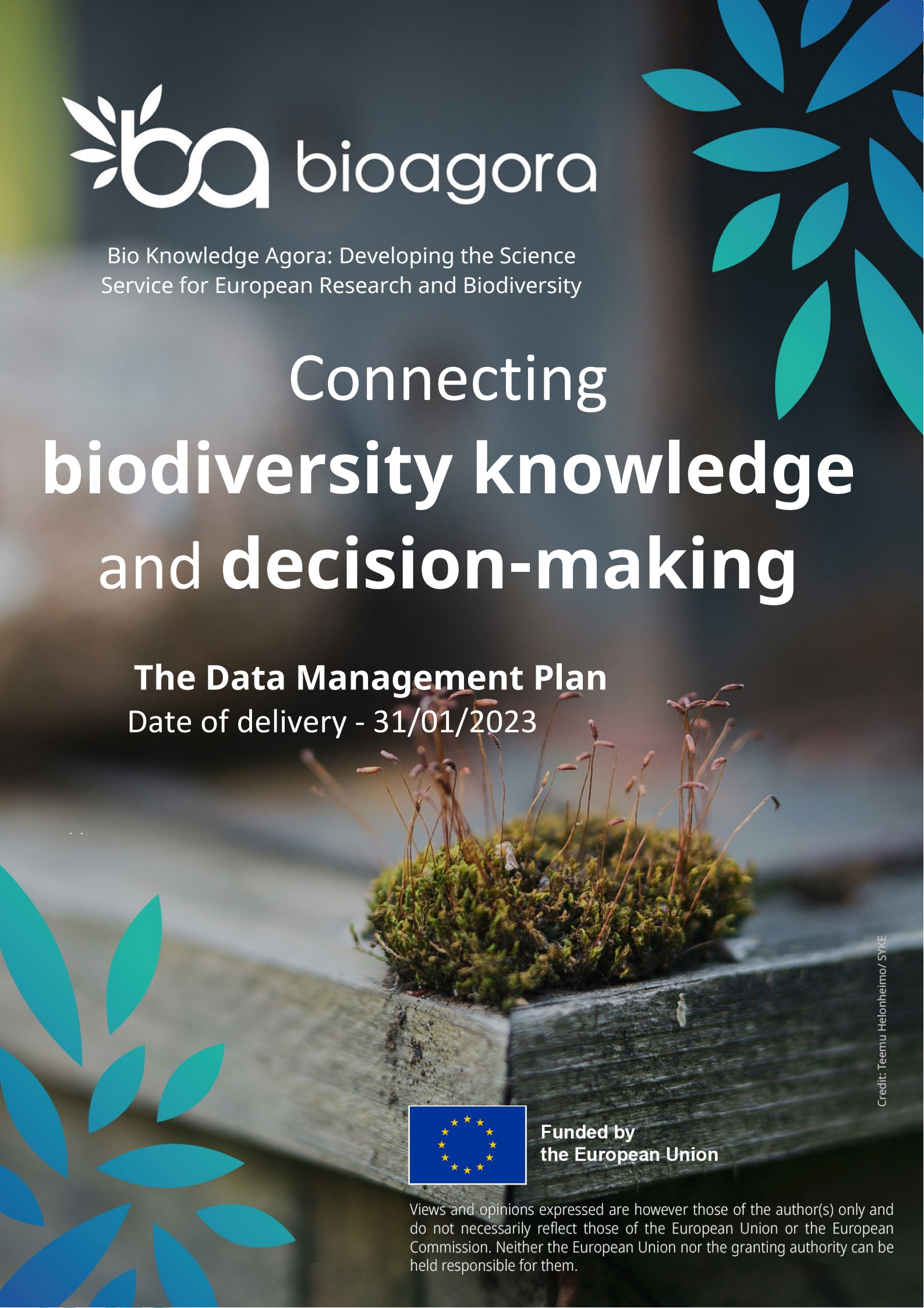
bioagora

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

Connecting **biodiversity knowledge** and **decision-making**

The Data Management Plan

Date of delivery - 31/01/2023



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	D6.2
Work Package number	WP6
Deliverable title	The First Data Management Plan
Lead beneficiary	Finnish Environment Institute (Syke)
Author(s)	Maria Söderholm and Kati Vierikko (Finnish Environment Institute)
Due date	31/12/2022
Actual submission date	31/01/2023
Type of deliverable	DMP
Dissemination level	PU (Public)

VERSION MANAGEMENT

Revision table			
Version	Revision	Date	Description
1	Maria Söderholm/Finnish Environment Institute	26/01/2023	First draft
2	Kati Vierikko//Finnish Environment Institute	30/01/202	Internal review
3	Maria Söderholm/Finnish Environment Institute	31/01/2023	Update following review

LIST OF ACRONYMS AND ABBREVIATIONS

Acronym / Abbreviation	Meaning / Full text
BDS 2030	EU Biodiversity Strategy 2030
DCs	Demonstration Case
DMP	Data Management Plan
DoA	Description of the Action
DOI	Digital Object Identifier
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

The BioAgora project funded by the Horizon Europe programme aims to bring research on biodiversity better closer to policy making in Europe, and globally. BioAgora is a Research and Innovations Action, RIA project that brings together 22 top level organisations in a partnership to establish a European Science Service for Biodiversity (SSBD) to promote and protect biodiversity (Appendix 1). Together with a broad network of stakeholders, BioAgora is co-creating new ways to bridge the gap between science, practice and policy, and using this as a basis for developing the Service for future needs. To achieve this, the project will 1) analyse the existing landscape of science-policy interfaces, assess their policy tools and the current biodiversity knowledge in Europe, 2) engage with a broad range of actors, 3) develop a governance structure and system model the Service by testing it in real life, 4) provide capacity building for the EU community of decision makers to become empowered for transformative change for biodiversity. The project's mission is to develop a Science Service that brings together the biodiversity research community and the knowledge it produces to support decision-makers and other users with the information it provides. BioAgora has five main objectives to address these biodiversity knowledge needs:

1. The development of Science Service for Biodiversity (SSBD) as a scientific pillar for Knowledge Centre for Biodiversity (KCBD).
2. Uses multiscale demonstration cases (DCs) across EU Member States that ratchet up the EU Biodiversity Strategy 2030 (BDS2030) and to showcase real-life actions and processes in various biodiversity management or governance contexts.
3. Encouraging of socially justice practices means considering the pluralistic knowledge base and the relative values of biodiversity and integrating societal values alongside with scientific knowledge in the development of the SSBD.
4. Establish an inclusive governance model and feasible business model for SSBD that makes the service responsive, transparent, and accountable and supports its long-term operation.
5. Promote capacity building and co-learning processes for an inclusive and functional SSBD by ensuring openness and participation of stakeholders from all relevant societal branches.

2. Objectives of the Data management Plan

This document (D6.2) presents BioAgora's updated plan for managing data and other outputs. The previous DMP provided an overview of the project data and the general data management principles to be followed.

The First Data Management Plan aims to set out in as much detail as possible:

- The data collected and used by each work package.
- What data will be shared with project partners.
- What data will be published.
- Where and how data will be made publicly available.
- What data is collected and stored on the "Science Service" web platform.

BioAgora is committed to good scientific conduct, governance and to valid legislation. Ethical standards and guidelines will be strictly applied. We will follow the guidelines issued by the ALLEA, the European Code of Conduct for Research Integrity. Personal data are processed in accordance with the European General Data Protection Regulation (GDPR). In addition, the management of data and other outputs will follow the guidelines of the Horizon Europe Programme and the FAIR data principles.
The Data Management Plan is a living document. As the project progresses, the plan will get detailed form and will be kept up to date to support project data/output management activities during the project. The next updated version of this first plan, Deliverable 6.7, will be presented in June 2025.

3. Summary of data

Achieving BioAgora's objectives requires a broad and diverse knowledge base. From the outset, the project has identified data and information needs: what new and existing data and information are needed to achieve the objectives, and how to locate and use them. BioAgora generates new data and uses the existing data by synthesising and analysing it to provide all project participants with a solid knowledge base for the creation of a science service. There is a wealth of data and information produced by different actors and projects, which will be fully exploited in the project. The key approach of the project is to involve a wide range of stakeholders in the biodiversity community in both the generation of data and knowledge and their application in the development of the “Science Service for Biodiversity”.

The BioAgora project will also develop specific functions, tools and mechanisms to promote dialogue between the research community and decision-makers and the dissemination and use of biodiversity knowledge. Some of tools are developed together with the KCBD. These functions and tools include for example:

- Supporting tool integrated to the SSBD to take up the biodiversity related requests from the KCBD system that are made by EU level policy-makers decision-makers on biodiversity.
- Mechanism to harvest and rationalize knowledge generated by EU funded projects, initiatives, and platforms.
- Interactive and widely accessible web-platform and related functions for the SSBD.

This chapter summarises the data in as much detail as possible at this stage of the project (Table 1). We will certainly identify more data and other information that we need to develop the SSBD. This also applies to the data and other content that will be collected and stored on the “Science Service” web platform, which is also presented below (Table 2).

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

Work package	Types of data & Description
WP1 **Underpinning the Biodiversity Strategy – Case Study Hub**	**Primary data:**
	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)
	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
	Expert consultations with BioAgora project partners: face-to-face T1.1 workshop organized at the kick-off meeting, and online consultation using the shared Teams folder. Formats: text documents (.txt, .docx)
	Data from the ticket system of key BDS policy request ▪ Formats: text documents (.txt, .docx)
	Information on possible future research pathways through expert consultation, focus group, survey data, scenario analysis and other methods. ▪ Formats: .mp3, .mp4 (recordings), .docx, .txt (transcripts), .docx, .txt, .pdf, . xlsx

	Spatial data developed for Demonstration Cases. (Refers to processes and activities related to biodiversity management in areas such as the pollination, freshwater, nature-based solutions and urban green infrastructure.)
	▪ Nature based solutions – explicit spatial information on ecosystem conditions, ES supply, and demand, including mapping of beneficiaries.
	  ▪ Formats: .geotiff. .shp
	▪ 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.
	  ▪ Formats: .xlsx, .docx, arcgis, open street map
	▪ Pollination – to be defined later
	▪ Convention on Biological Diversity (CBD) – to be defined later

	Secondary data:
	Published scientific literature and grey literature such as project reports, publicly available information on websites. Formats: .docs, html, .pdf, .txt, . xlsx
	Collecting and coding knowledge on the state of art of BDS implementation challenges from key EU projects. ▪ Formats: .txt, .docx (a content analysis of projects' documents)
	Mapping the knowledge base available in the scientific literature on the selected BDS challenge addressed by each demonstration case (e.g., monitoring data, assessment methods, and implementation tools) using appropriate knowledge synthesis mechanisms (e.g., literature reviews, rapid evidence assessment, etc). ▪ Formats: .txt, .docx, .xlsx

	Spatial data for Demonstration Cases
	▪ Nature based solutions – sociodemographic (e.g., population density) and biophysical data (e.g., land use and land cover, soil properties, etc.).
	  ○ Formats: .geotiff, .shp
	▪ Freshwater – decision making processes
	  ○ on the installation of hydropower plant construction
	  ○ construction plans for river expansions
	  ○ plans for restorations on river connectivity
	  ○ Formats: .xlsx, .docx, arcgis, open street map

	▪ Pollination – to be defined later ▪ Convention on Biological Diversity (CBD)– to be defined later
	Data describing the network of actors that can be activated to answer policy requests. ▪ Formats: .xlsx
WP2 Landscape of actors, networks and policies	**Primary data:**
	Interviews with (1) the representatives of EU and international organisations and (2) the representatives of selected networks and demonstration cases. Data consists of voice and/or video recordings, text documents. ▪ Formats: .mp3, .mp4 (recordings), .docx, .txt, .MX20 (transcripts), .xlsx .MX20, .hpr; QDPX (aggregated/result data)
	Survey addressed to the representatives of selected networks and demonstration cases. Data consists of questionnaires ja excel database. ▪ Formats: .docx, .xlsx, .txt
	Secondary data:
	Databases listing organisations and personal contacts of organisations' key representatives/leaders. ▪ Formats: .docx, .xlsx.
	Policy documents and description of policy tools of employed by various networks. Analysis and results of policy implementation and policy tools performance (e.g., databases). ▪ Formats: .pdf, .docx, .xlsx.
WP3 State of knowledge	**Primary data:**
	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)
	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)
	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)
	Secondary data:
	Desk study on mapping knowledge synthesis mechanisms Eklipse evidence reports (Format: text documents: pdf) Evidence reports from Identified knowledge synthesis mechanisms ▪ Formats: .txt, .docx, .xlsx, .pdf
	Primary data:
WP4	Interviews and focus group discussions with science-policy interface actors contacted as key informants. Contact information of experts are collected from publicly available sources (i.e. webpages)

Inclusive and functional Science Service	▪ Formats: .mp3, .mp4 (recordings), .docx, .txt (transcripts), .docx, .txt, .pdf, . xlsx
	Lessons learned from different tasks, Demonstration Cases and Tailored Programmes to identify key actors and key functions ▪ Format most likely to be used: .txt, .docs, .xlsx, .ppt
	Interviews and focus group discussions to better understand interactions between future key actors and governance structure of the Science Service to identify and analyse any potential ethical risks influencing the reputation of the future science service ▪ Formats: .mp3, .mp4 (recordings), .docx, .txt (transcripts), .docx, .txt, .pdf, . xlsx
	The SSBD business plan and supporting documents produced by the project. Document provided may be considered entirely or partly sensitive and then only public version are accessible to all. ▪ Formats: .docx, .txt, .xlsx
	Secondary data:
	Published scientific literature and grey literature (e.g. research results) with a focus on how other science-policy interfaces are working ▪ Formats: txt, .docs, .pdf
	Documentation collected through web search, literature review, databases and personal contacts, for analysis and input into the primary data ▪ Formats: .txt, .docs, .xlsx, .ppt
	Desk study on analysing existing SPI ethical infrastructure ▪ Formats: .txt, .docs, .xlsx, .ppt
	Long list of potential ethical risks and corresponding ethical measures to tackle those risks coming from the KNEU EU project and H2020 Eklipse project.
	Documentation collected through web search, literature review, databases and personal contacts, for analysis and input into the primary data ▪ Formats: .docx, .txt, .xlsx
WP5 Empowering collective agency across scales and sectors	**Primary data:**
	Interviews with science-policy interface actors (incl. scientists, policy makers, business representatives) 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)
	Questionnaires addressed to the science-policy interface actors (incl. scientists, policy makers, business representatives) contacted as key informants. ▪ Formats: .docx, .xlsx, .txt
	Deliberative group discussions with citizens and capacity building events. Data consists of photos and video recordings. ▪ Formats: .mp4, .jpeg
	Secondary data:
	▪ Published scientific literature and grey literature such as project reports and publicly available information on websites. ▪ Online available course materials. ▪ Online information about ongoing citizen science projects and networks and impact assessment tools and published handbooks ▪ Formats: .docs, .html., .pdf, .txt

Table 2: Description of the data collected and stored on the “Science service web platform”.

Data type	Source	Formats
Science Service users	▪ Website sign-ups ▪ Submissions to online enquiry/ ticketing process	▪ Online forms
Tools and resources	▪ Uploads to platform ▪ Evidence assembled by project team	▪ Reports ▪ Spreadsheets ▪ Links to online tools

4. FAIR data

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. In general, attaching sufficient metadata to the data from the beginning:

- Makes data understandable.
- Helps to keep track of and remember what data has been produced.
- Ensures that data is stored and published smoothly in the repository when we are ready to do so.
- Also supports the use of data by project partners during the project.

The information represented by the metadata about the data is particularly important when the data is accessed through an archive/repository or database. However, the continuous and systematic production of metadata will also improve the use of the data during the project.

The repositories and archives we aim to use for making the outputs publicly available or index the metadata of data or other outputs require at least a minimum of metadata to be produced. We are therefore prepared to produce the following metadata:

1. Basic information about the data, such as the author/ creator and the name of the data and the year when it was created and/or the period to which the data relates.
2. Keywords that ensure firstly the findability and secondly the understandability of the data.
3. A brief description of the data makes the data understandable and helps re-users evaluate whether the data meets their needs. The description includes e.g. what the data is about, what its origin is, how it was produced, processed, etc.
4. Licence information to define the terms of use.
5. Level of openness/publicity, if necessary, i.e., the data is open access, embargoed, access to the data is restricted or access must be requested.
6. A persistent identifier, such as DOI and HANDLE, will be assigned to the data when it is made publicly available in the repository/archive. The persistent identifier identifies the

data (and a specific version of it), facilitates citation of the data and as a link helps to locate the data.

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 identify 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. BioAgora "SSBD" 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 (see Initial Plan for Dissemination, Exploitation and Communication of Results of BioAgora (PDEC, D7.1)). Project's public data will be opened in the European data repository Zenodo. 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.

Only personal data and other outputs that contain personal information are restricted or not shared openly at all. Intellectual property rights for data and other outputs may also require limited open access or prevent sharing them altogether. So far, we have recognised that personal data, such as interviews, and survey require us to carefully consider in BioAgora whether we can open such data.

The main data types collected in the project is presented below from a sharing and opening perspective. Depending on the data, the following questions will be addressed:

- How the data will be shared with project partners.
- What data will be made openly available. If this is not possible why not.

- What actions will be taken that personal data is possible to share and make publicly available.
- How the data will be made publicly available.

Interviews, group discussions, expert consultations, and surveys

Data based on interviews, group discussions, focus groups, expert consultations and surveys are considered as personal data. The rights of research participants will be respected in accordance with good research practice, ethical guidelines and valid European Union and national legislation. The collection, processing, sharing and opening of data will be carried out in strict compliance with the GDPR and national legislation.

In accordance with the GDPR, interviewees, participants in discussions, etc., will receive a written privacy notice which (1) will provide participants adequate information about the research and their rights, (2) defines what personal data is collected and for what purpose it is used in the study, (3) defines the legal base for processing personal data.

Data publishing is made possible by (1) minimising the collection of personal data, i.e., collecting information related to personal identifiers only to the extent necessary for the project, (2) anonymizing the data, and if this proves impossible (3) providing aggregated version of data, i.e., summaries, spreadsheets, etc. However, decision to publish the data will be based on scientific consideration. Aggregated data will only be published in a repository if it is still comprehensible and allow valid interpretations to be made. If data cannot be made publicly available in a repository, in line with HE, access to data for verifying purposes is granted.

The non-anonymized data, original recordings and transcriptions, is stored on the secure server of the data collector. Recordings are deleted when they are no longer needed in the project. Anonymized/ aggregated data is shared with the partners through Teams provided by Syke.

Data to be public are made publicly available through the Zenodo repository or when a domain specific repository designed for interview and survey data is more suitable, we can use, for example, 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, organisations' key representatives/leaders, organisations and projects is data produced mainly for internal use and will not be published, i.e., made publicly available in a repository. BioAgora network data may be published in deliverables and in communication-related publications, but personal data must not be disclosed without the consent of the persons concerned. Data is shared with the partners mainly through Teams provided by Syke.

Data on Demonstration Cases

The data on demonstration cases (Freshwater, Nature based solutions, Pollination, Convention on Biological Diversity (CBD)) is very diverse. So far, data on the Freshwater and Nature based solutions demonstration cases is identified. Spatial data will be published in an open data repository, such as Zenodo. 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.

Publications and other written materials

Each work package in the project compiles, synthesis and analyses and draws of conclusions using different publication types and other written materials as data. This data consists, for example, of research of policies, and policy documents, scientific articles, information on EU projects and impact assessment tools, and course materials. The data is mainly secondary and publicly available. The data is shared with the partners in Teams. The outputs 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 deliverables**, reports, dissemination level are mainly public. We will consider publishing selected deliverables in a publication archive, where they can be given a permanent identifier (e.g., DOI), making them easier to find and refer to. Such deliverables could include, for example:

- Blueprint for a multilayered deliberative process to incorporate societal value choices in the Science Service (D5.4)
- Policy brief and recommendations to the Science Service on integrating citizen science into policy decision-making (D5.3)
- Report on the governing principles of a Science Service (D4.1)
- Systematic overview of forecasts and projections related to the BDS2030 and to specific policies (D1.4)

BioAgora's DoA does not include **scientific peer-reviewed publications**. However, if they are produced by the project, they will be published under CC-BY licenses in channels with immediate open access as well as the related research data when possible. To ensure permanent access to the peer reviewed articles, the latest possible peer-reviewed versions will be deposited immediately upon publication in a trusted repository/archive.

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/generated, processed, and synthetized in accordance with established practices to ensure the data quality. If necessary, guidelines will also be drawn up to ensure consistency of data. For example, a guide has been prepared for conducting interviews and document analysis. The common privacy notice and consent templates ensure the consistent approach to GDPR compliance. In addition, specific software, are used when appropriate to ensure solid data processing procedures (e.g., MAXQDA, Atlas.ti). The guidelines and other supporting documents are available to all project partners in Teams along with the data in question.

We follow four principles to ensure the re-use of data in our project:

1. Data and other outputs that can be made open access are freely available to all in the repositories/archives and databases.
2. The data will be attached with rich enough metadata.

Our practices related to repositories and metadata are described already in the previous chapters. The next principles and related practices are the matter of this section:

3. The data will be documented as necessary.
4. The data and other outputs (where applicable) will be licensed under the Creative Common license CC-BY 4.0 and the metadata under the license CC0.

Documentation of data ensures the (re)use of data, also during the project but especially when the data is made openly available for all potential re-users in a repository/archive database. How adequate documentation will be provided in BioAgora depends on the data, its intended use and domain practices. It can be provided, for example, by including it in the respective publication/deliverable. The additional documentation can also be presented, for example in read-me-file and provided along with the respective data. Additional documentation for data is needed if the description in publications is not sufficient, there is not related publications or, in case of public data the metadata and data itself is not sufficient to ensure the understandability and validation of the data.

The documentation should be done accurately on the data in question. It may include, for example, detailed methodological description of how the data was collected and analysed information about the methods used for the collection and processing the data, description of variables, and more general information, e.g., about the project and purposes to produce the data. In case of interviews, for example, summaries, key topics of the analysis is collected in an excel file, included date, time and length of the interview, country and work sector of the interviewee is provided.

Public data, peer-reviewed and other publications to be agreed will be freely available to all in the repositories/archives. The re-use of these outputs will be promoted by publishing them under the Creative Common license CC-BY 4.0., which also preserves sufficient intellectual property rights for the authors/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 output and result 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. It is still to be determined whether we need other common platforms and tools, and how to manage their maintenance. 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 ensuring that partners are instructed to store and share only materials for which Teams is secure. In general, the ITC services of the partner organisations are responsible for providing secure storage and backup services for those working on the project. In line with good and secure practice, the backup is mainly done automatically on a regular scheduled basis. It is also good practice that the project participants can use password-protected servers in their organisations to store data and other materials, so that only authorised people have access to them.
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 in cases where data and other materials are shared within BioAgora and handled together. The needs of sensitive data, such as interviews, will also be taken into account and, if necessary, a specific storage and secure data transfer solution will be provided if, for example, non-anonymised interview data is to be shared with project partners.

7. Ethical and legal considerations

BioAgora is committed to good scientific conduct, governance and to valid legislation. Ethical standards and guidelines will be strictly applied. We will follow the guidelines issued by the ALLEA, the European Code of Conduct for Research Integrity. As the in the Description of the Action’s chapter 4. Ethics self-assessment of BioAgora it is stated that human participants are involved in the research, and we produce and use personal data. These data include interviews, survey and contact information data. In addition we collect and handle discussion data and diverse materials

from workshops and other events. All project partners are committed to generating and processing the personal data in accordance with good research practices, ethical guidelines and principles of EU General Data Protection Regulation (GDPR). Thus, we follow datamanagement practises that fully safeguard the identity of research subjects. The practices includes, for example that the participation in the study is voluntary, and the participants will give their informed consent to the participation. The participants are granted sufficient information (privacy notice) related to the study and all legal rights to the data concerning them, e.g. to review the data and stop participating in the data collection at any point in time. Also the data collection and handling methods as well as the data transfer will ensure the privacy: the collection of identifiers is minimized and only anonymised data is made openly available. Data containing personal identifiers will only be shared between project members using a secure transfer system. We also take into account the fact that the project involves participants from non-EU countries when sharing personal data.

Appendix 1. Partners of the BioAgora project

SYKE (Finnish Environment Institute)
University of Bucharest
University of Trento
INRAE (National Research Institute for Agriculture, Food and Environment)
INBO (Research Institute for Nature and Forest)
PBL Netherlands Environmental Assessment Agency (Ministry of Infrastructure and Water Management)
Norwegian Institute for Nature Research
Helmholtz Center for Environmental Research
Delbaere Consulting
Forschungsverbund Berlin e.V.
CREAF (Centro de Investigacíon Ecologíca y Aplicaciones Forestales)
ESSRG (Environmental Social Science Research Group)
Pensoft Publishers
ERCE PAN (European Regional Center for Ecohydrology)
Euronovia
Wageningen University and Research
European Citizen Science Association
Alfred Wegener Institute
UNEP-WCMC*
UK Center for Ecology and Hydrology*
* 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.