bioagora

Bio Knowledge Agora: Developing the Science Service for European Research and Biodiversity

D6.7 The 2nd Data Management Plan

Date of delivery – 27/06/2025

Maria Söderholm and Kaisa Korhonen-Kurki
(Finnish Environment Institute)

Credit: Teemu Helonheimo/ SYKE

Funded by the European Union

DOCUMENT TRACKS DETAILS

Project acronym	BioAgora
Project title	Bio Knowledge Agora: Developing the Science Service for European Research and Biodiversity Policymaking
Starting date	01/06/2022
Duration	60 months
Call identifier	101059438
Grant Agreement No	101008626

Deliverable Information	
Deliverable number	D31
Work Package number	WP6
Deliverable title	The 2nd Data Management Plan
Lead beneficiary	Finnish Environment Institute (Syke)
Author(s)	Maria Söderholm & Kaisa Korhonen-Kurki (Finnish Environment Institute)
Due date	30/06/2025
Actual submission date	27/06/2025
Type of deliverable	DMP
Dissemination level	PU (Public)

VERSION MANAGEMENT

Revision table			
Version	Revision	Date	Description
1	Maria Söderholm / Finnish Environment Institute	13/02/2025	First draft
2	Kaisa Korhonen-Kurki / Finnish Environment Institute	25/02/2025	1st internal review

	BioAgora WP leaders and co-leaders	5/6/2025	2nd internal review
3	Maria Söderholm / Finnish Environment Institute	27/06/2025	Update following reviews

LIST OF ACRONYMS AND ABBREVIATIONS

Acronym / Abbreviation	Meaning / Full text
BDS 2030	EU Biodiversity Strategy 2030
KENs	Knowledge Exchange Networks
DMP	Data Management Plan
DoA	Description of the Action
DOI	Digital Object Identifier
EC	European Commission
FAIR	The principles related to Findable, Accessible, Interoperable and Re-usable data
GDPR	General Data Protection Regulation
KCBD	Knowledge Centre for Biodiversity
RIA	Research and Innovation Action
SSBD	Science Service for Biodiversity
URN	Uniform Resource Name
WP	Work Package

Table of contents

1. Project description

BioAgora is a collaborative European project funded by the Horizon Europe programme. As a Research and Innovations Action RIA, project BioAgora brings together 22 leading organisations (see Appendix 1) to connect biodiversity research with policy needs through targeted dialogue between scientists, knowledge holders, and policymakers.

The primary outcome will be the development of a Science Service for Biodiversity (SSBD). This new service will support the ecological transition required by the European Green Deal and the European Union's Biodiversity Strategy for 2030 (EU BDS). By strengthening interactions between science and policy, the project seeks to enhance the implementation of the EU BDS and improve the integration of expert knowledge into EU policymaking and its execution.

Through close collaboration with a broad network of stakeholders, BioAgora strengthens the link between scientific knowledge and policymaking. It makes biodiversity research more accessible to policymakers through a coordinated set of actions that form the foundation for the development of the Science Service. These actions include:

- Assessing the current landscape of science-policy interfaces, evaluating existing policy tools and analysing biodiversity knowledge across Europe.
- Engaging a wide range of actors to build a network of experts capable of responding to requests directed to the Science Service.
- Testing the Science Service through case studies called Knowledge Exchange Networks (KENs).
- Developing and piloting a governance model for the Science Service in real world context.
- Synthesising policy-relevant knowledge from sources within and outside the European Commission.
- Supporting the EC's Knowledge Centre for Biodiversity (KCBD) by assisting in the setup of a ticketing system for EC staff to submit requests.
- Providing capacity building to empower the EU decision-makers to drive transformative change for biodiversity.

2. Objectives of the Data management Plan

This data management plan (DMP, D31) outlines the data, other outputs and data management practices of the BioAgora project. Its primary focus is on data and outputs produced by the project. Additionally, the DMP provides an overview of the Science Service for Biodiversity (SSBD) data management (see Appendix 2). As a living document, the DMP is continuously updated to reflect ongoing identification of data and other outputs, as well as the definition and refinement of management practices that support the project's data and output management activities.

BioAgora is committed to good scientific conduct, sound governance and compliance with applicable legislation. In the production and management of data and other outputs, we adhere to the guidelines issued by the ALLEA, specifically the European Code of Conduct for Research Integrity. Personal data are processed in accordance with the European General Data Protection Regulation (GDPR). Furthermore, our activities align with the guidelines of the Horizon Europe programme and the FAIR data principles.

The aim of this DMP is to describe as accurately as possible:

- The data collected and used by each work package
- The practices in place to ensure compliance with the FAIR principles, GDPR and ethical guidelines
- The data to be shared with project partners
- The data to be published and how it will be made publicly available

3. Summary of data

Achieving BioAgora's objectives requires a broad and diverse knowledge base. From the outset, the project has identified its data and information needs: specifying what new and existing data and information are necessary to meet these objectives, and how to locate and utilize them effectively. BioAgora generates new data and leverages existing data through synthesis and analysis, providing all project participants with a solid knowledge foundation to develop the Science Service for Biodiversity (SSBD).

There is a wealth of data and information produced by various actors and projects, which BioAgora will fully exploit. A key approach of the project is to engage a wide range of stakeholders from the biodiversity community, involving them in both the generation and application of data and knowledge to support the development of the SSBD.

Additionally, BioAgora project will develop specific functions, tools and mechanisms to foster dialogue between the research community and decision-makers, enhancing the dissemination and practical use of biodiversity knowledge. Some of these tools are developed together with the Knowledge Centre for Biodiversity (KCBD) and include, for example:

- A support tool integrated to the SSBD to handle biodiversity related requests originated from the KCBD system, submitted by EU level policy and decision-makers.
- A mechanism to harvest and rationalise knowledge generated by EU-funded projects, initiatives, and platforms.
- An interactive and widely accessible web-platform with related functionalities for the SSBD.

This chapter summarises, in much detail as possible, the data generated and used by the work packages (see Table 1.). Details concerning the data management of the Science Service web platform are presented in Appendix 2.

Table 1. Detailed description of the data generated and used in BioAgora.

WP1: Underpinning the Biodiversity Strategy – Case Study Hub				
Task	**Title of data produced by the project**	**Description**	**Publishing / reporting channel**	**Reasons for not publishing the data and limiting access**
T1.1	Interviews with the biodiversity policy experts and policy makers	Data consists of voice and/or video recordings, text documents prepared as anonymized summaries. Formats: .mp3, .mp4 (recordings), .docx, .txt (transcripts), .docx, .txt, .pdf, .xlsx (aggregated/result data)		The .mp3 and .mp4 files are deleted after transcription, in accordance with the GDPR-compliant consent form. The full transcripts contain confidential information. Deleting the sensitive parts would render the data unintelligible, and therefore the transcripts will not be shared openly.
T1.1	Interviews with EC officers on BDS implementation challenges	Data consists of voice and/or video recordings, text documents prepared as anonymized summaries. Formats: .mp3, .mp4 (recordings), .docx, .txt (transcripts), .docx, .txt, .pdf, .xlsx		The .mp3 and .mp4 files are deleted after transcription, in accordance with the GDPR-compliant consent form. The full transcripts contain confidential information. Deleting the sensitive parts would render the data unintelligible, and therefore the transcripts will not be shared openly.
T1.1	Expert consultations with BioAgora project partners	Data based on face-to-face workshop organised at the kick-off meeting, and online consultation using the shared Teams folder. Formats: text documents (.txt, .docx)		The data contains personal information. Removing both direct and indirect identifiers would result in the loss of crucial context, making the data unusable for its intended purpose.
T1.1	Non-systematic literature review of challenges that hinder the	Data summarises the results of a non-systematic literature review as well as expert opinion shared on the initial results by project consortium members.	Zenodo: https://zenodo.org/records/7685109	

		implementation of specific targets of the BDS 2030	Data is based on published scientific literature and grey literature such as project reports, publicly available information on websites. Formats: .docs, html, .pdf, .txt, . xlsx		
	T1.2	Mapping the knowledge base on selected BDS challenges addressed by the Knowledge Exchange Networks	Data consists of scientific literature synthesised using appropriate methods such as literature reviews, rapid evidence assessment. Formats: .txt, .docx, .xlsx	TBC	
	T1.2	Urgent requests	This dataset includes a record of all urgent requests received and processed by the DCs. It details the requester, the BioAgora focal points involved, submission and delivery dates, the title of the request, the type and scope of the deliverable. Formats: text documents (.txt, .docx)		We adhere to strict data protection practices in accordance with the GDPR for all personal data. Consequently, access to the data is restricted to justified use only, and data containing personal information are, in principle, not made publicly available.
	T1.2	Actors of the topical networks	This dataset provides information on organizations and networks that operate at the science-policy-society interface and are relevant to the thematic focus of Task 1.2. It includes knowledge holders, users, brokers, and funders who may contribute to various SSBD functions, such as responding to policy requests, horizon scanning, and setting research priorities. The dataset captures the name of the organization/network, website,	TBC	Only anonymised data can be made public. Contact information and other potentially sensitive details will be removed before publication.

		type and level of network, contact information for the entry point, and details of their involvement in DC activities. Formats: Spreadsheet (.xlsx)		
	Experts involved in answering requests	This dataset includes the names, affiliations, email contacts, and areas of expertise of people invited to participate in expert groups formed to address urgent policy requests, as well as those involved in related dialogues or workshops. This includes scientists, policymakers, practitioners, NGOs, businesses, and citizen scientists. Formats: Spreadsheet (.xlsx)		The dataset contains personal data that cannot be fully anonymised without compromising its interpretability. If effective anonymisation becomes feasible in the future, data publication may be considered.
	Research priorities	This dataset includes a list of priority research topics and questions identified for their importance and potential impact on the EU Biodiversity Strategies. Formats: text documents (.txt, .docx)	TBC	
T1.2 /T1.3	Spatial data generated for Knowledge Exchange Networks ▪ Refers to processes and activities related to biodiversity management in areas such as the pollination, freshwater, nature-based solutions and urban green infrastructure. ▪ Freshwater – data on river systems, existing barriers and plans to reach 25.000 km free flowing rivers. Connectivity to floodplains and plans to extend or reinstall meandering wings. ▪ Pollination ▪ Marine ▪ Monitoring and scenarios ▪ Transformative change		TBC	

	▪ Formats: .geotiff., .shp, xlsx, .docx, arcgis, open street map			
T1.3	Landscape DC: Network of actors responding to policy requests and invited to workshops/dialogues	Data describes the network of actors/organisations that can be activated to answer policy requests and/or to be invited to workshops/dialogues. Formats: .xlsx		Contact information was collected solely for project-related communication and development. Due data protection considerations, consent for publishing this information was requested. Therefore, the data is restricted to internal project use only.
T1.4	Information on possible future research pathways	Data acquired through expert consultations, focus groups, surveys, scenario analyses, and other methods. Formats: .mp3, .mp4 (recordings), .docx, .txt (transcripts), .docx, .txt, .pdf, . xlsx		Due to GDPR-compliant data protection practices, access to personal data is restricted to justified use only and is not made publicly available.
	Other outputs			
T1.3	Landscape DC: Mapping the knowledge base on selected BDS challenges addressed by the Knowledge Exchange Networks	Shiny app format, using different public available databases (e.g. Cordis) under development.	Will be openly available through the SSBD portal.	
WP2: Landscape of actors, networks and policies				
Task	**Title of data produced by the project**	**Description**	**Publishing / reporting channel**	**Reasons for not publishing the data and limiting access**
T2.1	Interviews and survey with (1) the representatives of EU and international organisations and (2) the representatives of	Data consists of voice and/or video recordings, aggregated data from interviews, text documents and survey data Formats: .mp3, .mp4 (recordings), .docx, .xlsx, (aggregated/result data), .pdf.	D'Amato et al. 2023. D2.1. Assessing the community of key actors for the Biodiversity Strategy for 2030. https://bioagora.eu/storage/app/media/BioAgora%20Deliverable%202.1_20231130.pdf	The data is published in aggregated form to prevent opinions from being traced back to individual persons (although they may still be linked to organisations). Personal and sensitive data collected cannot be published in their raw form.

		selected Knowledge Exchange Networks		An update of the deliverable was published in Summer 2025. https://bioagora.eu/storage/app/media/BioAgora Deliverable 2.1_05.2025.pdf D'Amato et al. 2025. A social network analysis of the European science–policy–society interface on biodiversity https://doi.org/10.1111/cobi.70023	
	T2.1	Database of organized actors at the biodiversity science-policy-society interface	The data covers a broad range of organisations operating at the interface of science, policy, and society, with an impact on biodiversity and sustainable development. While the focus is on actors operating at the EU level, the dataset also includes particularly important international, regional, and national organisations. Formats: .xlsx	D'Amato, D., Rantala, S., & Korhonen-Kurki, K. (2025). Organized actors at the biodiversity science-policy-society interface [Data set]. Zenodo https://doi.org/10.5281/zenodo.10077773	
	T2.2	Interviews with representatives of EU and international organisations and the representatives of selected networks.	Data consists of voice and/or video recordings, aggregated data from interviews, text documents and survey data Formats: .mp3, .mp4 (recordings), .docx, .txt (transcripts), .pdf	Falco et al. 2024. D2.2. A Science brief on policy instruments to support the EU Biodiversity Strategy for 2030. https://bioagora.eu/storage/app/media/D2.2_BioAgora.pdf	The data is published in aggregated form to prevent opinions from being traced back to individual persons (although they may still be linked to organisations). Personal and sensitive data collected cannot be published in their raw form.
	T2.3	Workshops with DCs and networks	Formats: .mp4, .docx, .pdf, Miro (https://miro.com/app/board/uXjVP-zBnXw=/?share_link_id=46119014535)	Westerink et al. 2023. D2.3. A Framework for the Assessment of Transformative Potential of Networks https://bioagora.eu/storage/app/media/BioAgora_D2.3%20Transformative%20Potential%20of%20Networks.pdf	The data is published in aggregated form to prevent opinions from being traced back to individual persons (although they may still be linked to organisations). Personal and sensitive data collected cannot be published in their raw form.

T2.4	Policy documents analysis + expert workshop	Formats .docx., .xlsx., .mp3, .mp4 (recordings)	Adamescu et al. 2024. D2.4. Report on biodiversity policy mismatches and overlaps with recommendations on how to foster cross sectoral dialogue for transformative Science Service https://bioagora.eu/storage/app/media/BioAgora-Deliverable%202.4.pdf	The data is published in aggregated form to prevent opinions from being traced back to individual persons (although they may still be linked to organisations). Personal and sensitive data collected cannot be published in their raw form.
WP3: WP3 State of knowledge				
Task	**Title of data produced by the project**	**Description**	**Publishing / reporting channel**	**Reasons for not publishing the data and limiting access**
T3.1	Knowledge database collecting "actionable information" that can be mobilised to answer requests	Database prototype (spreadsheet) with the collection of actionable information from two documents related to freshwater. The spreadsheet is not usable per se, but can be used to build a database that would be hosted by the Science Service web-platform. Formats: .xlsx	Villero, D., & Costa, A. (2024). BioAgora D3.1 Annex - Freshwater DC actionable information elements [Data set]. Zenodo. https://doi.org/10.5281/zenodo.14018608	
T3.2	Harvesting knowledge of EU funded research projects	A spreadsheet has been created that has harvested publicly available data on EU funded research projects from CORDIS and project final reports including potential outcomes Formats: .xlsx, .txt	TBC. The spreadsheet can be made publicly open.	The spreadsheet may require redaction of coding and other elements used in the analysis before it can be shared publicly.
T3.3	Desktop study and interviews on the actors and methods of research prioritisation landscape	The data includes interviews with experts involved in mapping biodiversity information needs and prioritizing research, literature aimed at identifying challenges and	Deliverable(s)	We adhere to strict data protection practices in accordance with the GDPR for all personal data. Consequently, access to the data is restricted to justified use only. Data containing personal

		opportunities, and information on relevant initiatives, networks and organisations. Formats: .docx, .xlsx, .pdf		information or data from third parties are, in principle, not made publicly available.
T3.4	Potential interviews or survey with knowledge synthesis mechanisms to understand how they work and complete the desk-study (secondary data)	Format most likely to be used: voice or video recordings (mp3 or 4), text documents (.txt, .docs)	TBC	
T3.4	Potential workshop/focus group discussion with knowledge synthesis mechanisms to discuss how to better orchestrate them and develop a responsive future network (with potential recommendations on "when using who")	Format most likely to be used: voice recordings (mp3) or video recordings (mp4), text documents (.txt, .docs, .xlsx, .ppt)	TBC	
T3.4	Design of Tailored EU research and knowledge synthesis programmes based on lessons learned from T3.1, T3.2, T3.3 and T2.1 and T6.4)	Format most likely to be used: regular exchange with partners, text documents (.txt, .docs, .xlsx, .ppt, .pdf)	TBC	

WP4: Inclusive and functional Science Service				
Task	**Data produced by the project**	**Description**	**Publishing / reporting channel**	**Reasons for not publishing the data and limiting access**

	T4.1	Online interviews with science-policy actors at EU and international level	Data consists of audio and video recordings and transcripts of interviews with science-policy interface actors, or in case consent is not granted, written notes. Formats: .mp3, .mp4 (recordings), .docx, .txt (transcripts), .docx, .txt, .pdf (results of the analysis)	Data will be reported in Deliverable D4.1.	The .mp3 and .mp4 files are deleted after transcription, in accordance with the GDPR-compliant consent form. The full transcripts contain confidential information. Deleting the sensitive parts would render the data unintelligible, and therefore the transcripts will not be shared openly.
	T4.1	Design thinking workshops with key stakeholders (EC officers and scientists)	Data consists of written notes taken during the workshops Formats: .docx, .txt	Data will be reported in Deliverable D4.1.	We adhere to strict data protection practices in accordance with the GDPR for all personal data. Consequently, access to the data is restricted to justified use only, and data containing personal information are, in principle, not made publicly available.
	T4.2	Interviews and focus group discussions	Data will be collected to better understand interactions between future key actors and governance structure of the Science Service, in order to identify and analyse potential ethical risks that may affect its reputation. Formats: .mp3, .mp4 (recordings), .docx, .txt (transcripts), .docx, .txt, .pdf, . xlsx	TBC	
	T4.2	Lessons learned from different tasks, Knowledge Exchange Networks and Tailored Programmes	Data will be used to identify key actors and key functions. Format most likely to be used: .txt, .docs, .xlsx, .ppt	TBC	
	T4.2	Desk study on ethical infrastructure	The existing SPI ethical infrastructure will be analysed Formats: .txt, .docs, .pdf		

T4.3	Collected background material for the SSBD business plan	Documentation collected through web search, literature review, databases and personal contacts, for analysis and input into the primary data Formats: .docx, .txt, .xlsx	To be integrated into the Business Plan.	May include sensitive information.
	Other outputs			
T4.2	Ethical risks and mitigation measures identified in the KNEU and H2020 Eklipse project	A long list of potential ethical risks and corresponding ethical measures to tackle those risks coming from the KNEU EU project and H2020 Eklipse project.	TBC	
WP5: Empowering collective agency across scales and sectors				
Task	**Title of data produced by the project**	**Description**	**Publishing / reporting channel**	**Reasons for not publishing the data and limiting access**
T5.1	Interviews with science-policy interface actors	Actors including scientists, policy makers and business representatives have been contacted as key informants. Data consists of voice and/or video recordings, text documents prepared as anonymized summaries. Formats: .mp3, .mp4 (recordings); .docx, .txt (transcripts); .docx, .txt, .pdf, . xlsx (aggregated/result data)	Aggregated and anonymized data is reported in Deliverable D5.1: https://zenodo.org/records/10600408	The .mp3 and .mp4 files are deleted after transcription, in accordance with the GDPR-compliant consent form. The full transcripts contain confidential information. Deleting the sensitive parts would render the data unintelligible, and therefore the transcripts will not be shared openly.
T5.1	Survey on capacity building	Questionnaires addressed to the science-policy interface actors (incl. scientists, policy makers, business representatives) contacted as key informants.		The data may contain personal information that cannot be fully anonymised or removed.

		Formats: .docx, .xlsx, .txt		
T5.3	Deliberative group discussions with citizens and capacity building events	Data consists of photos and video recordings. Formats: .mp4, .jpeg	TBC	
T5.4	Knowledge database of citizen science projects and projects with citizen science component was created	A spreadsheet with projects concerning the four Knowledge Exchange Networks Freshwater, Marine, Nature based Solutions and Pollination. The spreadsheet can be used to complement the database hosted by Science Service web-platform	TBC – The task is ongoing.	
T5.4	Interviews with project co-ordinators. Also, online survey will be administered	Data consists of voice and/or video recordings, text documents prepared as anonymized summaries. Formats: .mp3, .mp4 (recordings), .docx, .txt (transcripts), .docx, .txt, .pdf, .xlsx	TBC – The data collection is underway.	If anonymisation of personal data is possible, the data will be published.
	Other outputs			
T5.2	In-person SPI dialogue workshops, dialogues series and trainings	Workshops, training sessions, and dialogues are concepts to be developed and tested throughout the project to support capacity building, contributing to the development of the SSBD functions, specifically ‘building and activating the clusters’ and ‘mainstreaming.’		

4. FAIR data

The Fair principles will be followed as closely as possible to ensure that our data is Findable, Accessible, Interoperable and Reusable. Below is a summary of the FAIR practices we will implement:

- **Findability:** We enhance findability and re-usability of data by publishing it in open repositories whenever possible and by providing sufficient metadata.
- **Accessibility:** Whenever data can be made openly available, it is licensed under a CC-BY licence, which also promotes reusability.
- **Interoperability:** We use open and widely accepted formats wherever possible to support interoperability.
- **Reusability:** We will publish data in repositories that support FAIR principles, such as those that assign persistent identifiers and allow for inclusion of sufficient documentation. In addition, we will follow practices that ensure data quality.

The following sections describe in more detail what implementing the FAIR principles means in practice.

4.1. Making data findable through metadata

The data to be published will be accompanied by adequate metadata. Metadata is produced mainly for data that will be opened publicly, but also for data that is not opened if necessary.

Metadata and documentation are essential for several reasons:

1. They make data understandable.
2. They help track and recall what data has been produced.
3. During the project, metadata and documentation facilitate data exchange, enabling project partners to use data effectively.
4. Creating adequate metadata and documentation from the early stages ensures smooth archiving and publishing, as well as the findability and understandability of published data.
5. Documenting and openly sharing data collection and processing methods, methodological choices, and workflows help data users evaluate the data's usefulness and assess the validity of both the data and the conclusions drawn from it.

Metadata is particularly important when data is accessed through a repository or database. However, systematically producing metadata throughout the project also improves data usability during research activities.

The repositories we plan to use for publishing the data or indexing metadata of data require at least a minimum set of metadata. We are therefore prepared to provide the following metadata elements. Elements 1 to 3 should be recorded as soon as the data is generated or collected:

1. **Basic information:**
   - Authors/ creators.
   - Name of the data.
   - Author affiliation(s)
   - Year when data was created / period to which the data relates.
2. **A brief description:**
   - A summary of the data to make it understandable and help users evaluate whether it meets their needs.
   - Includes details such as the purpose of the data, origin, production, and processing.
3. **Data ownership:**
   - Information about the owner or responsible party for the data.

When decided to publish the data, the following elements 4 to 6 are added:

4. **License:**
   - Information defines terms of re-use (e.g. the funder (EU) recommends the CC-BY license, the most permissive CC license. Other licenses can be used as needed. If it would be about something other than data (or publications), there are other possibilities for licensing. E.g. MIT license for the code.
5. **Access details:**
   - Specifies the level of openness, open access, embargoed, restricted, or request.
6. **Keywords:**
   - Ensures that data is findable and understandable.

At the publication, a Persistent Identifier (PID) is assigned:

7. **A PID, e.g., DOI and HANDLE**
   - A unique persistent identifier (PID) is assigned for the data that facilitates citation, ensures versioning, and provides a permanent link to locate the data in the repository/archive.

In addition, specific metadata may be produced from some of the data. Such metadata could relate, for example, to interviews, group discussions and event participants, and could be about the participants' country, the organisation they work for, their field (policy, science, practice, business) and location. As the metadata concerns personal information, it is collected, stored and processed with particular care and not made publicly available in a way that could reveal individuals.

In the repositories/archives and databases we aim to publish our data uses standard metadata schemas and data transfer protocols. They provide a platform where the metadata is both human-understandable and machine-readable, allow harvesting and indexing of metadata.

Data to be collected and published on the Science Service for Biodiversity web platform will also be accompanied by standard metadata. For more details on the metadata on the SSBD web platform, and data management practices, see Appendix 2. The Science Service web platform.

4.2. Making data accessible

In BioAgora the accessibility of public data will be ensured by making them available in open repositories/archives and data bases. Similar accessibility practices will be used to other outputs when applicable. Project outputs and results are also presented through various communication channels, for example the project website and social media tools. More detailed information can be found in the Final Plan for Dissemination, Exploitation and Communication of Results of BioAgora (D4.7). Project's public data will be opened mainly in the European data repository Zenodo. Bioagora has establishes a Zenodo Community. All possible outputs are planned at being published through Zenodo, in order to include them in the Zenodo community, ensuring their findability and accessibility. Other repositories are used where appropriate, for example to improve the accessibility and usability of data by opening them in a domain specific repository. Zenodo is recommended because it is easy to use, complies with the requirements of the Horizon Europe programme and the grant agreement, and ensures open access to public data in accordance with the FAIR principles.

Access to outputs containing personal data or other sensitive information is restricted or, in some cases, not shared openly at all. Intellectual property rights related to data and other outputs may also require limited access or prevent sharing altogether. So far, we have identified that data containing personal information, such as interviews, workshops, and surveys, requires careful consideration in BioAgora to determine whether it can be made openly available.

Collection, processing, accessing and publishing practices

The main data types collected in the project is presented below. In general, the data collected and produced by the project forms an essential basis for reports and other publications and is made available or reported in manner appropriate to the data type and the purposes of each publication (e.g. aggregate reporting for data containing personal information). The main data types are outlined below and described from general viewpoint on sharing and open access. The following questions will be addressed:

- How the data will be shared with project partners.
- What data will be made openly available and if not, the reasons why.
- What steps will be taken to enable sharing and public availability of data that originally contained personal data.
- How the data will be made publicly available.

Interview, group discussion, workshop, expert consultation, and survey data

Data derived from interviews, group discussions, focus groups, workshops, expert consultations and surveys is considered personal data. The rights of research participants will be respected in accordance with good research practices, ethical guidelines and applicable European Union and national legislation. The collection, processing, sharing and publication of data will comply with the General Data Protection Regulation (GDPR) and national laws.

In line with the GDPR, all contributors to data collection activities will receive a written privacy notice that:

1. Provides adequate information about the project and participants' rights
2. Specifies what personal data is collected and for what purpose in the project
3. Defines the legal basis for processing personal data
4. Provides clear contact information in case participants have questions about the project, how their personal data is used, or wish to raise a concern.

Participants are clearly informed that participation is voluntary and that they can refuse or withdraw from the study. At the same time as providing the privacy notice, before any personal data is collected, participants' consent is requested.

Special attention will be given to how the data is collected and processed to enable publication. We will carefully assess which data (and which version of it) can be made publicly available, following best practices for data collection and processing. To enable publishing, we will:

1. Minimise the collection of personal data, ensuring that only necessary identifying information is collected,
2. Anonymise the data, and if anonymisation is not feasible,
3. Provide aggregated version of data (e.g., summaries, spreadsheets).

However, the decision to publish the data will be based on scientific consideration. Data will only be published in a repository if it contains sufficient information to allow valid interpretations. If data cannot be made publicly available, in line with Horizon Europe (HE) guidelines, access for verification purposes will be granted.

The following practices will be followed for data storage and publishing:

- Personal data will be retained as long as necessary for research purposes and will not be stored for more than 10 years after collection.
- Non-anonymised data (original recordings and transcriptions) will be stored securely on data collectors secure server. Recordings will be deleted once they are no longer needed for the study.
- The data containing personal data is shared within the project for justified research purposes.

- Anonymised, aggregated data and data summaries can be shared with partners via Teams provided by Syke.
- Publicly available data will be published in Zenodo. If domain specific repository is more suitable (e.g., for interview and survey data), we will consider using an archive provided by the Consortium of European Social Science Data Archives (CESSDA).

Data facilitating the development of the BioAgora network

The network and contact information collected on various biodiversity actors, such as experts, key representatives of organisations, organisations themselves, and relevant projects, is produced primarily for internal use and will not be made publicly available in a repository. Some of the BioAgora network data may be shared through project deliverables or communication materials, where appropriate. As the data includes personal information, it must not be disclosed without the consent of the individual concerned. Data is mainly shared with the partners. The Teams platform provided by Syke will be used when appropriate.

Data on Knowledge Exchange Networks (KENs)

The data on KENs (Freshwater, Nature based solutions, Pollination, Marine is very diverse. Spatial data will be published in an open data repository, such as Zenodo, when it is considered useful for scientific research and the broader public. The secondary data, e.g., on sociodemographic and decision-making processes will be published through deliverables. If necessary, the data is shared with the partners in Teams.

Science service request processing sharing and storage

At this stage of the Science service development phase, the cloud of the Helmholtz Centre for Environmental Research (UFZ) will be used for storage, with a dedicated folder created for each request. Focal points and experts will have access to their specific request folder via a provided link and password, while the request management team will have access to all folders. Additionally, each focal point will manage working documents, such as report drafts, in Google Drive, ensuring that all relevant links are shared with the request management team.

Publications and other written materials

Each work package in the project will compile, synthesise and analyse and draw conclusions using various publication types and other written materials. The data includes, for example:

- Research on policies and policy documents
- Scientific articles
- Information on EU projects and impact assessment tools
- Course materials

The data is primary secondary and publicly available. It is shared with the partners via Microsoft Teams. The results are mainly published as project deliverables, while respecting the intellectual property rights of authors, creators, and publishers.

According to the Description of Action (DoA), the project's deliverables, reports are public. We will publish the deliverables in the Zenodo, where they will be assigned persistent identifiers (e.g., DOIs), and included in the BioAgora's Zenodo community. Persistent identifiers, and access via the Zenodo repository ensure that the deliverables are easy to find and cite. Examples of such deliverables include:

- A Framework for the Assessment of Transformative Potential of Networks (D2.3) https://doi.org/10.5281/zenodo.10406918
- Mapping the needs of decision-makers to tailor capacity development activities (D5.1) https://doi.org/10.5281/zenodo.10600408
- Typology of challenges that hinder the implementation of BDS 2030 (D1.1) https://doi.org/10.5281/zenodo.7685651

Although BioAgora's Description of Action (DoA) does not include scientific peer-reviewed publications, any that are produced during the project will be published under CC BY licenses in immediate open access channels, together with the related research data when possible. To ensure permanent access, peer-reviewed versions will be deposited without delay in a trusted repository or archive.

For data publication, we have identified repositories that may be suitable for our data. When data processing has progressed to the stage where publication is being considered, the details are documented in Table 1. As BioAgora is a service development project, it generates a substantial amount of data that serves as source material and background for deliverables and other outputs. Table 1 also indicates which publications have used specific datasets. This is particularly useful when the data cannot be made openly available in a repository. In such cases—mainly due to the inclusion of personal information. Table 1 provides the reasons why the data cannot be published.

4.3. Making data interoperable

Data and other outputs are stored and shared in different formats depending on their type and the software used to produce them. However, their (re)usability is ensured by using formats that are widely used, preferably non-proprietary formats. Already during the project, data will be stored in a format that is easily accessible to other project participants. Where appropriate, for example when the original data or analysis results are produced using specific software, outputs are produced in common formats.

In BioAgora, most of our data are text documents, thus most of the data are presented in .docx, .txt, .pdf and .xlsx format. For example:

- Interview, group discussion and expert consultation data consists of several data types and thus wide range of format:
    - Video and audio data will be stored in non-proprietary file formats such as .mp3, .mp4
    - Anonymized/aggregated versions of data and results are presented as text documents and spreadsheets in non-proprietary file formats (.txt and .pdf), in common formats (.docx, . xlsx) and in software specific formats (MX20, .hpr; QDPX).
- Survey data is presented in .docx, .txt, .xlsx formats.
- The spatial data is presented in .xlsx, arcgis, .geotiff, .shp, open street map.
- The data collected for the creation of networks on organisational personal data is in .pdf, .docx, .xlsx. format.

Public data will be shared in repositories/archives mainly in commonly used formats to ensure the interoperability and re-usability of our data.

4.4. Increase re-use of data

In general, the data is collected, processed, and synthetised according to established practices to ensure high data quality. When necessary, guidelines are developed to maintain data consistency. For example, a guide has been prepared for conducting interviews and document analysis. To ensure a consistent approach to GDPR compliance, a common privacy notice and consent templates are used. Additional, specific software tools e.g., MAXQDA, Atlas.ti) are employed when appropriate to support data processing procedures. The guidelines and other supporting documents are available to project partners in Teams, along with the data relevant data.

Our key principles for ensuring data reuse:

1. Open access: Data and other outputs that can be made openly available are freely accessible to all in repositories and databases.
2. Rich metadata: Data is accompanied by comprehensive metadata to enhance findability and usability.
3. Adequate Documentation: Data is documented as necessary to ensure proper interpretation and re-use.
4. Licensing: Data and other outputs (where applicable) are published under the Creative Commons CC-BY 4.0 license, while metadata is released under CC0.

Proper data documentation supports both the re-use of data during the project and its long-term accessibility when made publicly available for potential re-users in a repository/archive database. The level and type of documentation in BioAgora depend on the type data, its intended use and domain specific practices. Documentation can be provided in different ways, such as:

- Including relevant details in publications / deliverables
- Attaching additional documentation (e.g., a README file) to the dataset

Additional documentation is needed if the publication descriptions are insufficient, there are no related publications or, in case of publicly available data the metadata and data itself are not enough to ensure the understandability and validation.

The documentation should be accurately and detailed, covering aspects such as:

- Methodology: How the data was collected and analysed
- Collection and Processing Methods: Information on the techniques used
- Variable Descriptions: Key details about dataset variables
- Project Information: General details about the project and its objectives

For interview data, documentation may include:

- Summaries and key topics of analysis
- Interview questions
- Interview metadata: Date, time, duration, country, and work sector of the interviewee

Public data, peer-reviewed publications and other agreed-upon outputs will be freely available to all in the repositories or archives. The reuse of these outputs will be ensured by publishing them under the Creative Common license CC-BY 4.0., while preserving intellectual property rights for authors and creators.

5. Responsibilities and allocation of resources

Syke as the project coordinator has a general responsibility for data management and compliance with common guidelines and EC requirements. A data management manager appointed by Syke is responsible for preparing the data management plan, drawing on the expertise of the partners. The data manager also supports the implementation of data management as agreed in the project. On the WP level the overall responsibility lies with the WP leader. The WPs and Tasks will ensure that the information on the data generated and used by them and other outputs produced by the project are updated in the DMP. In practice, all participants in the project are responsible for ensuring that the commonly agreed management guidelines are followed.

Our project has not identified any specific costs associated with data management between partners. The external data sharing through websites and newsletter are budgeted under WP7 activities. Pensoft together with Euronovia develops the content for the project website. The platform for website maintenance is provided by external service provider. We will refine the information on the use of other shared platforms and tools as necessary. The day-to-day data management practices are mainly integrated into research and included in the project budget as salary costs. We also don't purchase equipment or services, also the facilities are free of charge or provided by the partner

organisations. For example, data repositories and archives, which are used to make data and other outputs openly available, are free of charge. In addition, Syke provides Teams/Sharepoint, which is used to share and organise data and documents and as a collaboration platform, without charging the project directly.

6. Data and information security

Syke provides the Teams collaboration platform for the project and is therefore responsible for instructing partners to share only materials for which Teams offers an appropriate level of security. In general, the IT services of the partner organisations are responsible for providing secure storage and backup services for their staff involved in the project. Backups are typically performed automatically on a regular schedule, in line with good and secure practice. Project participants are also encouraged to store data and other materials on password-protected servers within their organisations, ensuring access is restricted to authorised users. Data and other electronic materials will be stored both in the servers of the partner organisations responsible for them and Syke Teams. Syke Teams is used when materials are shared and collaboratively handled within BioAgora. For sensitive data, such as interviews, appropriate safeguards will be in place. If needed, a dedicated secure storage and data transfer solution will be used, for example, in cases where non-anonymised interview data must be shared with project partners.

7. Ethical and legal considerations

BioAgora is committed to good scientific conduct, responsible governance and compliance with relevant legislation. Ethical standards and guidelines will be strictly followed. We adhere to the ALLEA European Code of Conduct for Research Integrity. As stated in Chapter 4, Ethics self-assessment DoA, project involves human participants and includes the production and use of personal data. These data include, for example, interview, survey and contact information data. All project partners are committed to processing the personal data in line with good research practices, ethical standards, and principles of EU GDPR. We apply data management practises that safeguard the identities of research participant. The practices include ensuring voluntary participation and obtaining informed consent. The participants are provided with sufficient information about the study and are granted all legal rights to their personal data, such as the right to access their data and to withdraw from the study at any time. Data collection, handling and transfer procedures are designed to ensure participants' privacy. Data containing personal identifiers will only be shared among project partners using secure transfer systems. We also take into account the implications of involving participants from non-EU countries when handling and sharing personal data. A detailed description of the procedures for handling personal data is provided in Section 4.2.

Appendix 1. Partners of the BioAgora project

Finnish Environment Institute, Syke

University of Bucharest

University of Trento

National Research Institute for Agriculture, Food and Environment, INRAE

Research Institute for Nature and Forest, INBO

Netherlands Environmental Assessment Agency, PBL (Ministry of Infrastructure and Water Management)

Norwegian Institute for Nature Research

Helmholtz Center for Environmental Research

Delbaere Consulting

Forschungsverbund Berlin e.V.

Centro de Investigacíon Ecologíca y Aplicaciones Forestales, CREAF

Environmental Social Science Research Group, ESSRG

Pensoft Publishers

European Regional Center for Ecohydrology, ERCE PAN

Euronovia

Wageningen University and Research

European Citizen Science Association, ECSA

Alfred Wegener Institute

The UN Environment Programme World Conservation Monitoring Centre, UNEP-WCMC*

UK Center for Ecology and Hydrology, UKCEH *

* Associated partners

Appendix 2. The Science Service web platform

The information presented in this appendix is preliminary until the specific requirements of the web-platform and the content/data it will contain are defined.

The Science Service web-platform will gather and store data for the following purposes:

1. To support the Science Service user community and enquiry/ticketing process
2. To collate and share tools, resources and case studies with the EU Knowledge Centre for Biodiversity and Oppla

Data type	Source	Formats	
Science Service users	▪ Website sign-ups ▪ Submissions to online enquiry/ ticketing process	▪ Online forms	▪ Stored securely in web platform ▪ Public profiles available to other members if consented ▪ Users can access all stored information and delete if desired
Tools and resources	▪ Uploads to platform ▪ Evidence assembled by project team	▪ Reports ▪ Spreadsheets ▪ Links to online tools	▪ Shared through the web platform

The Science Service web-platform will follow the principles of **FAIR data management** where possible, ensuring data is findable, accessible, interoperable and re-usable. The underlying principle is that data should be "as open as possible, as closed as necessary". All personal information will be managed in accordance with GDPR.

Keywords will be used to categorise and assist the retrieval of data stored on the web-platform. Each piece of data will be assigned a persistent and unique link that will ensure that it is available in perpetuity after the end of the project.

All data published by the web-platform will be accompanied by **metadata** following the Dublin Core Metadata Initiative. We will adopt ISO 15836-1:2017 that establishes 15 core metadata elements for cross-domain resource description. These terms are part of a larger set of metadata vocabularies maintained by the Dublin Core Metadata Initiative.

Dublin Core Metadata Element Set:

Title	**Contributor**	**Source**
Creator	**Date**	**Language**
Subject	**Type**	**Relation**
Description	**Format**	**Coverage**
Publisher	**Identifier**	**Rights**

The full details of the Dublin Core Metadata Initiative terms can be found at

http://www.dublincore.org/specifications/dublin-core/dcmi-terms

The following **data formats** will be used for data hosted by the Science Service web-platform. If data is shared by third parties in other formats, then it will be converted provided this does not degrade the quality of the data.

Data type	Formats
Text	▪ HTML ▪ Open Document Format Text (.odt)
Spreadsheet	▪ Open Document Format Spreadsheet (.ods)
Presentations	▪ Open Document Format Presentation (.odp)
Drawings / diagrams	▪ Open Document Format Drawing (.odg)
Raster images	▪ Tagged Image File Format (.tiff) ▪ Joint Photographic Experts Group (.jpeg) ▪ Portable Network Graphics (.png)
Vector images	▪ Scalable Vector Graphics (.svg)
Video	▪ MPEG-4 Part 14 (.mp4)
Audio	▪ MP3 (.mp3)

Wherever possible, content hosted by the web-platform will be shared as **open data.** This will include data in a wide range of formats.

It is recommended that the web-platform adopts a Creative Commons Attribution-ShareAlike license for managing open data. This type of license has the advantage over a fully open, or public domain license in that it ensures that the originators of the content are given due credit – for instance, including where EU funding has supported content development. Appropriate attribution of content will also encourage partners from the wider network beyond the project team to share content with the Science Service. The requirement for any derived outputs to also be shared under a similar licence may be waived for some outputs if that would reduce their usability, or convene any proprietary licenses or protected IP.

Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)

Summary: https://creativecommons.org/licenses/by-sa/4.0/

License: https://creativecommons.org/licenses/by-sa/4.0/legalcode