OASEES D1.2 Data Management Plan

Work package	WP1: Project Management
Task	T1.1 Project Coordination T1.2 Technical & Innovation Management T1.3 Monitoring and Quality Control
Authors	**Akis Kourtis - NCSRD**, Sofia Karamitsiani - NCSRD, George Xylouris – NCSRD, Christian Bolzmacher – CEA, Paride D'Ostilio – ASM, Davide Balducci – EMOT, M. Boğaç Kaya – SENSO, Marta Millet – ROBOT, Raed Bousalman – SCM, Andrea Fontalvo – CAP, Diego Cugat - CAP
Dissemination level	Public
Status	Final
Due date	30/06/2023
Document date	12/12/2023
Version number	1.0
	Funded by the European Union.

Revision and history chart

Version	Date	Main author	Summary of changes
1.0	30/06/2023		Final submitted version

Table of contents

Index of figures

Index of tables

LIST OF ABBREVIATIONS AND ACRONYMS

Acronyms/ Abbreviations	Description
AI	Artificial Intelligence
AMR	Autonomous Mobile Robot
CA	Consortium Agreement
CFS	Certificate on the Financial Statements
D	Deliverable
DAO	Decentralized Autonomous Organization
DMP	Data Management Plan
DPIA	Data Protection Impact Assessment
DPO	Data Protection Officer
EC	European Commission
GA	Grant Agreement
HITL	Human-in-the-Loop

IDS	International Data Spaces
IoT	Internet of Things
KPI	Key Performance Indicator
M	Month
ML	Machine Learning
OS	Open Science
PoC	Proof of Concept
SNN	Spiking Neural Network
T	Task
WP	Work Package

1. EXECUTIVE SUMMARY

The OASEES Data Management Plan (DMP) outlines a comprehensive strategy for managing critical data generated and collected during the project, while also optimizing access to and reusability of research data. The OASEES DMP follows the Horizon Europe DMP template, which is intended for use in any Horizon Europe project that produces, collects, or processes research-related data. The OASEES DMP adheres to the principles of FAIR and open access data in line with the Horizon Europe guidelines. The first DMP of the OASEES project details data management principles and strategies, tools and data types, ownership, intellectual property management, and access procedures. It also specifies procedures that will be implemented for data collection, storage, access, sharing policies, protection, retention, and destruction in accordance with EU standards as described in the Grant Agreement and the Consortium Agreement. The DMP provides insight into how the OASEES consortium intends to address the requirements of data management, protection, and security in the context of the current stage of the project. In line with the EU Open Access[1] practice, the OASEES project provides free online access to scientific information and results of specific project activities. These resources are also reusable for external users, as agreed in the Grant Agreement and Consortium Agreement. The OASEES consortium is committed to achieving open access when publishing papers and articles, thereby promoting the widespread dissemination and impact of research findings. The OASEES Data Management Plan (DMP) serves as the reference document that outlines the various procedures and strategies for managing the data generated and collected throughout the project. This document is advised upon to ensure that the data is effectively managed and accessible, both during and after the project. The OASEES consortium recognizes the importance of proper data management, and as such, the DMP is a critical component of the project's overall success.

[1] https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en

2. INTRODUCTION

2.1. PURPOSE AND SCOPE

The purpose of this deliverable is to outline the initial Data Management Plan (DMP) for the OASEES project, which aims to manage three major types of data: Administrative, Technical, and Use Case data. The project has a twofold objective: a) to utilize the data generated and collected within the project, and b) to disseminate OASEES scientific results as widely as possible. The development of this deliverable is based on the Guidelines for Open Access to Scientific Publications and Research Data in Horizon 2020, as well as the General Data Protection Regulation, and aligns with the Horizon Europe FAIR Data Management Plan. We offer guidance principles and specific details on the use case data descriptions, laying the groundwork for the dataset descriptions that will be included in our final DMP at the project's completion. The feedback received from all project partners has been consolidated in this deliverable, which includes provisions for the datasets contributed by each partner. Additionally, our plan for data security, adherence to FAIR principles, and legal and ethical considerations for data protection are outlined.

2.2. STRUCTURE OF THE DOCUMENT

The structure of this document is as follows:

- Section 1: Executive Summary
- Section 2: Introduction
- Section 3: Brief presentation of OASEES project
- Section 4: OASEES Data Management Plan Overview
- Section 5: OASEES Data
- Section 6: Fair Data
- Section 7: Allocation of Resources
- Section 8: Data Security
- Section 9: Legal, Data Protection and Ethical Aspects
- Section 10: Conclusion

3. BRIEF PRESENTATION OF OASEES PROJECT

Currently, data processing services are mainly centralized in the cloud and rely on large entities for IT infrastructure, limiting users from governing their data and managing their identities. In response to this issue, the OASEES project seeks to create an open, decentralized, intelligent, and programmable edge framework for Swarm architectures and applications. This project leverages the Decentralized Autonomous Organization (DAO) paradigm and integrates Human-in-the-Loop (HITL) processes for efficient decision-making. The OASEES vision is to provide open tools and secure environments for swarm programming and orchestration in various fields, in a completely decentralized manner. The project also aims to implement a portable and privacy-preserving ID federation system for edge devices and services that is fully compliant and compatible with GAIA-X federation and IDSA trust directives and specifications for identification and identity management. To address the lack of open management frameworks to address the heterogeneity of cloud computing and limited commercial solutions for hybrid core/edge management, the OASEES project aims to deliver and promote a European, fully open-source, decentralized, and secure Swarm programmability framework for edge devices. This project leverages various AI/ML accelerators such as FPGAs, SNNs, and Quantum while supporting a privacy-preserving Object ID federation process.

OASEES will manage the lifecycle of services across the compute continuum by orchestrating heterogeneous resources in the cloud, WAN, edge, and smart device domains. The project will promote the development of decentralized ML/AI edge services by means of an SDK and in the form of Decentralized Applications (DApps) in user-friendly notebook-style abstractions for data scientists and engineers.

The OASEES project will support multi-actor/multi-domain deployments by enforcing security and trustworthiness, enabling the federation with peer OS instances in other administrative domains (multi-domain operation), and fostering monetization by advertising/trading capabilities and resources in third-party Marketplaces (including the Marketplace of the European Open Science Cloud).

OASEES will be fully open-sourced, and its capabilities will be demonstrated in a diversity of proof-of-concept (PoC) deployments in six highly relevant vertical applications. The open-source community will also be invited to leverage its capabilities for building and managing innovative edge services. OASEES envisions a holistic approach for edge data processing that aims to disrupt current practices that heavily rely on non-European cloud AI data processing and push AI training and inference at the edge of the network while being vertical agnostic.

4. OASEES DATA MANAGEMENT PLAN OVERVIEW

The OASEES Data Management Plan (DMP) presents a comprehensive approach to managing critical data generated and collected throughout the project while optimizing accessibility and reuse of research-related data. The OASEES DMP is a living document that outlines how accumulated data will be handled during and after the project and will be reviewed and updated regularly. OASEES framework is designed with an open architecture and an open software stack, providing an integrated cross-domain data analysis platform that builds upon open cloud-oriented software tools and standards. OASEES will offer APIs that can be deployed on different data sources, promoting abstract application development related to new data analysis tools that can be easily integrated as plug-ins in the proposed analysis tools suite. The execution of all tools will be feasible through existing remote computing systems integrated into the infrastructure. OASEES platform's core includes an innovative and transparent AI/ML mechanism based on federated learning, natural language processing, and text mining techniques, enabling accessibility, interoperability, translation, transcription, and analysis of data. OASEES' novel explainable deep learning mechanism will extract knowledge from multidisciplinary data domains to empower the use and re-use of data for efficient and cost-effective decision-making procedures and workflows that contribute to a flourishing knowledge society. OASEES, as a European-funded project, will disseminate as much information and results as possible. All consortium partners are committed to providing open access to research data generated during the project unless it goes against their legitimate interests. The main purpose of the DMP is to ensure the accessibility and intelligibility of the data generated during the project.

Throughout the ongoing research efforts, collaboration and knowledge sharing have been the top priority to produce reusable and verifiable outputs to ensure OASEES' advancement and further the next-generation research. By adopting Open Science (OS) practices, OASEES concentrates its efforts on enriching the corresponding research area and ensures open-access research. Furthermore, the project's aim is to define and determine research plans to thrive at the earliest possible and share the plans publicly through preregistration. To ensure that the outputs and scientific publications under development meet the highest quality standards, all of them will be subject to peer review before the data collection process in the form of registered reports. OASEES also aims to make available to the public all the manuscripts demonstrating the results as pre-prints after the completion of the project. By making research and outcomes publicly available at the earliest possible moment, OASEES aims to contribute to the rapid pace of scientific progress. OASEES research outputs and data will comply with the FAIR principles. All scientific results, algorithms, tools, models, and publications will become openly accessible, except if restricted by data protection, security, confidentiality, and intellectual property rights, by deposition in trusted repositories widely known throughout the scientific community. OASEES will reinforce open peer-review practices of the articles

generated during the project to support utmost transparency during the reviewing process by making the identities of reviewers and authors known to each other and publishing reviews along with scientific publications. OASEES aims to involve relevant actors, from the conceptualization of the idea to its development, to contribute research outcomes that are beneficial to society and future generations. By adopting OS practices, the project aims to enrich public knowledge and support scientific evolution with insightful research results and technological advancements.

4.1. RESEARCH DATA MANAGEMENT AND MANAGEMENT OF RESEARCH OUTPUTS

Data generated within OASEES' use-cases will be collected in accordance with the FAIR principles, which include Findability, Accessibility, Interoperability, and Reusability. OASEES will manage digital assets generated through its processes by indexing data and metadata in searchable sources and assigning unique identifiers to ensure that they can be easily utilized by both humans and machines. OASEES will ensure that digital assets are easily retrievable and accessible. Reusability of digital assets will be achieved through clear and detailed descriptions for the data, as well as the provision of metadata that describes the context of data generation. OASEES will also censor metadata to protect data creator privacy when applicable. To demonstrate adherence to FAIR principles, OASEES will provide a living document detailing the consortium's plan for research data management during and after the project, with detailed information on the project data lifecycle, privacy, and policies for data collection, storage, access, sharing, protection, retention, and destruction (see Figure 1). WP1 and its respective tasks will maintain and update the data management plan and monitor/report on its implementation. Data management will be respectful of internal ethics and data protection measures. OASEES is expected to consider various data types, including qualitative and quantitative research data from pilots, administrative data such as participants' details, communications, and identity management data, data from public sources such as legislation, government guidance, codes of practices, and results of ethical horizon scanning, open-source data collected from publicly available sources that require ethical overview and agreement with owners, and publications and dissemination data related to open peer-reviewed publications, interviews, reports, proceedings, stakeholders, capacity programs, contact details for webinars/workshops, dissemination contacts, and inquiries. These data types will be categorized in more specific contexts in Section 5.

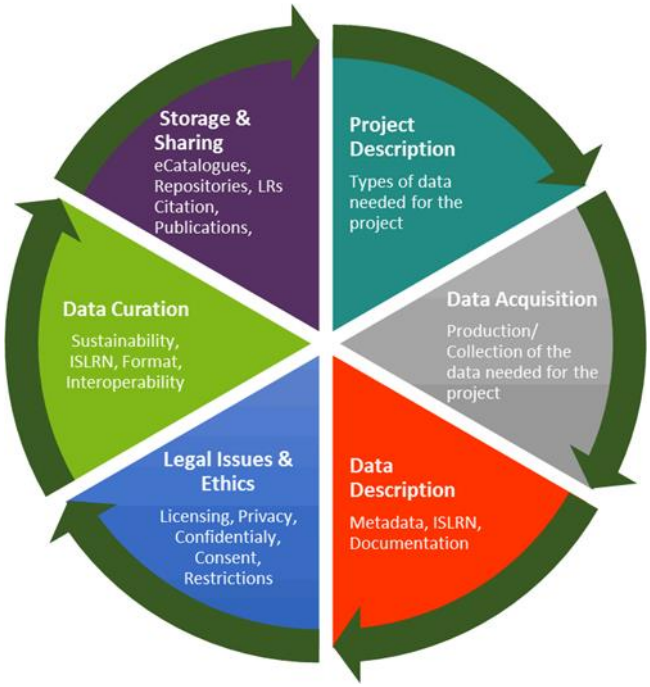

Figure 1: Data lifecycle

4.2. GUIDING PRINCIPLES

The Data Management Plan for OASEES is developed under Work Package 1 and adheres to the FAIR open access principles as outlined in the Horizon Europe guidelines, as depicted in Figure 1.

The primary principles for the DMP of OASEES are as follows:

I. The DMP is based on the "**Horizon Europe Data Management Plan Template**"[2] and has been created accordingly.
II. The DMP is an official project Deliverable (D1.1) due in June 2023, and it will be updated periodically based on significant changes and relevant reporting stages.
III. The DMP outlines how the OASEES consortium plans to address the data management, protection, and security requirements at the current stage of the project.
IV. The type of data, storage, confidentiality, data protection, and access procedures will conform to EU standards as specified in the Consortium Agreement and Grant Agreement. This includes policies for data collection, storage, access, sharing, protection, retention, and destruction.

4.3. DATA MANAGEMENTS STRATEGY AND DATE TYPES

The primary objective of the DMP for the OASEES project is to implement FAIR (Findable, Accessible, Interoperable, and Reusable) data management protocols. This plan outlines the approach for each data set that will be collected, processed, and/or generated in the project and includes the following elements:

- **Database reference source:** description of the data set to be produced, including OASEES use cases and pilot datasets.
- **Database/Dataset description:** description of the data generated or collected, including its origin, nature, and scale, as well as its potential usefulness to others and whether it is the basis for a scientific publication. Also, information on similar data and potential integration and reuse.
- **Standards and metadata:** reference to existing suitable standards or an outline of how and what metadata will be created if suitable standards do not exist. Specialized for each possible use case (WP2).
- **Data sharing/access:** outline of how data will be shared, access for enabling use, sharing, reuse, and whether access will be open or restricted to specific groups. Identification of any repository where data will be stored, indicating the type and security mechanisms (institutional, standard repository for the discipline, etc.). If the dataset cannot be shared, reasons will be stated (e.g., ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related).
- **Archiving and preservation procedures:** the plan for long-term preservation of the data, including how long the data should be preserved, the approximate end volume, associated costs, and how these costs will be covered.

4.4. OASEES DATA SUMMARY

OASEES aims to improve Self-Sovereign technologies by using Blockchain for decentralized identifiers and Homomorphic Encryption for cross-discipline data federation. The project seeks to provide a framework that allows data consumers to perform search queries across the federated OASEES data repositories without revealing their identity or equipment. Robust authentication and authorization mechanisms ensure secure communication, while the system only discloses the necessary data for any given transaction or interaction. OASEES prioritizes privacy preservation, green and responsible management, and empowers individuals and organizations to fully own their digital and analogue identities and control the sharing and use of their personal data. The project enables search in the encrypted domain, while maintaining data privacy and confidentiality through encryption. OASEES has also developed a detailed methodology to build high-quality data solutions and datasets more quickly and reliably. In this section, we summarize the data and address the purpose of data collection/generation, its relation to the project

[2] https://enspire.science/wp-content/uploads/2021/09/Horizon-Europe-Data-Management-Plan-Template.pdf

objectives, the types and formats of data generated/collected, the origin of the data, and the data's expected size and utility. We also present a summary of the current status, which will be further analysed in Section 5.

4.4.1. PURPOSE OF DATA COLLECTION/GENERATION AND ITS RELATION TO THE OBJECTIVES OF THE PROJECT

OASEES project will aim to collect and harmonize multidisciplinary data from several geographically distributed databases through the design and implementation of a multi-layered distributed database federation. Data-scraping services will be utilized to gather data from the appropriate databases. The collected data will then be processed and analysed by a data processing layer, which will create context-aware ontologies to enable complex and meaningful queries that can span across the OASEES data federation and various disciplines. The objective of the project is to collect data across Europe, harmonize them, and use them to achieve the OASEES objectives as specified in the grant agreement and summarized below. Section 5 will provide a detailed analysis of the current status of the project's data collection efforts, including the purpose of data collection, relation to the project objectives, types and formats of data, existing data being re-used, data origin, expected size of the data, and data utility.

4.4.2. RE-USE OF EXISTING DATA THROUGH OASEES

As stated, all OASEES collected data (except for private or prohibited data) will be available for research within the project's scope. The goal of OASEES is to aggregate diverse data types from multiple sources, using a multi-layered data fusion middleware to achieve a comprehensive and consistent representation of the data. Each layer of the middleware will represent a different data fusion approach to enable dynamic and complex processes. The middleware will serve as a mediator between incoming data from interdisciplinary domains, ensuring that all data is managed and stored in a uniform and meaningful way that complies with all national and EU ethics and legal requirements, as outlined later in this document.

4.4.3. TYPES AND FORMATS OF DATA THE PROJECT WILL GENERATE, USE OR RE-USE

In this deliverable, the different types of OASEES data that are relevant to the project's scope are discussed in detail in Section 5 on a per pilot case.

4.4.4. PURPOSE OF DATA GENERATION USE OR RE-USE AND THEIR RELATION WITH THE PROJECT

Section 5, along with its corresponding use case subsections, provides a detailed explanation of the rationale and objectives behind each type of data and dataset considered within OASEES, highlighting their connection with the overall goals and objectives of the project.

4.4.5. EXPECTED SIZE OF DATA

As OASEES progresses, the decision to make technical and use case data publicly available and accessible will be made by the owners of the respective technical components. The expected size of the data will depend on the extent and nature of the data that are made available for the OASEES purposes, which will be evaluated during the course of the project. The final version of the DMP will identify the size of the data accordingly.

4.4.6. WHAT IS THE ORIGIN/PROVENANCE OF THE DATA

The source of OASEES data is discussed comprehensively in Section 5 and illustrated in Figure 2. The data can be traced back to the administrative procedures of the project, the technical data required to establish the OASEES infrastructure, and the use-case data provided by each respective use-case leader (which will be discussed in detail in Section 5).

4.5. PARTICIPATION IN THE PILOT ON OPEN RESEARCH DATA

4.5.1. DATA AVAILABILITY

The primary objective of the project is to ensure that its research data is made publicly available, albeit with varying levels of access for different types of data. To safeguard sensitive data regulated by data protection laws, such as personal data, appropriate measures will be taken to obscure it. The notes, recordings, and survey results from project meetings and workshops will also be anonymized to protect the privacy of participants. The plan is to make the anonymized data publicly accessible, while technical details and results will be made available based on the sensitivity of the information. The following sections provide a detailed description of the data types and the applicable rules for each dataset.

4.5.2. OPEN ACCESS TO SCIENTIFIC PUBLICATIONS

Unless there are specific requirements or constraints that prevent it, all scientific publications resulting from the project will be publicly available and open access.

4.5.3. OPEN ACCESS TO RESEARCH DATA

In order to comply with the open access policy and ensure accessibility to the research and professional communities, research data that are not restricted by other rights will be uploaded and stored on the Zenodo platform as well as the EC publications and data repository. Zenodo, a digital repository, will be responsible for archiving and making the research data available.

4.6. EXPLOITATION, AVAILABILITY OF DATA, RE-USE AND ARCHIVE

OASEES has established guidelines, as part of WP6, for the dissemination of project results and developments and is committed to making as much material open access as possible. All consortium members have agreed to adhere to open access rules for research contributions, with the exception of confidential information or data subject to intellectual property rights or data protection regulations, as defined in the GA and CA. Accordingly, deliverables, publications, and reports will be made publicly available according to the terms outlined in the GA. As the project involves various stakeholders, including internal and external individuals and representatives from the research community, OASEES will ensure that ethical and security issues are addressed appropriately, as described in T1.4. OASEES will also provide special APIs to enable the use of its data through the OASEES platform while ensuring that privacy and ethical considerations are met.

The OASEES consortium is taking necessary steps to comply with national and EU legislation related to data manipulation. OASEES partners will make every effort and use organizational mechanisms to implement institutional and technical measures to protect the rights and freedoms of individuals who may be affected by data manipulation. They are committed to using state-of-the-art technologies to ensure secure storage, delivery, and access to personal information, as well as to manage the rights of OASEES solution users. All consortium members

guarantee that the data they curate, including administrative, technical, use case, and deliverables data, will be managed by authorized personnel at the right time, with clearly defined rights. The coordinator will ensure that all OASEES partners provide adequate information on these measures to ensure privacy and confidentiality for data collection, storage, access, sharing, protection, retention, and destruction in accordance with the GA and CA. In addition, private OASEES databases will be restricted to authorized personnel only and will have limited access during and after the project until the data is properly destroyed. The coordinator has established this service and holds an access log to ensure proper use of project data, as well as backup mechanisms.

5. OASEES DATA

In this section, we discuss the classification and mapping of data that will be created, collected, or obtained in the OASEES project. The data generated, obtained, or produced in the project will be classified into one of the major categories mentioned below, as illustrated in the figure.

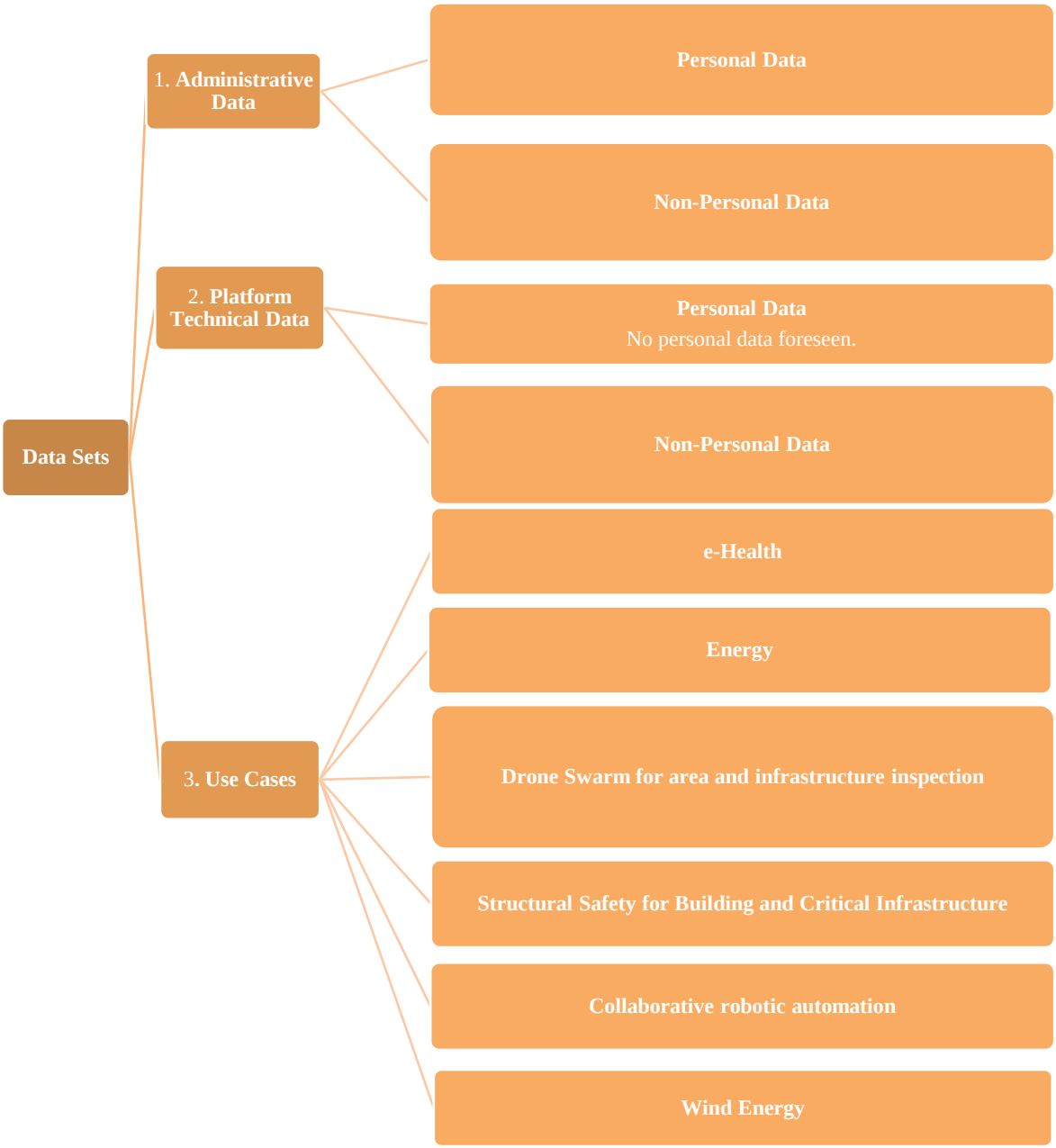

Figure 2: OASEES Date Categorization

5.1. DATA SET 1: OASEES ADMINISTRATIVE DATA

Administrative data in the OASEES project refers to information that is collected or generated for organizational, managerial, transactional, and record-keeping purposes. This data set encompasses various types of data, such as contact details of partners, project planning data including coordination of partners' work, documentation of communications among partners, templates of deliverables and reports, financial and administrative management data from each partner, as well as marketing, dissemination, and commercialization process data. The data is intended for internal use only and will be kept throughout the project's lifecycle. The data is created and stored in formats such as PDF, Word, Excel, and PowerPoint in a joint project's data repository created by the Project Coordinator on OASEES's private cloud (with restricted access limited only to consortium members) and locally on partners' servers and on the EC Portal. NCSRD has implemented a strict and limited access scheme that authorizes only pre-authorized personnel from consortium partners to access these repositories.

The data is made accessible to all OASEES consortium members through the data repository provided by NCSRD to ensure data reproducibility. The volume and velocity of this data set cannot be estimated accurately at this stage of the project, and it will be evaluated further and included in the final version of the DMP.

Pre-existing data that may be used in the OASEES project include publicly available templates and reports from similar projects, legal acts, and ethical guidelines that appear to be applicable to the project.

5.2. DATA SET 2: OASEES PLATFORM TECHNICAL DATA

The OASEES project also produces technical data which is created or collected during the technical development of the OASEES platform. This type of data falls under the non-personal data category and includes the following sub-categories:

- Source code of the developed components and services
- Data obtained from the analysis of state-of-the-art technologies
- Certification and auditing data

Similar to the administrative data, the technical data will be used for internal purposes and the public availability of this data will be determined by the respective owners of the technical components. As the project progresses, the size, format, velocity, and reproducibility of this data category will be more precisely defined and included in the final version of the Data Management Plan, since this stage of the project (M6) marks an intermediate milestone of the technical work packages (WP2-5) for the OASEES project.

5.3. DATA SET 3: OASEES PILOT DATA

5.3.1. E-HEALTH

The dataset for Pilot 1 for the Analysis of Voice, Articulation and Fluency disorders in Parkinson Disease comprises several data sources:

- **Parkinson disease patients.** Recordings of the patients' voice when they repeat or read standardized sentences, words, expressions. These recordings will be made during the ON phase when the symptoms are well alleviated by the levodopa treatment, contrary to the OFF phase when the treatment is no longer effective enough and the symptoms reappear in a very disabling way. The audio recordings are stored at HOPALE and CEA for AI training purposes. The variables characterizing intelligibility, fluency,

prosody... are extracted and constitute the only stored variables on the devices: fundamental frequency, maximum phonation time, harmonics/noise ratio, standard deviation of the fundamental frequency...

- **Control group.** The recordings of the voice of the control subjects when they repeat or read these same sentences, words and standardized expressions.
- **Patient-related data.** Data likely to influence the quality of the voice or to reflect the progress of the Parkinson's disease on the motor level: sex, age, weight, smoker/non-smoker, date of diagnosis of the disease, physiotherapy check-up, medication, presence of transcranial implants, heart rate and heart rate variability to evaluate the level of stress. This data will be anonymized.
- **Serious games**. Eventually, we plan to develop a self-education tool in the form of a serious game whose playful aspect will reinforce the patients' motivation to follow the proposed management. The data concerning the progress in the game, the history of the evolution of the patient's speech, and the parameterization of the game by the therapist will also be stored.
- **User tests.** User tests will be carried out at HOPALE and CEA in France. Collected data will be in form of usual audio formats and text files (in form of questionnaires). Anonymized data (features) can be shared with project partners.
- **Blockchain ledger.** The blockchain ledger serves as a secure and transparent platform for storing and sharing data coming from the recording and diagnostic edge device. Information such as device ID, patient ID, patient-related data, audio readings, extracted features, and other data. The blockchain ledger can be accessed by different stakeholders, such as edge device operators, speech therapists, and patients. It will support the decentralized identification of the edge devices and the organization of the pilot DAO operation.
- **AI/ML outputs.** The results of the real-time data analysis performed by the algorithms used on the edge devices. Raw speech data and extracted features will be stored at CEA and HOPALE for further analysis and verification.

Relationship to OASEES technologies

- **Data Processing at the Edge:** GPU edge accelerators will be used on the Parkinson Smart Edge-Connected Node. The raw data and processed outputs from these devices will be managed and processed by the OASEES edge framework.

- **Decentralized AI/ML Services:** A decentralized ML/AI edge services approach will be applied for this use case. Data will be generated at different locations at the patient’s home addresses or during hospital stay. The AI/ML output will be managed and optimized using the OASEES tools for decentralized AI/ML services.

- **Lifecycle Management of Services:** The framework aims to manage the lifecycle of services across the compute continuum. All the data generated by the sensors and the ML/AI outputs would be managed as part of this lifecycle, from creation and storage to processing and eventual archiving or deletion.

- **Security and Trustworthiness:** Security is a critical aspect of OASEES, particularly given the volatile and dynamic nature of edge infrastructures. Patients' data is one of the most sensitive data that needs to be protected accordingly. The blockchain ledger used in the pilot could be integrated with OASEES's security protocols to ensure data integrity and privacy.

- **Decentralized Autonomous Organization (DAO) and Human-in-the-Loop (HITL) Processes:** OASEES leverages the DAO paradigm and integrates HITL processes for efficient decision-making. The DAO is used in pilot 1 to oversee the data provided by the patients, update the exercises and to schedule

physical appointments if necessary. The HITL processes allows the speech therapists reviewing and making decisions based on the ML/AI outputs.

5.3.2. ENERGY

PoC2 dataset related to UC2 - EVs fleet coordinated recharging to support optimal operation of electricity grid comprises of several data sources:

- **Smart Meter data -** Measurements of voltage, currents, and power of consumers, producers and prosumers provided by ASM, collected through an MQTT protocol via public IP and / or LAN/VPN connections. To ensure secure access to the dataset, credentials will be required. Data format is JSON and data size is about 1 MB per day
- **Power Quality Analyzers data** - Measurements of voltage, currents, and power derived by ASM substations, collected through HTTP protocol only via LAN/VPN connections, which provide built-in security access features. Data format is JSON and CSV; data size is about 1KB per day.
- **Phaser Measurement Units data** - Measurements of voltage, currents, and power derived by ASM substations, collected through HTTP protocol only via LAN/VPN connections, which provide built-in security access features. Data format is JSON and data size is about 1 MB per day.
- **Charging Station data** - Real-time and historical data collected from the charging stations provided by EMOT. The data will include information about the charging station and each charging session (voltage, currents, power). Data format is JSON and data size is about 1 MB per day.
- **Electric Vehicle data** - Real-Time and historical data collected from electric vehicles provided by EMOT. The data will include battery capacity, state of charge, odometer, speed. Data format is JSON and data size is about 1 MB per day.

Relationship to OASEES technologies

- **Edge Data Processing:** OASEES provides an edge framework that can handle extreme data processing demands. This corresponds to the data collected from Smart Meters, Power Quality Analyzers, Phaser Measurement Units, Charging Stations, and Electric Vehicles in the EV pilot. This data, which is collected in real-time and with significant volume, can be processed at the edge, allowing for more efficient and responsive decision-making.
- **Decentralized ML/AI Services:** OASEES promotes the development of decentralized ML/AI services at the edge. This is directly relevant to the EV pilot, as the data generated is decentralized and requires local processing for optimal operation. For instance, Machine Learning algorithms could be employed to analyze data from the EVs and charging stations to optimize charging schedules and grid operations.
- **Lifecycle Management of Services:** The OASEES framework manages the lifecycle of services across the compute continuum. In the context of the EV pilot, this would include managing data from the various sources, such as smart meters, EVs, and charging stations. The data lifecycle would encompass stages from collection and storage to processing and eventual archiving or deletion.
- **Security and Trustworthiness:** Given the sensitive nature of the data collected in the EV pilot, such as power consumption data and EV status information, security is a critical concern. The OASEES framework emphasizes security and trustworthiness, which directly corresponds to the need for secure handling of the EV pilot data.
- **Decentralized Autonomous Organization (DAO) and Human-in-the-Loop (HITL) Processes:** OASEES leverages the DAO paradigm and integrates HITL processes for efficient decision-making. In the context of the EV pilot, this could involve human operators making decisions based on the collected data, such as adjusting charging schedules or grid operations. Meanwhile, the DAO paradigm could be used to automate

certain aspects of the EV fleet management, such as optimizing charging times based on grid conditions and EV usage patterns.

5.3.3. DRONE SWARM FOR AREA AND INFRASTRUCTURE INSPECTION

The dataset for Pilot 3 Drones swarm for area and infrastructure inspection comprises of several data sources:

• **Drone sensor data.** This data is generated by the sensors and cameras on the drones used for inspection. This data source includes images and video footage, as well as potentially other types of data depending on the sensors installed on the drones, such as thermal, infrared, lidar, or multispectral data. This data will be stored in a database or other storage system, with an API for retrieval and updates.

• **Drone flight logs.** These logs track the flight paths, distances, times, and other relevant flight data for each drone. This data could be used for performance tracking, compliance with flight regulations, and other analysis. Like the sensor data, this will be stored in a database or other storage system, with an API for retrieval and updates.

• **Blockchain ledger.** This could serve as a secure and transparent platform for storing and sharing data. It could store a variety of information, such as flight paths, sensor readings, inspection results, and other data. The blockchain ledger can be accessed by different stakeholders, such as drone operators, telecommunications companies, and regulators. It will support the decentralized identification of the drone devices and the .organization of the pilot DAO operation.

• **Infrastructure metadata.** This could include information about the telecommunications infrastructure that is being inspected, such as the location, size, construction year, and other relevant details. This data will be stored in a database or other system and linked to the drone inspection data.

• **Spiking Neural Network (SNN) outputs.** These are the results of the real-time data analysis performed by the SNNs on the drone. This will be stored alongside the raw sensor data for further analysis and verification. Spiking Neural Networks are a type of artificial neural network that models the way neurons in the brain communicate with each other using short bursts of electrical activity, known as spikes. Unlike traditional neural networks that process information continuously, SNNs process information in a discrete, event-driven manner, which can make them more efficient for certain tasks. In the context of a drone inspection pilot, SNNs will be used to analyze the data collected by the drone's sensors in real-time. This could include image data, lidar data, or any other type of sensor data that the drone collects. The SNN would process this data and generate outputs that represent its analysis.

• **Communication data.** Information about the communication links between the drone and the ground station. This could include data about the quality of the communication link, any disruptions or issues, and other relevant information. This could be stored in a database or other system for troubleshooting and analysis.

Relationship to OASEES technologies

- **Data Processing at the Edge:** The OASEES project emphasizes the need for handling extreme data processing demands at the edge, using various edge accelerators like GPU, NPU, SNN, and Quantum. This is directly related to the data sources from devices located at the edge, such as drones in the inspection pilot. The raw data and processed outputs from these devices, including the Spiking Neural Network (SNN) outputs, will be managed and processed by the OASEES edge framework.

- **Decentralized AI/ML Services:** OASEES aims to promote the development of decentralized ML/AI edge services. This aligns with the decentralized nature of the data sources in the drone inspection pilot, as data

is generated in different locations by individual drones. The SNN outputs, which can be considered a form of AI/ML output, can be managed and optimized using the OASEES tools for decentralized AI/ML services.

- **Lifecycle Management of Services:** The framework aims to manage the lifecycle of services across the compute continuum. This includes managing data sources from cloud to edge devices, such as drones. All the data generated by the sensors and the SNN outputs would be managed as part of this lifecycle, from creation and storage to processing and eventual archiving or deletion.

- **Security and Trustworthiness:** Security is a critical aspect of OASEES, particularly given the volatile and dynamic nature of edge infrastructures. This relates directly to the data sources in the inspection pilot, as the drone sensor data, flight logs, and SNN outputs all need to be securely stored and transmitted. The blockchain ledger used in the pilot could be integrated with OASEES's security protocols to ensure data integrity and privacy.

- **Decentralized Autonomous Organization (DAO) and Human-in-the-Loop (HITL) Processes:** OASEES leverages the DAO paradigm and integrates HITL processes for efficient decision-making. This will directly relate to the drone inspection data. For instance, the HITL processes could involve human operators reviewing and making decisions based on the SNN outputs and other drone data. Meanwhile, the DAO paradigm could be used to automate certain aspects of the drone inspections, such as scheduling flights or flagging areas for further inspection based on the sensor data and SNN outputs.

5.3.4. STUCTURAL SAFETY FOR BUILDING CRITICAL INFRASTRUCTURE

The dataset for monitoring structural safety of critical infrastructure comprises of various data sources:

- **Structural characteristics data as thresholds.** This is a set of data relevant to the earthquake limits of the structures, that need to be embedded as thresholds to the monitoring software. Once each threshold is exceeded, different information flow and action follow.
- **Run logs.** The decision support software runs in every couple of minutes via a watchdog that detects if there has been an earthquake in the last couple of minutes. There are logs kept for the results of the watchdog, whether a threshold is exceeded or not, and about the sensor behaviour.
- **Communication data.** One a threshold earthquake is exceeded, relevant people are informed via e-mail and SMS. Those communications are kept as backup data.
- **Reporting data.** Once a threshold is exceeded in a structure, a technical report is prepared automatically. These reports are stored in the location where the decision support software is running.
- **Sensor data.** Once a threshold is exceeded, the relevant data from all the sensors are downloaded or copied in a specific folder. Usually, the time the earthquake occurred is detected, and all sensor data 0.5minute before and 2.5minute after that exact time are cropped, merged and stored.

Relationship to OASEES technologies:

- **Data Processing at the Edge.** Edge-based data processing is a fundamental aspect of the decision support in critical structures, requiring the use of multiple edge accelerators like GPU, NPU, SNN, and Quantum. The processing ability at the edge allows direct data interpretation, and data and decision exchange with other devices and even other and geographically distributed client locations. The OASEES edge framework will handle and process the raw seismic data at the location, and based on the interaction with the other nodes in swarm mode, it will give output for a safe post-earthquake decision-making.

- **Decentralized AI/ML Services**. Decision of whether an earthquake occurred or not, where and what magnitude it was, and how the effects on a specific structure/location was, will all be processed using AI/ML models that will be trained at the edge. Digital twin, which can be updated as the data flow in, will then be possible.
- **Lifecycle Management of Services.** The lifecycle management of services encompasses the comprehensive handling of data, from its generation by sensors to its storage, processing, and eventual archiving or deletion.
- **Security and Trustworthiness.** To ensure data integrity and confidentiality, the pilot implementation will integrate the blockchain ledger with OASEES's security protocols.
- **Decentralized Autonomous Organization (DAO) and Human-in-the-Loop (HITL) Processes**. OASEES adopts an approach that enhances decision-making efficiency by combining a decentralized autonomous organization (DAO) model with human-in-the-loop (HITL) processes. This is an extremely important aspect for critical infrastructure, since the operational process is divided to two streams, i) first the output from a autonomous decision support system, and ii) properly informed human getting into the loop to take the action. For example, a detailed report is produced by the decision support software, DAO in the case of OASEES, and an alarm is produced (yellow, orange or red). The person in charge at the tunnel, where Senso monitoring system is working, will then decide to close the barrier and stop the traffic and new vehicle entrance to the tunnel or not. The DAO-HITL sequence of OASEES perfectly matches this concept.

5.3.5. COLLABORATIVE ROBOTIC AUTOMATION

The dataset for collaborative robotic automation comprises of several data sources:

- **Autonomous Mobile Robot (AMR) sensor data.** This data is generated by the sensors and cameras on the AMR. This data source includes images, videos, LiDAR, GPS, and odometry data. This information can be accessed via the robot's computer.
- **Arm logs.** These logs track the movements, warnings and errors of the collaborative arm. This data could be used for performance tracking, and other process analysis. Like the AMR data, this information can be accessed via the arm's computer.
- **Blockchain ledger.** This technology has the potential to function as a reliable and open system for the storage and exchange of data. Its versatility allows for the inclusion of various types of information, including maps, quantities of reliable or substandard components, and user authentication. Multiple entities, including operators of the robotic systems, wood processing firms, and other personnel, could access the distributed ledger.
- **Wood pieces metadata.** This could include information about the wood pieces that are being processed: size, composition, time of processing, quality control, etc. This data can be stored in a database or other system.
- **Computer Vision outputs.** These are the results of the real-time data analysis performed by the quality control camera in the working table. In the context of the collaborative robotic automation pilot, this will mainly include image data. The AI would process this data and generate outputs that represent its analysis.
- **Communication data.** Information about the communication links between the AMR and the robotic arm, like the quality of the communication, disruptions or issues.

Relationship to OASEES technologies:

- **Data Processing at the Edge.** Edge-based data processing is a crucial aspect of the OASEES project, necessitating the utilization of diverse edge accelerators like GPU, NPU, SNN, and Quantum. This focus

directly pertains to data derived from edge devices such as AMR and collaborative robotic arms. The OASEES edge framework will handle and process both the raw data and the resulting outputs from these devices.

- **Decentralized AI/ML Services**. The management and optimization of Computer Vision outputs, which can be seen as a form of AI/ML output, can be accomplished using OASEES tools designed for decentralized AI/ML services.
- **Lifecycle Management of Services.** The lifecycle management of services encompasses the comprehensive handling of data, from its generation by sensors to its storage, processing, and eventual archiving or deletion.
- **Security and Trustworthiness.** To ensure data integrity and confidentiality, the pilot implementation could seamlessly integrate the blockchain ledger with OASEES's security protocols.
- **Decentralized Autonomous Organization (DAO) and Human-in-the-Loop (HITL) Processes**. OASEES adopts an approach that enhances decision-making efficiency by combining a decentralized autonomous organization (DAO) model with human-in-the-loop (HITL) processes. In this context, human operators review and make decisions based on quality control camera results, while specific aspects of robotic processing, such as path scheduling or the selection of wood pieces for further processing, are automated using the DAO paradigm.

5.3.6. WIND ENERGY

The following data sources shall be considered in the dataset for PoC 6: Smart Swarm Energy harvesting and Predictive Maintenance Wind turbines:

- **Device Sensor Data**: This data source includes the audio data acquired with the microphones and environmental data such as wind speed, wind direction, temperature and humidity. This data will be stored in a hard drive storage in the device and it may be accessed remotely.
- **System logs:** The device will keep information about internal errors and events with the associated timestamp, the battery status and the operation of the wind turbine.
- **Audio processed data**: The data of the filtered audio signal to remove wind influence (noise), and audio features calculated for the audio processing algorithm that can be employed also on the failure detection algorithm. This data will be stored in the device and sent to the failure detection and predictive algorithms.
- **Blockchain ledger:** The data that will be published in the blockchain is related to the Stakeholder interest. This includes the reports of the failure detection, prediction algorithms, blade status and audio timeseries for a specific or multiple wind turbines. Also, the HITL decisions will be evaluated in the blockchain, for unexpected anomalies that the algorithms may not be able to classify.
- **Wind turbine metadata:** Wind turbine infrastructure data is not priority data, but it may help on wind turbine identification and to compare results with similar wind turbines. The availability of this data will depend on wind farm data policy, but some of the possible information will be the wind turbine model, manufacturer, tower height, rotor diameter, power generation, year of installation, and other additional information if it is provided, such as design or operational parameters.
- **Communication data:** The information about the communication between IoT devices for the Federated Learning configuration and the information of the communication with other related OASEES edge nodes and cloud services.

This data sources are related with the following OASEES technologies:

- **Data Processing at the Edge:** An important aspect that OASEES takes on account is the need for process computation of the data at the Edge using different edge accelerators. In the Wind Energy pilot, part of the

data will be processed in the devices to improve the quality of the analysed audio signal that will be used for failure detection and prediction. Furthermore, the identification and predictive algorithm will be running in the smart nodes, i.e on the available infrastructure in charge to supply the computational resources to run and train the identification and prediction algorithms.

- **Decentralized AI/ML Services**: From the OASEES perspective, decentralized machine learning shall be applied to data sources with decentralized nature, which fits with the swarm IoT device infrastructure that can be deployed in a wind farm.
- **Lifecycle Management of Services:** The framework aims to manage the lifecycle of services across the compute continuum. This applies to the wind energy use case as some services will be deployed in the edge devices such the audio processing algorithms, and some services will be deployed in the cloud such as the failure prediction algorithms.
- **Security and Trustworthiness:** Security is a fundamental aspect for the OASEES data management, such all the data sources need to be securely stored and transmitted through the OASEES environment. In the wind energy will be essential to protect and control the failure and prediction results, because wind turbine status can be very sensitive data for a wind farm owner.
- **Decentralized Autonomous Organization (DAO) and Human-in-the-Loop (HITL) Processes:** On one hand the DAO paradigm can be used in the wind energy pilot for create a secure environment where stakeholders can access to the data reports and propose polls with the main objective of improve the service with new functionalities in the algorithms, final reports, user interaction, and many other functionalities. On the other hand, the IoT devices could interact with the DAO through voting in the polls for temporally disconnection, stand by modes for power supply saving, and other functionalities that will be defined in next steps. Also, HITL processes will be part of the DAO to make decisions about unexpected data anomalies and will there be some automated decision making from the devices, such as scheduled and unscheduled disconnections of the devices.

6. FAIR DATA

OASEES aims to establish a systematic approach for developing high-quality data solutions and datasets more efficiently and dependably. The success of the project is contingent not only on the speed but also on the quality of the input data. OASEES promotes critical cross-functional collaboration and automation to construct rapid and dependable data pipelines. The project is anchored on three pillars of innovation: Foster Collaboration, Build Trusted Data Solutions, Automate Testing and Monitoring. OASEES seeks to merge technological and social innovations for secure and sustainable data operations, such as optimizing/minimizing/decentralizing data processing, transfer, and storage, and avoiding unnecessary data manipulations, as well as implementing technologies and solutions for fair and ethical collection, processing, and manipulation of data. This approach aligns with the principles of responsible/trustworthy AI by using a co-development methodology as the backbone.

OASEES proposes a secure-by-design Federated Platform in line with the EU data strategy (COM (2020) 66) and the main EU reference architectures (GAIA-X, EOSC, EGI) in the sector, capable of ensuring interoperability, enabling cross-border scenarios, and scaling various AI-based applications by using open APIs. The aim is to make the EU the most secure and trustful data hub in the world. To achieve this, OASEES leverages an innovative homomorphic approach that ensures user-friendly, safe, trustworthy, compliant, fair, transparent, accountable, and sustainable collection, storage, processing, querying, and delivery of data.

All consortium partners are committed to ensuring that the data produced, collected, and processed aligns with the FAIR Principles definition[3], making it findable, accessible, interoperable, and reusable. To achieve this, OASEES proposes the following:

I. **Making data findable** – The DMP should provide details on how to make data easily discoverable, identifiable, and referenceable through standard identification mechanisms. This includes specifying naming conventions, outlining the approach towards search keywords, ensuring clear versioning, and defining standards for metadata creation (if applicable).
II. **Making data openly accessible** – The DMP must clearly state which data will be accessible to the public and which data will be kept restricted, along with the rationale for any restrictions. Additionally, the plan should explain the methods or software required to access the data, where the data, metadata, documentation, and code will be deposited, and the steps for accessing restricted data, if any.
III. **Making data interoperable** – The DMP needs to specify which data and metadata vocabularies, standards, or methodologies will be employed. Additionally, it should be addressed whether all data types in the datasets will follow standard vocabulary to enable inter-disciplinary interoperability.
IV. **Increasing re-usability** - The DMP should provide documentation regarding data licensing, specifying the conditions under which the data can be reused by others, including the timing and availability of the data after the project's end. Furthermore, it should indicate the duration for which the data will remain reusable, as well as any data quality assurance processes, if relevant.

The OASEES platform is dedicated to promoting responsible and environmentally sustainable data management and processing through advanced digital technologies. The platform leverages cutting-edge solutions to ensure compliance with data privacy regulations and secure data preservation. The ultimate goal of OASEES is to facilitate seamless communication and exchange of data among various EU data spaces through the use of open science policies and procedures. The platform offers user-friendly and transparent ICT services for data collection, storage, processing, and delivery. The use of homomorphic encryption technology ensures secure sharing and manipulation of data in compliance with GDPR and other EU legislation. OASEES prioritizes social innovation and privacy impact assessment by minimizing unnecessary data manipulation while adhering to responsible and trustworthy AI principles. The platform is designed using a Software Oriented Architecture to allow for future upgrades and evolution. To provide a comprehensive set of information services to different user groups, OASEES integrates existing domain knowledge and related services.

Each of the data types outlined in Section 5 has its own corresponding solutions for the issues mentioned above. It is important to note that the approach towards these solutions may be subject to change as the project progresses and evolves. The consortium is currently discussing these issues, and decisions will be made at a later stage of the project, which will be reflected in future updates to the DMP or in other Deliverables.

Throughout the duration of the project, the principles of FAIR data management will be implemented to the fullest extent feasible for both the infrastructure's data and AI models, while adhering to national and European legal frameworks, data protection regulations, and ethical standards. This section is established on the directives for proficient data management during a Horizon Europe project, as recommended by the European Commission[4].

[3] Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18, https://www.nature.com/articles/sdata201618

[4] https://webgate.ec.europa.eu/funding-tenders-opportunities/display/OM/Online+Manual

6.1. MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA

6.1.1. ADMINISTRATIVE DATA

The initial stage of the FAIRification process involves ensuring that data can be located within large pools of information. To achieve this, both data and metadata should be easily identifiable by humans and machines alike. The OASEES website (oasees-project.eu) will list all the project's deliverables, and we will use social media and other appropriate channels to publicize how the output can be accessed. Furthermore, for public deliverables, a link will be established between the project website and the relevant open repositories where the data is submitted.

6.1.2. PLATFORM TECHNICAL DATA / USE CASE DATA

As the OASEES project advances, the owners of the technical components will make decisions regarding the public availability and open access of technical data. In accordance with these decisions, the final version of the Data Management Plan (DMP) will provide further details on how to make this data more findable.

6.2. MAKING DATA ACCESSIBLE

This principle emphasizes the importance of providing users with easy access to data, which may require authentication and authorization. The OASEES project is not foreseen to process, store, modify any confidential data. It should be noted, that in the frame of OASEES in a project-wide policy aspect, personal data, particularly special categories of personal data as defined in Article 9(1) of the GDPR, may only be made accessible in accordance with data protection requirements. Access to confidential data may only be granted in accordance with the confidentiality requirements outlined in the Grant Agreement and Consortium Agreement. As multiple repositories may be used to store data, a policy outlining how to grant access to restricted results will be developed throughout the OASEES project and documented in future DMPs.

6.2.1. ADMINISTRATIVE DATA

In order to ensure that the ongoing scientific publications and outputs meet the highest standards of quality, they will undergo peer review in the form of registered reports prior to the data collection process. The administrative data, on the other hand, are accessible only to the OASEES consortium members and are stored on the local cloud provided by the coordinator (NCSRD) for internal use within the project.

6.2.2. PLATFORM TECHNICAL DATA / USE CASE DATA

As the OASEES project advances, decisions regarding the public availability and open access of technical and use case data will be determined by the owners of the corresponding technical components. Consequently, the accessibility of these data will be further developed and explained in the final version of the Data Management Plan (DMP).

6.3. MAKING DATA INTEROPERABLE

6.3.1. ADMINISTRATIVE DATA

The OASEES consortium members will adhere to common standards and formats for the creation and sharing of administrative data, which will be stored on a cloud repository provided by the coordinator. Any necessary modifications can be made by individual partners and shared with the rest of the consortium. Additionally, email communication will be used to exchange data as needed. To ensure consistency, references to pre-existing data or templates will be included in deliverables, reports, and state-of-the-art analyses stored on the local data repository provided by the coordinator.

6.3.2. PLATFORM TECHNICAL DATA / USE CASE DATA

As the OASEES project moves forward, decisions regarding the public availability and open access of technical and use case data will be made by the owners of the respective technical components. To ensure interoperability, which involves the ability of different systems to communicate and exchange data, the final version of the Data Management Plan (DMP) will provide more details on how to achieve this. This may involve providing documentation or other means to facilitate the sharing and use of the data.

6.4. INCREASE DATA RE-USE

6.4.1. ADMINISTRATIVE DATA

The publicly accessible deliverables of the project will be published on Zenodo, in accordance with the open data strategy. The deliverables will be structured and standardized according to the prescribed standards, as indicated in the deliverables. To ensure that the quality of the disseminated deliverables conforms to the standards established for Horizon Europe projects and those set by the scientific community, the quality assurance processes outlined in D1.1 "Project handbook" will be employed.

These publicly available deliverables will be accessible and can be reused by external stakeholders, even after the conclusion of the project.

6.4.2. PLATFORM TECHNICAL DATA / USE CASE DATA

As the OASEES project advances, the decision regarding the public availability and open access of technical and use case data will be made by the owners of the corresponding technical components. Accordingly, the extent to which these data can be reused will be further elaborated in the final version of the DMP. The primary objective of the FAIRification process is to enhance the data reusability. To achieve this objective, both metadata and data must be thoroughly described and adequately licensed to allow for replication and/or exploitation in different contexts.

7. ALLOCATION OF RESOURCES

Our projection is that there will be no cost associated with making the OASEES data sets FAIR. This is due to the fact that Zenodo, which is an EU-funded and supported repository, provides its service free of charge. However, private repositories such as the project repository supported by the coordinator may require a fee, which the OASEES consortium partners will cover using their own budgets. The general coordination and supervision for data management in the project will be the responsibility of the coordinator and technical manager, as outlined in our handbook. Consortium partners are required to ensure that their activities align with all applicable local, government, and international laws, regulations, and guidelines related to the project workplan.

8. DATA SECURITY

8.1. ADMINISTRATIVE DATA

The OASEES consortium members will restrict the sharing of administrative data to within the consortium. The project coordinator, NCSRD, has established a SharePoint-based project-collaboration space, and a mailing list on a local cloud and data repository. This private space has been designed as a document repository to be accessed exclusively by consortium partners. This will allow them to access and share research and project documentation, from final deliverables to presentations, internal documents, and other relevant information.

8.2. PLATFORM TECHNICAL DATA

The datasets related to the OASEES project, such as software code of different components specified in WP4 and WP5, will be kept in a secure cloud infrastructure that can only be accessed by authorized members of the consortium. A backup strategy will be developed and applied in a secure storage server. If the data is stored in an external certified repository, the security standards of that repository will be followed. Additionally, when designing the system architecture, privacy by design principles will be taken into account and secure communication protocols will be established.

9. LEGAL, DATA PROTECTION AND ETHICAL ASPECTS

Ensuring ethical compliance is a top priority for all research activities supported by the European Union. Consequently, ethical evaluation is necessary right from the proposal's conceptual stage, and the activities carried out under the Horizon Europe Framework must adhere to ethical principles from start to finish. For projects such as OASEES, adhering to legal, data protection, and ethical considerations is essential. The objective of this document is to identify any potential issues that could affect data sharing within the OASEES consortium. In Deliverables D1.3, and under WP1, the legal and ethical requirements for designing and developing the OASEES platform, as well as those that future users of the platform must comply with, will be thoroughly addressed.

9.1. PERSONAL DATA IN OASEES

The OASEES project strictly adheres to a comprehensive set of rules when processing personal data, as required by the General Data Protection Regulation (GDPR) of the European Union[5]. According to Article 4(1) of the GDPR, personal data refers to any information relating to an identified or identifiable natural person, including factors specific to their physical, physiological, genetic, mental, economic, cultural or social identity. Only natural persons, including employees of businesses and public authorities, are protected under the GDPR. To distinguish between personal and non-personal data, which includes anonymous data, is crucial for all project activities. The project does not generate or collect any personal.

The GDPR applies to organizations established in the EU that process personal data, whether they are functioning as processor or controller. The GDPR also applies to companies that are not in Europe under certain conditions, such as when data processing activities are related to offering goods or services to data subjects situated in the EU or monitoring the behaviour of such data subjects[6]. In the OASEES project, 21 partners are established in the EU

[5] Regulation (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

[6] GDPR, Art 3.

and are thus subject to GDPR rules. The project will ensure compliance with legal and ethical requirements concerning the protection of personal data.

9.2. PRINCIPLES OF THE PROCESSING OF PERSONAL DATA

The GDPR's Article 5 outlines seven principles that must be adhered to whenever personal data is processed. By following these principles, the legal requirements for data protection can be met, and a sufficient level of accountability and protection can be maintained. Moreover, this approach will help safeguard the rights of data subjects.

The following principles must also be considered in the OASEES project:

I. **Lawfulness, fairness, and transparency** - Processing of personal data must be legal, fair, and transparent to the data subject. The data subjects must be informed about the processing (as required in Article 12 to 14 GDPR).
II. **Purpose limitation** - Personal data can only be processed for the purposes specified at the time of data collection, and any further processing must be compatible with those purposes.
III. **Data minimization** - Only relevant personal data should be processed. Pseudonymization and anonymization of data should be used whenever possible.
IV. **Accuracy** - Personal data must be accurate and up-to-date.
V. **Storage limitation** - Personal data should only be stored for as long as necessary and deleted or anonymized afterwards, unless an exception applies.
VI. **Confidentiality, integrity, and data security** - Appropriate security measures must be in place to protect personal data from unauthorized or unlawful processing, accidental loss or damage, and to ensure data integrity and confidentiality.
VII. **Accountability** - The data controller is responsible for demonstrating compliance with all of the above principles.

This document will provide some explanation of the essential terminology in the context of the GDPR principles specific to the OASEES project and the roles that partners may play.

9.2.1. LAWFULNESS AND LEGAL BASIS

The first principle outlined above mandates that any processing of personal data within the OASEES consortium must be grounded on at least one of the legal bases specified in Article 6 of the GDPR. In addition, special categories of personal data, as defined in Article 9 of the GDPR, must be processed based on one of the exceptions listed in Article 9(2) of the GDPR. The legal basis for processing must be established before the initiation of a processing activity, and it must be deemed necessary for a specific purpose. To fulfil this requirement, each party processing the data must identify a legal basis for the processing.

The following are the legal bases defined in Article 6 of the GDPR:

a) The data subject has given consent to the processing of their personal data for one or more specific purposes (Article 6(1)(a) GDPR);
b) Processing is necessary for the performance of a contract to which the data subject is a party or in order to take steps at the request of the data subject prior to entering into a contract (Article 6(1)(b) GDPR);
c) Processing is necessary for compliance with a legal obligation to which the controller is subject (Article 6(1)(c) GDPR);
d) Processing is necessary in order to protect the vital interests of the data subject or of another natural person (Article 6(1)(d) GDPR);

e) Processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller (Article 6(1)(e) GDPR);
f) Processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require the protection of personal data, particularly where the data subject is a child (Article 6(1)(f) GDPR).

When processing personal data, it is essential for the consortium partners to identify the appropriate legal basis with the assistance of their respective Data Protection Officers (DPOs) or legal departments, along with support from NCSRD. The personal data should be processed explicitly within the framework of this legal basis and for the purposes associated with it. In cases where none of the mentioned legal bases are applicable, the processing of personal data is not allowed, and the project must refrain from all personal data processing without a valid legal basis.

9.2.2. DATA MINIMISATION

In the OASEES project, the principle of data minimisation should be followed when dealing with personal data. Consortium partners must limit the processing of personal data to what is strictly necessary for the intended purpose, and not process beyond this limit. When acting as data controllers, they should be able to provide a clear and justifiable reason for collecting and retaining personal data or, alternatively, consider using aggregated, anonymised, or pseudonymised data[7] (if adequate protection is provided).

If data can be anonymised and still serve the intended purpose, it is not necessary to process identifiable personal data.

9.2.3. ACCURACY

As previously mentioned, this principle mandates that personal data processed should be precise and continuously updated to ensure accuracy. Personal data can be considered inaccurate if it is misleading or incorrect in any way. The consortium partners must adhere to this principle when handling personal data within the OASEES project and should delete or correct any personal data that is no longer valid or accurate.

9.2.4. STORAGE LIMITATION

This principle stipulates that personal data must not be stored for any longer than is strictly necessary[8]. Once the purpose for which the data was processed has been achieved, it should be deleted, except where retention is required by law (as specified in Article 5(1)(e) and Article 89 GDPR). Consortium partners have the responsibility to assess the duration for which the personal data, provided by the partners or external stakeholders, will be stored on their own local repositories, and on the joint repository provided by NCSRD.

9.2.5. CONFIDENTIALITY, INTEGRITY AND DATA SECURITY

[7] Article 29 Data Protection Working Party, Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679 (WP 251 2017) at 11.

[8] GDPR, Recital 39.

The GDPR contains a principle[9] that plays a critical role in data security. Maintaining the security of personal data is an essential component of protecting the right to data privacy. Both the controller and processor bear the responsibility of assessing relevant risks associated with data processing and implementing appropriate measures and techniques to ensure the security, confidentiality, and integrity of the data.

[…] the controller should adopt internal policies and implement measures which meet in particular the principles of data protection by design and data protection by default. Such measures could consist, inter alia, of minimising the processing of personal data, pseudonymising personal data as soon as possible, transparency with regard to the functions and processing of personal data, enabling the data subject to monitor the data processing, enabling the controller to create and improve security features. When developing, designing, selecting and using applications, services and products that are based on the processing of personal data or process personal data to fulfil their task, producers of the products, services and applications should be encouraged to take into account the right to data protection when developing and designing such products, services and applications and, with due regard to the state of the art, to make sure that controllers and processors are able to fulfil their data protection obligations.[10]

Examples of technical measures that can be implemented include the use of multi-factor authentication or end-to-end encryption by employees who have access to personal data. Organizational measures include staff training, adding a data privacy policy to the employee handbook, and limiting access to personal data on a need-to-know basis with proper authorization. The principles of data protection by design and by default (Article 25 GDPR) should also be incorporated into the project at all stages of development, especially during the establishment of the OASEES solution. The OASEES partners are required to apply technical solutions that ensure data confidentiality, integrity, and security in compliance with the Grant Agreement and Consortium Agreement. The measures will be evaluated by the Coordinator and Technical Manager, following the procedures outlined in D1.1 "Project Handbook." Additional legal and data protection requirements for the system architecture of OASEES data space will be included in D1.3.

9.2.6. ACCOUNTABILITY

The GDPR's principle of accountability mandates that data controllers establish and implement appropriate technical and organizational measures and be able to demonstrate their compliance and effectiveness when requested. To this end, the OASEES consortium should be prepared to implement a comprehensive personal data protection system capable of mitigating the risks associated with project activities and substantiating GDPR compliance. This involves allocating sufficient resources to budget, compliance tools, procedures, staffing, technology, and security. A set of instruments that enable accountability includes maintaining accurate documentation on the processing of personal data (such as what data is collected, how, for what purpose, for how long, how it is used, where it is stored, and who is responsible for processing it), implementing documented processes and procedures that address data protection issues early on in the project, designating data protection responsibilities to the team, having appropriate Data Processing Agreements in place with third-party data processors, and appointing a Data Protection Officer if required by law. The OASEES consortium should consider these instruments whenever personal data processing is involved.

9.3. DATA SHARING

[9] GDPR, Art. 5(1)(f).

[10] GDPR, Recital 78.

In compliance with the General Data Protection Regulation and the Grant Agreement and Consortium Agreement, the consortium partners will execute Data Sharing Agreements, when necessary, prior to sharing data within the consortium. However, as of the current stage of the project (M6), it does not appear that Data Sharing Agreements such as Joint Controllership Agreements or Data Processing Agreements are necessary. Administrative data, which is being processed by each partner for their own purposes within the scope of the OASEES consortium (e.g. information exchange), is covered by the Grant Agreement and Consortium Agreement and considered as being processed by each partner as a controller.

9.3.1. RESPONSIBILITIES OF THE CONTROLLER AND PROCESSOR AND JOINT CONTROLLERS

In the case of processing personal data, it is crucial for the development activities of the OASEES project to consider the responsibilities of the controller and processor as outlined in the GDPR. Both parties have specific obligations set out by the GDPR.

The distinction between the data controller and data processor is based on their roles in relation to the data subject and the personal data being processed, as well as their liability for inappropriate processing of the data under the GDPR. The data controller determines the purposes and means of processing the data. This includes decisions on how the data is processed, stored, and transferred.

On the other hand, the data processor acts on behalf of the controller and within the obligations specified by the controller. The duties of the processor towards the controller must be specified in a Data Processing Agreement. A typical "controller-processor" relationship occurs when research institutions collaborate with external cloud storage providers.

In light of the current relationship between the OASEES partners and their responsibilities regarding the personal data processed in the project, it is considered that each partner would be the controller of their own data and would process the data of other OASEES team members for their own purposes (i.e., communication with the team).

If OASEES partners jointly establish the purposes and means of processing personal data in the consortium, they will become joint controllers, as defined in Article 26 of the GDPR. This relationship shall be established in a transparent manner by means of an arrangement between the joint controllers, such as a memorandum of understanding or a joint controllership agreement. This arrangement shall transparently determine the respective responsibilities for compliance with the obligations under the GDPR, particularly concerning the response to the rights of the data subjects and provision of the information about the data processing to the data subjects. The joint controllers have the duty to disclose all necessary information to ensure fair and transparent processing.

If the joint controllership relationship occurs in the consortium, appropriate arrangements shall be made by the OASEES partners.

9.4. DATA PROTECTION IMPACT ASSESSMENT (DPIA)

When engaging in activities that are considered high-risk, the controller or a processor instructed by them should conduct a Data Protection Impact Assessment. This assessment describes the processing of personal data and

evaluates its necessity and proportionality11. It also identifies potential risks and proposes methods for mitigating them. Throughout the DPIA process, the aim is to minimize any risks to the rights and freedoms of the data subjects.

The initial step for the OASEES consortium is to ascertain if a DPIA is required for the OASEES project. To this end, Article 35(1) and 35(3) of the GDPR outline the principal criteria that, when met, necessitate a DPIA:

- If the processing of personal data using new technologies is likely to pose a high risk to the rights and freedoms of individuals, the controller must conduct an assessment to evaluate the impact of the processing operations on the protection of personal data, considering the nature, scope, context, and purposes of the processing. It is possible to conduct a single assessment for a set of similar processing operations that present similar high risks.
- A DPIA is required, specifically in cases of: (a) extensive and systematic assessment of personal information of individuals that rely on automated processing, including profiling, leading to decisions that have significant legal effects or similarly affect individuals; (b) processing a large amount of special categories of data as specified in Article 9(1), or personal data relating to criminal convictions and offenses as defined in Article 10; or (c) systematically monitoring a publicly accessible area on a large scale.

Based on the current amount and type of personal data involved (limited to team members and not including sensitive/special categories of data), it seems that a DPIA may not be necessary for the project. However, ongoing monitoring will be necessary as the project develops to determine if a DPIA becomes necessary.

9.5. RECORDS OF DATA PROCESSING ACTIVITIES

As per the provisions of Article 30 of the GDPR, the controller is required to maintain records of processing activities in cases involving the processing of sensitive data or where the organization has 250 or more employees.

The records should include the following information:

a) Contact details of the controller
b) Description of processing purposes
c) Description of types of data collected/processed
d) Categories of data recipients
e) Transfers of data, including the ones to third countries
f) Storage and erasure details
g) Applied security measures

Moreover, the controller is obligated to maintain a record of the given or revoked consent[12] and also document any personal data breaches, including relevant facts, impacts, and measures taken in response to those breaches[13].

Likewise, data processors must maintain a record of their processing activities conducted on behalf of the controller. This includes information such as details of the processors, controllers, and DPOs, categories of processing

[11] Article 29, Data Protection Working Party, 'Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679. Adopted on 4 April 2017, at 6.

[12] GDPR, Art 7.

[13] GDPR, Art 33(5).

activities, transfers of data to third countries or international organizations, and descriptions of the technical and organizational measures implemented in accordance with Articles 28 and 32 of the GDPR. Although the records should be in writing, some jurisdictions may accept electronic form as an equivalent. These records must be made available to the competent supervisory authority upon request.

To comply with this requirement, each partner recognized as a controller or processor in the OASEES project must document their data processing activities in accordance with the DMP and in collaboration with their DPOs.

9.6. PRIVACY / DATA PROTECTION AND SECURITY BY DESIGN AND BY DEFAULT

In accordance with Article 25(1) of GDPR, privacy protection must be integrated from the early stages of the project and all data processing activities. Each OASEES partner must follow the GDPR safeguards when planning a processing activity that involves personal data to comply with this requirement. To protect the rights of data subjects such as workshop participants, GDPR principles must be followed, and appropriate technical and organizational measures should be applied.

Article 25(2) of the GDPR mandates data controllers to establish default settings that guarantee only the necessary personal data are processed for a particular purpose. The OASEES project will apply this type of protective setting whenever, and if personal data processing is involved. Additionally, the OASEES partners must ensure that data subjects' rights are easily exercisable, including the right to object to data processing and the right to access information about the data subject processed by the OASEES partners. The partner who collects data directly from data subjects will be accountable for responding to the exercise of their rights and sharing information about the requested actions with the rest of the consortium.

9.7. INFORMED CONSENT TO PARTICIPATE IN THE RESEARCH PROJECT

Obtaining informed consent is a fundamental ethical principle in research. It involves providing participants with an explanation of the research project, outlining their involvement and any potential risks that may be associated with it. Only after this information has been clearly communicated to the participants and they have fully comprehended it, can the project request and receive their explicit consent to participate in the research project (in accordance with Articles 4(11) and 7 of the GDPR).

The Horizon Europe guidelines on Ethics and Data Protection[14] highlight the importance of providing appropriate information to research participants outside the consortium and obtaining their free, specific, and unambiguous consent. The OASEES consortium recognizes this requirement and commits to fulfilling it. Whenever necessary, the OASEES partners will inform research participants about the research and data processing activities planned, and ensure that they provide their consent freely and without ambiguity. The forms for obtaining consent will be developed and shared in the subsequent versions of the DMP.

9.8. OPEN DATA

The OASEES project has taken into consideration the following recommendations regarding open data access:

[14] European Commission, Horizon Europe, Ethics and data protection, 5 July 2021, p. 11, available at https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-and-data-protection_he_en.pdf

The Open Data Directive

The Open Data Directive amends the Directive 2003/98/EC on the re-use of public sector information and is aimed at allowing better utilization of "the potential of public sector information" in the internal market. The Directive addresses the differences in "rules and practices in the Member States relating to the exploitation of public sector information resources.[15]" The Directive defines re-use as the use of documents held by public sector bodies or public undertakings for "commercial or non-commercial purposes other than the initial purpose" for which the documents[16] were produced[17].

The Member State transpositions of this Directive should always be taken into account in the consideration of this Directive.

In terms of the research data associated with the OASEES project to be held by Consortium partners who could be regarded as public sector bodies or public undertakings, it is important to consider the rules outlined in the Open Data Directive. While the national policies of Member States regarding the publication of research data funded by public funds are also relevant, OASEES partners may adopt an approach in alignment with the principles of the Directive and its conditions for the re-use of datasets, such as free of charge re-use of documents, making documents accessible in open, machine-readable, findable and reusable formats for the benefit of the scientific community and beyond.

OASEES aims to comply with the Open Data Directive[18] and other relevant legislation that requires public sector data to be released in open and free formats. OASEES recognizes the importance of a research project adhering to the directive on open data by providing open and trusted services for data silos and facilitating the use and reuse of information (including public and protected information, as well as metadata) using common rules within an expanding European market for government-held and federated data.

The Open Data Directive sets out several key principles that are relevant to OASEES, including:

- Releasing non-personal data in open formats and standards.
- Making data available in real-time and via APIs where possible.
- Charging rules: free reuse is now a principle, especially in terms of commercialization and exploitation.
- Reusing publicly funded research data, which is a focus area for OASEES use cases.
- Discouraging exclusive arrangements and data lock-in.
- Allowing the reuse of data held by public undertakings, such as public utilities and transport providers.

OASEES also monitors the efforts related to the concept of high-value datasets (in areas such as geospatial, earth observation and environment, meteorological, statistics, companies and company ownership, mobility, etc.). Defined as documents, the re-use of high-value datasets is associated with important benefits for the society and economy. They are subject to a separate set of rules ensuring their availability free of charge, in machine readable formats. OASEES is positioned in between these datasets and their use (through specific Application Programming

[15] Open Data Directive, Recital 15.

[16] Documents, according to Article 2(6) of the Directive, can be any content regardless of their medium (i.e. paper or electronic form) or any part of their content.

[17] Open Data Directive, Article 2(11).

[18] https://digital-strategy.ec.europa.eu/en/policies/legislation-open-data

Interfaces (APIs) and will exploit the possibility of using its SSI framework to strengthen their thematic scope of value.

Commission's Recommendation on access to and preservation of scientific information of (C(2012) 4890 final)

The EU Member States are advised to appoint a National Point of Reference, whose responsibility is to report on the implementation of open access within the respective Member State. As part of a package of measures aimed at enhancing access to scientific information produced in Europe and aligning them with the Commission's policy for Horizon 2020[19], the 2012 Recommendation on access to and preservation of scientific information (2012/417/EU) was introduced.

Recommendation C(2018)2375, adopted on April 25th 2018

The recommendation explicitly takes into account various developments in areas such as research data management, including the concept of FAIR data (i.e., data that is Findable, Accessible, Interoperable and Reusable), Text and Data Mining (TDM), technical standards enabling re-use and incentive schemes. It considers the ongoing developments at the EU level of the European Open Science Cloud and recognizes the growing role of data analytics in research. Moreover, it clearly distinguishes between two separate points: the issue of reward systems for researchers to share data and commit to other open science practices on the one hand, and the skills and competences of researchers and staff from research institutions on the other hand.

9.9. LINKING OASEES WITH OTHER OPEN RESEARCH INFRASTRUCTURES, EUROPEAN RESEARCH INFRASTRUCTURES AND INTERNATIONAL DATA SPACES, GAIA-X AND EUROPEAN OPEN SCIENCE CLOUD

Efforts will be made to ensure seamless interoperability with other relevant European infrastructures. Specifically, the project aims to explore interoperability with International Data Spaces, Gaia-X, and European Open Science Cloud (EOSC) data spaces.

The EOSC is a federated, open environment that allows the scientific community to access, store, manage, analyze, and reuse digital research outputs (such as publications, data, metadata, and software) for research, innovation, and educational purposes[20]. It offers cloud-based services for open science by integrating and consolidating e-infrastructure platforms, federating existing European research infrastructures, and scientific clouds.

Funded through the Horizon 2020 initiative and officially launched in November 2018, the EOSC provides access to services via their EOSC Portal[21]. Special attention will be given to ensuring interoperability with the EOSC to fully leverage the benefits of this innovative platform.

[19] http://ec.europa.eu/research/openscience/pdf/openaccess/background_note_open_access.pdf#view=fit&pagemode=none

[20] Eudat, Liber, OpenAIRE, Egi, Geant, European Open Science Cloud for research [Internet], Position Paper, 2015 Oct. Available at http://libereurope.eu/wp-content/uploads/2015/11/OSC_Position_Paper-final-30.10.15.pdf

[21] https://www.data-infrastructure.eu/GAIAX/Navigation/EN/Home/home.html

OASEES takes into consideration the de-centralized approach of Gaia-X, which comprises of multiple individual platforms following a common standard known as the "Gaia-X standard", while exploring technical and governance requirements.

The International Data Spaces (IDS) initiative focuses on cross-sectoral data sovereignty and data interoperability, with an emphasis on a Reference Architecture Model that implements open standards and contributes to global standards. The IDS also establishes the terms and conditions for the data economy, and it ensures efficient adoption by, for example, publishing open source software codes. The IDS Association (IDSA) and its member organizations participate in more than 20 European research projects, mainly funded under the Horizon 2020 (and Horizon Europe) program. The technologies continue to develop, mature, and are deployed in several productive environments. The IDSA is also part of GAIA-X, which aims to establish a Federated Data Infrastructure for European ecosystems[22]. Moreover, OASEES will ensure interoperability with other relevant European infrastructures, including the European Open Science Cloud (EOSC) data spaces, as it explores ways to expand the benefits of digital research outputs for research, innovation, and educational purposes.

The OASEES consortium established a taskforce group to investigate all three initiatives and explore potential collaborations with them. These meetings occur within the context of WP2, which aims to provide educational resources to less experienced partners and generate ideas for aligning OASEES with the initiatives' standards.

9.10. EU CLASSIFIED INFORMATION

The security of classified information within the EU is governed by various acts. Among these is the Council Decision No 2013/444[23], which establishes the fundamental principles and minimum criteria for safeguarding EU Classified Information (EUCI).

The Decision defines four levels of the classification[24]:

Table 1: Levels of the Classification

The EU classification levels:	What an unauthorised disclosure of the EU information or material could do
TRES SECRET UE/EU TOP SECRET	Cause exceptionally grave prejudice to the essential interests of the Union or of one or more of the Member States

[22] IDSA, Implementing the European Strategy on Data Role of the International Data Spaces (IDS), Position Paper, Version 1.0, April 2022, available at https://internationaldataspaces.org/wp-content/uploads/IDSA-Position-Paper-Implementing-European-Data-Strategy-Role-of-IDS1.pdf

[23] Council Decision 2013/444 of 23 September 2013 on the security rules for protecting EU classified information with further amendments

[24] Ibid., Article 2(1).

SECRET UE/EU SECRET	Seriously harm the essential interests of the Union or of one or more of the Member States
CONFIDENTIEL UE / EU CONFIDENTIAL	Harm the essential interests of the EU or of one or more of the Member states
RESTREINT UE / EU RESTRICTED	Be disadvantageous to the interests of the EU or of one or more of the Members states

The Council Decision No 2013/444 is applicable to the Council and the GSC, and it requires the Member States to comply with their national laws and regulations, ensuring an equivalent level of protection for EU Classified Information (EUCI). However, given the current situation and the General Assembly, the OASEES project does not intend to utilize any data or information classified as EUCI at any stage of the project. Even the most sensitive use cases will not involve the use of such information.

10. CONCLUSION

This deliverable provides an initial overview of the Data Management Plan (DMP) for the OASEES project, which outlines the intended data sets that will be collected and generated, as well as their potential to become FAIR data. The document includes detailed information about the content of the data sets and their intended use. To ensure the data's security, the project will use secured data storages and implement technical and organizational measures for secure data processing.

11. REFERENCES

1. https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en
2. https://enspire.science/wp-content/uploads/2021/09/Horizon-Europe-Data-Management-Plan-Template.pdf
3. Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18, https://www.nature.com/articles/sdata201618
4. https://webgate.ec.europa.eu/funding-tenders-opportunities/display/OM/Online+Manual
5. Regulation (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).
6. GDPR, Art 3.
7. Article 29 Data Protection Working Party, Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679 (WP 251 2017) at 11.
8. GDPR, Recital 39.
9. GDPR, Art. 5(1)(f).
10. GDPR, Recital 78.
11. Article 29, Data Protection Working Party, 'Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679. Adopted on 4 April 2017, at 6.
12. GDPR, Art 7.
13. GDPR, Art 33(5).
14. European Commission, Horizon Europe, Ethics and data protection, 5 July 2021, p. 11, available at https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-and-data-protection_he_en.pdf
15. Open Data Directive, Recital 15.
16. Documents, according to Article 2(6) of the Directive, can be any content regardless of their medium (i.e. paper or electronic form) or any part of their content.
17. Open Data Directive, Article 2(11).
18. https://digital-strategy.ec.europa.eu/en/policies/legislation-open-data
19. http://ec.europa.eu/research/openscience/pdf/openaccess/background_note_open_ access.pdf#view=fit&pagemode=none
20. Eudat, Liber, OpenAIRE, Egi, Geant, European Open Science Cloud for research [Internet], Position Paper, 2015 Oct. Available at http://libereurope.eu/wp-content/uploads/2015/11/OSC_Position_Paper-final-30.10.15.pdf
21. https://www.data-infrastructure.eu/GAIAX/Navigation/EN/Home/home.html
22. IDSA, Implementing the European Strategy on Data Role of the International Data Spaces (IDS), Position Paper, Version 1.0, April 2022, available at https://internationaldataspaces.org/wp-content/uploads/IDSA-Position-Paper-Implementing-European-Data-Strategy-Role-of-IDS1.pdf
23. Council Decision 2013/444 of 23 September 2013 on the security rules for protecting EU classified information with further amendments
24. Ibid., Article 2(1).

Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.