of Threatened species; Time period: 1992-2024; Spatial Resolution: Points (Presence data); Temporal Resolution: annual

Description: Species Habitat Index (SHI), Species Protection Index (SPI), Species Status Information Index (SSII)
Source: GeoBon; Time period: 2001-2020; Spatial Resolution: 1km x 1km; Temporal Resolution: annual

Description: Air temperature, Total precipitation Source: ERA5-Land Monthly Averaged – ECMWF Climate Reanalysis; Time period: 1992-2024; Spatial Resolution: Points; Temporal Resolution: monthly

Description: Extreme precipitation Source: ERA5-Land Daily – ECMWF Climate Reanalysis; Time period: 1992-2024; Spatial Resolution: 11.132 x 11.132 km; Temporal Resolution: daily

Description: Palmer drought severity index (PDSI) Source: TerraClimate; Time period: 1992-2024; Spatial Resolution: 4.638 x 4.638 km; Temporal Resolution: monthly

Description: Land cover Source: Corine Land Cover, Copernicus Land Monitoring Service; **Time period:** 2006, 2012, 2018, 2022; **Spatial Resolution:** 100 x 100 m; **Temporal Resolution:** annual

Description: Air pollution Source: Bayesian geostatistical model-based estimates; **Time period:** 2006-2022; **Spatial Resolution:** 1 x 1 km; **Temporal Resolution:** annual

Description: Soil pollution Source: LUCAS database; Time period: 2009, 2012, 2015, 2018, 2022; Spatial Resolution: Points; Temporal Resolution: annual

Description: Population density Source: Global Human Settlements layers; Time period: 1990, 2000, 2015, 2025, 2030; Spatial Resolution: 1 x 1 km; Temporal Resolution: 5-year-steps

Description: Specially Protected Areas (SPAs) for birds, adopted Sites of Community Importance (SCIs), and Special Areas of Conservation (SACs) for habitats and other species Source: Natura2000; Time period: 1992-2021; Spatial Resolution: 1:100'000; Temporal Resolution: annual

Description: Macroalgae, Angiosperms, Fungi, Reptiles, Amphibia, Invertebrates, Fish, Plants, Mammals, Birds, Macrophytes, Phytoplankton, Plant community Source: EASIN data; Time period: 2014-2021; Spatial Resolution: Points (Presence data); Temporal Resolution: annual

Description: <u>Habitat Conservation Source:</u> Wild Bird Indicator; **Time period:** 2018-2022; **Spatial Resolution:** Points (Presence data); **Temporal Resolution:** annual

Description: <u>Habitat Connectivity Source:</u> wildlife corridor network, Switzerland; **Time period:** 2021; **Spatial Resolution:** Polygon Data; **Temporal Resolution:** annual





Data related to light pollution modelling and mapping.

The "Atlas of Colour Light Pollution Map of Europe" aims to create a comprehensive collection of maps depicting Europe at night using satellite data. This map leverages imagery from the International Space Station (ISS) and the SDGSAT-1 satellite to trace and analyse light pollution emissions across the continent. The lead beneficiary - Universidad Complutense de Madrid (UCM) is responsible for generating new observational data, which will be stored and managed according to the following data management plan.

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: The data is sourced from ISS images and SDGSAT-1 satellite images.
- **Purpose**: The primary objective is to trace and analyse light pollution emissions, contributing valuable insights into environmental and urban planning initiatives.
- New/Reused: The data generated for this project is entirely new.

Data Specifications

- Digital/Non-Digital: The data is digital.
- Data Type: Observational 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**: The data will be stored on hard disks and network drives to ensure accessibility and security during the project lifecycle.
- **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. But will be. Either via Zenodo or the
 publication. 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: Access to the data after the project's conclusion will be restricted.
 Interested parties will need to obtain a contract and approval to access the data, ensuring that sensitive information and usage terms are adequately managed. Mainly, SDGSAT-1 data have specific usage restrictions, as only non-commercial use is allowed currently. There are no restrictions on the usage of the ISS data produced.

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.
 - Data related to noise pollution modelling and mapping.

The "European natural areas noise exposure maps" aims 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 data management plan.

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: Vegetation/woods based on "Dynamic World", building footprints based on "Microsoft's Planetary Computer datasets", and other ground types than vegetation 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)"
- Meteorological data: using "ECMWF Climate Reanalysis" and "ECV Inventory"

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. But will be. Either via Zenodo or the
 publication. 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 light and noise pollution with regard to biodiversity, (2) to perform special statistics about influence and increasing impact of light and noise pollution 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 about light and noise pollution itself, drivers of light and noise pollution (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 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.org, pangea.de or 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, we will use widespread well-





documented and open formats and structures 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 have to be paid. The German Research Foundation (DFG) gives financial support for Open Access publication of results to institutions. The repositories www.zenodo.org and pangea.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 light and noise pollution; (2) to understand human behaviour, impacting light and noise pollution and (3) to suggest the best social and regulatory solutions for mitigation light and noise pollution impacts on terrestrial biodiversity and ecosystems.

The data management plan 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 <u>EUR-Lex: EU law</u>, the <u>UN Official Document System (ODS)</u>, and other legal depositories) related to light and noise pollution. 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.

The data is stored in a common secured project workspace - SharePoint. Each partner involved is also responsible for storing the data in the secured storage provided by their organisation.

Sensitive personal data deserve special attention and PLAN-B project implementation will ensure that data processing and management will respect the General Data Protection Regulation (GDPR) provisions. An informed consent form will ensure 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 give their consent to the management of the data produced.

All the data and Deliverables of the WP4, excluding personal data and other data protected under the current EU legislation, will be provided under open access and will be published on the 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 light and noise pollution to avoid significant disturbance to terrestrial biodiversity and ecosystem services (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. Quantitative data will be collected through monitoring LP and NP levels using specialized equipment such as luminance meters and noise measurement devices.

The project will reuse existing data from previous studies and reports on light and noise pollution, leveraging data from international, national, and local guidelines, policies, and legal instruments. The origin of the data will include both primary data collected through project activities and secondary data sourced from literature reviews and existing databases. 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.

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 optimize 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.





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 Minimization**: 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.
- Anonymization and Pseudonymization: Personal data, including geolocation points, will be anonymized or pseudonymized 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 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).

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).





Table 2: Overview of data storage options available through UGent to support PLAN-B activities.

	Microso	oft 365	SharePoint	Sh	ares	HF (only showing	
Name	OneDrive FB	SharePoint Online - Teams	UGent SharePoint	H:/	Project Share	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?	✓	\checkmark	✓				✓
Suitable for TB datasets					✓		✓
Data in use	✓	\checkmark	✓	✓	✓	✓	✓
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.⁵

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)*.

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



_



7. Data Management Risks and Mitigation Solutions

Component	Description	Extended Threats (Low Frequency/Low Risk)
	How often threats are expected to occur.	Infrequent Threats: Some threats may occur rarely due to robust security measures, low attractiveness to attackers, or infrequent access to sensitive data. For example, highly specialized attacks requiring significant resources and expertise may occur less frequently.
	Frequency of interaction between threat agents and data assets.	Rare Interactions: 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.
1	Likelihood that a threat agent will take action against an asset.	Low Probability of Action: 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.
Vulnerability	Susceptibility of data assets to threats.	Low Vulnerability: 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.
	Skill and resources available to threat agents.	Limited Threat Capability: 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.
	How often loss events are expected to occur.	Rare Loss Events: 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.
_	Direct losses resulting from an event.	Minimal Direct Losses: 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.
Loss Event	How often do secondary loss events occur due to the primary event?	Infrequent Secondary Events: 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.
Loss	Indirect losses, such as reputational damage or legal costs.	Minor Secondary Losses: 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.

Asset Valuation

Component	Description	Extended Threats (Low Frequency/Low Risk)
Criticality	Effect on the organization's productivity.	Low Productivity Impact: 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.
Cost		Low Replacement Cost: 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.
	of the information, including embarrassment, competitive advantage loss, legal/regulatory	Low Sensitivity: 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.

Threat Actions

Component	Description	Extended Threats (Low Frequency/Low Risk)
Access	Line and a second second second alega	Limited Unauthorized Access: 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.
Misuse	Improper use of data.	Minimal Data Misuse: 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.
Disclose	Unauthorized sharing of data.	Low Disclosure Risk: 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.
Modify	Unauthorized changes to data.	Low Modification Risk: 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.
Deny Access	Preventing legitimate access to data.	Low Denial Risk: 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.

Types of Loss

Component	Description	Extended Threats (Low Frequency/Low Risk)
Productivity		Low Productivity Impact: 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.
Response	Resources spent addressing data breaches or losses.	Low Response Costs: 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.
Replacement	Costs to restore lost or corrupted data.	Low Replacement Cost: 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.





Fines and Judgments (F/J)	Legal expenses resulting from data breaches.	Minimal Legal Penalties: 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.
-	Loss of research opportunities or intellectual property.	Low Competitive Impact: 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.
Reputation	Damage to the project's and involved institutions' reputations.	Minor Reputational Damage: 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.