NOUS

Empowering
Europe's Data
Future

D1.3 DATA MANAGEMENT PLAN

30/05/2024

Funded by
the European Union

PARTNERS

AIR INSTITUTE

AETHON ENGINEERING

VNiVERSiDAD D SALAMANCA

DEMOKRITOS

ICOOR

1859

UNIMORE UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

TIM GRUPPO TIM

netcompany intrasoft

itml innovation applied

UNIVERSITÀ DI PISA

UNPARALLEL

KU LEUVEN CiTiP CENTRE FOR IT & IP LAW

Hewlett Packard Enterprise

ECLIPSE FOUNDATION

F6S

ΔEH

ARCTUR

AEGIS IT RESEARCH

CS a Sopra Steria company

INESCTEC

Grant Agreement No.: 101135927
Call: HORIZON-CL4-2023-DATA-01
Topic: HORIZON-CL4-2023-DATA-01-02
Type of action: HORIZON - RIA

D1.3 DATA MANAGEMENT PLAN

Work package	WP 1
Task	T1.4
Due date	31/05/2024
Submission date	31/05/2024
Deliverable lead	AETHON
Version	V.1
Authors	Chara Nikolaou (AETHON), Anna Kontini (AETHON)
Reviewers	Costas Kalogiros (AEGIS), Marco Mamei (UNIMORE)
Abstract	The NOUS project represents a significant advancement in Europe's digital infrastructure, centered on implementing the FAIR (Findable, Accessible, Interoperable, and Reusable) principles. Through integrating advanced technologies like quantum computing and blockchain, NOUS enhances high-performance computing and secure data processes, thereby fostering European digital sovereignty. The project's comprehensive approach includes detailed analyses of data types, security needs, and ethical considerations, culminating in a robust Data Management Plan (DMP) that exceeds Horizon Europe's standards. By engaging stakeholders and conducting extensive surveys, NOUS ensures its DMP is practical, forward-looking, and adaptable, setting new benchmarks for data management in research. Future efforts should leverage these insights to further refine data practices and explore new technological frontiers, maintaining a commitment to open science and broad data sharing to drive innovation and enhance Europe's digital economy.
Keywords	Data Management, FAIR Principles

Document Revision History

Version	Date	Description of change	List of contributor(s)
V1	12/04/2024	Table of Content	AETHON
V2	15/05/2024	First Draft	AETHON
V3	27/05/2024	Deliverable Review	UNIMORE, AEGIS
V4	29/05/2024	Second Draft	AETHON
V0	30/05/2024	Final Document	AETHON

DISCLAIMER

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

COPYRIGHT NOTICE

The NOUS Consortium is the following:

Participant number	Participant organisation name	Short name	Country
1	FUNDACION INSTITUTO INTERNATIONAL DE INVESTIGACION EN INTELIGENCIA ARTIFICIAL Y CIENCIAS DE LA COMPUTACION	AIR	ES
2	AETHON ENGINEERING SINGLE MEMBER PC	AETHON	EL
3	UNIVESIDAD DE SALAMANCA	USAL	ES
4	NATIONAL CENTER FOR SCIENTIFIC RESEARCH “DEMOKRITOS”	NCSR “D”	EL
5	CONSORZIO INTERUNIVERSITARIO PER L’ OPTIMIZZAZIONE E LA RICERCA OPERATIVA	ICOOR	IT
6	POLITECNICO DI TORINO	POLITO	IT
7	UNIVERSITA DEGLI STUDI DI MODENA E REGGIO EMILIA	UNIMORE	IT
8	TELECOM ITALIA SPA	TIM	IT
9	NETCOMPANY – INTRASOFT SA	NET-INTRA	LU
10	IOTAM INTERNET OF THINGS APPLICATIONS AND MULTI LAYER DEVELOPMENT LTD	ITML CY	CY
11	UNIVERSITA DI PISA	UNIPI	IT
12	UNPARALLEL INNOVATION LDA	UNPARALLEL	PT
13	KATHOLIEKE UNIVERSITEIT LEUVEN	KU Leuven	BE
14	HEWLETT – PACKARD HELLAS ETAIREIA PERIORISMENIS EFTHINIS	HPE Hellas	EL
15	ECLIPSE FOUNDATION EUROPE GMBH	ECL	DE
16	F65 NETWORK IRELAND LIMITED	F65 IE	IE
17	DIMOSIA EPICHEIRISI ILEKTRISMOU ANONYMI ETAIREIA	PPC	EL

18	ARCTUR RACUNALNISKI INZENIRING DOO	ARCTUR d.o.o	SI
19	AEGIS IT RESEARCH GMBH	AEGIS	DE
20	CS GROUP – FRANCE	CS GROUP-FRANCE	FR
21	INESC TEC- INSTITUTO DE ENGENHARIADE SISTEMAS E COMPUTADORES, TECNOLOGIA E CIENCIA	INESC TEC	PT

EXECUTIVE SUMMARY

NOUS marks a pivotal endeavour aimed at enhancing Europe's digital infrastructure through the NOUS project. Central to its mission is the rigorous implementation of data management practices aligned with the FAIR (Findable, Accessible, Interoperable, and Reusable) principles. NOUS is distinguished by its innovative approach to integrating advanced computational technologies within the European data infrastructure, providing a blueprint for leveraging these technologies in fostering European digital sovereignty. It focuses on the development and integration of cutting-edge technologies such as quantum computing and blockchain, promoting the seamless interplay between high-performance computing and secure data processes. The outcomes are expected to have a profound impact on future research and development efforts, driving innovation, enhancing data security, and promoting sustainable and ethical data use across Europe.

Employing a comprehensive methodological approach, NOUS has undertaken extensive analyses of data types and security needs, detailed assessments of interoperability frameworks, and thorough evaluations of ethical implications. By emphasizing the FAIR data principles, the project commits to maintaining high standards of data accessibility and reusability, enhanced by strong data security measures and adherence to ethical standards.

A significant achievement of NOUS is the formulation of a robust DMP, setting new standards for data management in line with the objectives of Horizon Europe. The project underscores the critical role of the FAIR principles in modern data management, devises strategies for maximizing data utility, and navigates complex legal and ethical challenges associated with data governance.

The approach followed in formulating the NOUS DMP is characterized by thorough planning, stakeholder engagement, and iterative refinement. An extensive survey was conducted among NOUS consortium members to gather insights into their data management practices, needs, and challenges. The survey covered various aspects, such as data reuse, types and formats of data generated, expected data volumes, data origin and provenance, and the utility of the data for external stakeholders. This input was crucial in shaping a DMP that is both comprehensive and tailored to the specific requirements of the project. Furthermore, regular consultations with stakeholders, including researchers, data managers, and legal experts, ensured that the DMP aligns with the needs of the project and complies with relevant regulations. Lastly, the DMP is a living document, subject to regular review and updates. This iterative process allows for continuous improvement based on feedback from consortium members and changes in the regulatory landscape. To ensure effective implementation of the DMP, training sessions and capacity-building workshops were organized for all consortium members. This equipped them with the necessary skills and knowledge to manage data efficiently and in compliance with the DMP guidelines.

By following the aforementioned approach, NOUS project has developed a DMP that not only meets but exceeds the requirements of Horizon Europe, setting a new benchmark for data management in research projects. The use of a survey and active stakeholder engagement were pivotal in creating a plan that is both practical and forward-looking, ensuring that the data generated by the project is handled with the highest standards of quality and integrity.

Future initiatives should build on the insights gained from the NOUS project, continuously refining data management practices and exploring new technological advancements. A steadfast commitment to open science and widespread data sharing remains crucial in catalysing innovation and ensuring the broad applicability of the project's achievements.

This executive summary provides a clear and detailed overview of the NOUS project, designed to be accessible and informative for a diverse audience. It captures the project's strategic vision, key

accomplishments, and its transformative potential in strengthening the foundations of Europe's digital economy, emphasizing its pivotal role in advancing European digital sovereignty.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

DMP	Data Management Plan
FAIR	Findable, Accessible, Interoperable, and Reusable
AI	Artificial Intelligence
HPC	High-Performance Computing
R&D	Research and Development
PID	Persistent Identifier
GDPR	General Data Protection Regulation
IPR	Intellectual Property Rights
CA	Consortium Agreement
GA	Grant Agreement
LLM	Large Language Model
DOI	Digital Object Identifier

1 INTRODUCTION

The deliverable 1.3 provides a comprehensive analysis of a Data Management Plan (DMP). It includes guidelines for detailing the handling of research data during and after a project, emphasizing the principles of making data findable, accessible, interoperable, and reusable (FAIR). Sections cover data summary, adherence to FAIR principles, data storage and preservation, ethical and legal compliance, and allocation of resources for data management. It serves as a practical guide for researchers to plan and execute effective data management strategies in alignment with Horizon Europe's mandates.

The methodology for the Data Management Plan (DMP) in the NOUS project integrates best practices and standards required by Horizon Europe, focusing on the principles of FAIR data. Specifically, the key methodological steps are the following:

A. The initial phase involves gathering functional and non-functional requirements for data management from all project stakeholders. This includes understanding the data needs and compliance requirements of the consortium members and aligning them with Horizon Europe's guidelines.

B. Data Collection and Documentation: Establishing a standardized approach for data collection and documentation, ensuring that all data are properly recorded, labeled, and stored in predefined formats. This step includes the creation of metadata standards to support the FAIR principles.

C. Data Storage and Preservation Planning: Identifying and implementing robust storage solutions, both physical and cloud-based, to ensure data security and integrity. This includes setting up regular backup schedules and planning for long-term data preservation.

D. Compliance and Ethical Considerations: Ensuring that all data management practices comply with ethical guidelines and legal requirements, including GDPR. This involves obtaining informed consent from participants and implementing necessary data protection measures.

E. Training and Capacity Building: Providing training sessions for researchers and consortium members on effective data management practices and the use of data management tools and platforms.

F. Continuous Monitoring and Improvement: Regularly reviewing and updating the DMP to incorporate feedback and address any emerging challenges or opportunities. This ensures that the data management practices remain effective and aligned with project goals and Horizon Europe's mandates.

The DMP is closely related to several other deliverables within the NOUS project, such as D1.1, D2.1, D3.1, D4.2 and D5.2, ensuring that data management practices are seamlessly integrated into the broader project framework, enhancing overall project's efficiency and impact.

2 NOUS DMP SURVEY

The survey conducted within the NOUS consortium reveals insightful dynamics and strategies surrounding data management that are crucial for the project's success. These findings highlight a concerted effort among participants to manage data responsibly and securely, while also leveraging it for substantial research and development advancements.

2.1 OVERVIEW OF KEY FINDINGS

- The responses among NOUS participants indicate a deliberate approach to data stewardship and oversight, with specific responsibilities systematically allocated within the consortium. This structured allocation ensures that data management processes are both efficient and compliant with established standards.
- Ethical considerations are universally recognized among participants, with a strong emphasis on adhering to ethical guidelines and legal requirements. This includes the procurement of informed consent for data utilization and the observance of privacy regulations, underscoring the consortium's commitment to ethical research practices.
- Participants employ a dual strategy of utilizing both pre-existing data and generating new data, with a preference for accessible and interoperable data formats. Despite the unpredictability associated with data volume estimations, there is a significant focus on collaborative data collection, enhancing the utility and broader implications of the data gathered.
- While there is a general aspiration towards fulfilling the FAIR (Findable, Accessible, Interoperable, Reusable) data principles, participants note some obstacles, particularly in aspects such as data identification and metadata enrichment. These challenges highlight areas for improvement in ensuring complete compliance with the FAIR guidelines. Task T1.4 will facilitate this process by providing guidance, templates, and continuous support to consortium members.
- A proactive stance on managing and disseminating digital research outputs is evident, guided by internal regulations, GDPR compliance, and selective data sharing practices. This approach balances openness with the need to protect sensitive information, ensuring that digital outputs are shared responsibly and securely.
- The consortium members adopt a multi-layered approach to data security, integrating both physical and cloud-based storage solutions. This includes employing daily data backups and leveraging secure software platforms like SharePoint and Microsoft 365 to safeguard sensitive data and ensure its integrity throughout the project's lifecycle.

The survey findings from the D1.3 DMP within the NOUS consortium paint a picture of a well-considered and holistic approach to data management. These insights underscore the importance of ethical considerations, robust data security measures, and the strategic utilization of data to significantly contribute to the project's research and development objectives. Despite some challenges in fully aligning with the FAIR principles, the consortium's dedication to responsible and effective data practices promises to amplify the research outputs' value and impact, enhancing the project's overall success.

3 DATA SUMMARY

The survey conducted as part of the Data Management Plan preparation among NOUS consortium members has yielded insightful findings regarding their plans for handling, generating, and utilizing data within their respective projects. Below is an analysis of each aspect based on the responses provided.

3.1 REUSING EXISTING DATA

NOUS participants displayed a proactive stance towards leveraging existing datasets from previous projects, with some participants like AETHON engineering re-using data for metrics and performance benchmarks to enhance system design. This strategy not only maximizes the utility of prior data collection and analytical efforts but also has the potential to accelerate the initiation of new projects. The emphasis on integrating real-time and historical data demonstrates a dedication to continual improvement and a focus on creating robust, user-centric technological solutions.

3.2 DATA GENERATION – TYPES AND FORMATS

The preference for widely used file formats (e.g., JSON, XML, MP4, JPEG, DOCX, and PDF) among participants underscores the importance placed on data interoperability and accessibility. This also includes the use of standardized data schemas, which further enhance the ability to integrate and share data across different platforms and systems. For instance, the responses from FUNDACION INSTITUTO INTERNATIONAL DE INVESTIGACION EN INTELIGENCIA ARTIFICIAL Y CIENCIAS DE LA COMPUTACION (AIR) indicate a variety of multimedia formats are utilized to manage project documentation and meeting records. Opting for these common formats ensures that the data remains accessible to a broad audience, facilitating easier collaboration both within the consortium and with external entities such as academic groups and other research projects.

3.3 DATA VOLUME EXPECTATIONS

Responses varied significantly when it came to the anticipated size of the data, with estimates ranging from several megabytes to multiple terabytes, as detailed by POLITECNICO DI TORINO (POLITO) for their various use cases. This variability reflects the diverse nature of research and development endeavors within the NOUS project, where the extent of data generation can fluctuate based on the project scope, stage, and the exploratory nature of certain research activities.

3.4 DATA ORIGIN AND PROVENANCE

The sources of data, as reported by entities like UNIVERSIDAD DE SALAMANCA (USAL) and AETHON, include both internal project records and pilot partners, indicating a comprehensive and multi-dimensional approach to data collection. Such a strategy promises to enrich the dataset with a variety of perspectives and insights, bolstering the overall quality and utility of the information gathered. It also exemplifies the project's collaborative spirit, with contributions amalgamated from an array of stakeholders including project partners and external data sources.

3.5 EXTERNAL UTILITY OF DATA

The anticipation that the data might prove beneficial for external parties such as other research groups, security analysts, cloud service providers, and educational institutions signals a forward-looking approach to data management. Acknowledging the potential for data to extend its value beyond the confines of the specific project highlights the consortium's view of data as a crucial, enduring asset, poised to drive continuous learning, innovation, and value generation across various sectors.

Overall, the survey findings from the DMP preparation process illuminate the NOUS consortium members' thoughtful and strategic orientation towards data management. The insights gleaned underscore a collective commitment to leveraging existing data resources, ensuring data accessibility, accommodating the unpredictability of data generation, and recognizing data's extended utility. This strategic framework reflects a concerted effort towards efficient, collaborative, and progressive data management methodologies, setting a benchmark for future projects in similar high-tech domains. Detailed responses can be found in Appendix A.

4 FAIR DATA

FAIR data principles, standing for Findable, Accessible, Interoperable, and Reusable, were established to improve the infrastructure supporting the reuse of scholarly data across various disciplines. These principles are designed to enhance the ability of machines to automatically find and use data, as well as to support its reuse by individuals, thus facilitating better data management and stewardship. The FAIR data principles were first formally published and promoted in a 2016 article by Mark D. Wilkinson and colleagues in the journal Scientific Data (Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3. https://doi.org/10.1038/sdata.2016.18).

The promotion of the FAIR data principles began as a response to the urgent need for a more efficient and effective way to handle the vast amounts of data generated by research activities. With the exponential growth of digital data and the increasing complexity of scientific research, it became evident that traditional methods of data management were inadequate for ensuring the accessibility, reusability, and sustainability of research data. The FAIR data principles were developed through a collaborative effort involving a diverse set of stakeholders, including academia, industry, funding agencies, and scholarly publishers.

Core Principles of FAIR Data

Findable: Data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services.

Accessible: Once found, data should be accessible with clearly defined access conditions, whether open or under specific restrictions.

Interoperable: Data should be compatible with other datasets, analysis tools, and workflows for both inter- and intra-disciplinary use. This requires the use of standard, open formats and vocabularies.

Reusable: Data should be reusable and re-purpose able for future research and applications. This requires clear data and metadata documentation, including provenance information, to ensure that the data can be used by others.

The open science movement advocates for the democratization of scientific knowledge by ensuring that research data is made widely available and usable. This initiative is grounded in the belief that such openness will catalyse scientific discovery and innovation. The movement emphasizes the need for robust data management and stewardship practices, which are essential for enhancing the efficiency of research processes. By reducing redundancy in data collection and improving the quality and integrity of research data, the scientific community can achieve more reliable and verifiable outcomes.

Furthermore, open science fosters collaboration across disciplines and sectors by facilitating the sharing of data in a manner that is comprehensible across various fields. This interdisciplinary approach enables stakeholders, including researchers, policymakers, and the public, to engage in collaborative efforts to address complex challenges. Such collaboration is pivotal for the integration of diverse perspectives and expertise in the pursuit of innovative solutions.

The potential of big data and artificial intelligence (AI) in scientific research is another focal point of the open science movement. By making data more accessible and interoperable, researchers can leverage advanced analytical and interpretive techniques. This enhances the capacity for nuanced understanding and exploration of data, paving the way for groundbreaking discoveries. The strategic utilization of big data and AI underscores the transformative impact of open science on the research landscape, offering new horizons for exploration and understanding in various scientific domains.

4.1 CONTRIBUTION OF FAIR DATA TO NOUS PROJECT

The transition through various technological eras has consistently introduced groundbreaking advancements in digital infrastructure and cybersecurity, with the NOUS project marking a significant step towards the realization of European digital sovereignty. This evolution is underscored by the transformation from traditional data handling to innovative management processes that cater to enhanced security and operational efficiency, facilitated by state-of-the-art technologies such as quantum computing and blockchain. Within this context, the principles of FAIR data—Findable, Accessible, Interoperable, and Reusable—become pivotal, enhancing data management and stewardship to support robust, secure, and efficient digital systems.

The FAIR Guiding Principles, aimed at improving the infrastructure for data usability across scientific and technological disciplines, emphasize the necessity for machine autonomy in data discovery and utilization, alongside facilitating its reuse by humans. This framework is instrumental in the NOUS project, where the synergy between advanced computational solutions and data-driven decision-making is paramount. According to Wilkinson et al. (2016), the ability of machines to autonomously find and use data, coupled with their capability to support human reuse, underpins the transformative potential of technologies deployed in the NOUS project.

Furthermore, the integration of diverse data sources within the project—spanning quantum data outputs to blockchain transaction logs—benefits substantially from adherence to FAIR principles. By breaking down internal data silos and fostering efficient integration with advanced computational models, FAIR principles enable the development of innovative applications and robust cybersecurity measures characteristic of the NOUS project's goals. Vlijmen et al. (2020) elucidate how this integration is essential for the iterative innovation cycle in enhancing European digital infrastructure.

Emphasizing the security aspect within the digital framework of the NOUS project, the enhancement of data protection processes and the empowerment of data governance in compliance with GDPR are central themes. FAIR data principles facilitate this by ensuring data interoperability and reusability across systems, which in turn, enhances security protocols and decision-making processes. Akundi et al. (2022) discuss how the application of FAIR principles supports the development of more secure and efficient digital infrastructures, thereby optimizing the integrity and efficacy of data operations.

In the research and development (R&D) processes of the NOUS project, the application of FAIR principles is touted to significantly boost the efficiency and effectiveness of data management. By enabling seamless access to and learning from data by new analytical tools, such as AI and advanced data encryption methods, a digital transformation aligns with the core objectives of enhancing European digital sovereignty. Wise et al. (2019) argue that this digital transformation is critical for leveraging the vast potentials of data-driven insights in securing and optimizing digital services.

In conclusion, the FAIR data principles play an indispensable role in realizing the vision of the NOUS project, enhancing machine learning capabilities, facilitating cross-technology data integration, and promoting efficient and secure digital services. These principles are foundational in achieving the aspirations of the NOUS project, which include robust data protection, optimized data utilization, and the harmonious integration of advanced technologies with comprehensive data management strategies.

4.2 MAKING DATA FINDABLE, INCLUDING PROVISIONS FOR METADATA

The NOUS project articulates a comprehensive strategy that underscores the pivotal role of FAIR principles in sculpting a forward-looking data management paradigm. Central to this endeavor is the establishment of robust data management protocols which serve as a testament to the project's dedication to enhancing the interoperability and usability of data within the context of advancing European digital sovereignty. Through a structured compilation of deliverables and work packages, the project delineates a clear trajectory towards ensuring that data collected, processed, and analyzed adheres rigorously to the FAIR guidelines, thus epitomizing best practices in data stewardship.

The focus on these aspects is evident in the following key deliverable:

Deliverable D3.2 – Data Processing and Analysis: This deliverable undertakes a comprehensive analysis of data collected across various work packages, with a specific focus on data generated from quantum computing experiments and blockchain transaction logs. The goal is to extract insights that will inform the development of actionable recommendations for a broad spectrum of stakeholders, including academic researchers, industry partners, and policy makers. It encapsulates a dual focus on analyzing the data for insight generation and critically evaluating the data collection process itself, thereby ensuring that the data remains findable and accessible for subsequent analysis and interpretation. This deliverable also involves updating the data management protocols to ensure ongoing compliance with FAIR principles, particularly emphasizing the standardization of metadata to enhance data discoverability and usability.

Below presented the key aspects related to NOUS metadata provisions:

Metadata Standards and Descriptions: In the context of this DMP, the NOUS project emphasizes the importance of using standardized metadata formats and descriptions. This ensures that the data can be easily understood and used by others, including future researchers and stakeholders interested in the project's findings. The metadata includes details on the dataset reference and name, the origin, nature, and scale of the data, and any relevant information that would facilitate its reuse.

Task 3.1 – Data Processing and Analysis: This task involves analyzing the data collected in WP2 and includes an evaluation and update of the data management protocols based on the data analysis. As part of this process, the project will ensure that the data is organized in a standardized format with accompanying metadata. This metadata will detail the methodologies used for data collection and analysis, making the data not only interoperable but also reusable for similar future analyses or projects.

Metadata for Quantum Computing and Blockchain Data: For data derived from quantum computing and blockchain technologies, the metadata schema includes specific elements that describe the computational environment, the algorithms used, and the parameters of the blockchain network. This attention to detail ensures that the metadata is robust enough to allow other researchers and technologists to understand and replicate findings or extend the data's utility in future projects.

Each dataset will be assigned a persistent identifier, such as a Digital Object Identifier (DOI), which provides a permanent link to the data. This ensures that even if the dataset's location changes, the DOI will always direct users to the correct resource. Metadata and datasets will be uploaded to reputable public repositories, such as ZENODO. These repositories are indexed by major search engines, further enhancing the visibility and accessibility of the datasets.

Integration with European Data Infrastructure: The metadata strategy also aligns with the broader objectives of integrating with the European data infrastructure. By ensuring metadata is compliant with European standards and interoperable across different platforms and systems, the NOUS project supports seamless data sharing and collaboration across borders, enhancing the impact and applicability of the research conducted.

Making data findable through well-defined metadata and adherence to FAIR principles is a cornerstone of the NOUS project's data management strategy. This approach not only fosters transparency and accessibility but also significantly contributes to the project's overarching goal of enhancing European digital sovereignty and security. The structured and strategic application of metadata ensures that all data generated by the project is ready to be utilized in advancing the frontiers of digital technology and cybersecurity.

4.3 NOUS DATA ACCESSIBILITY STRATEGY

The NOUS project outlines a comprehensive strategy to elevate data accessibility, closely aligning with the FAIR principles. Central to this strategy is the development of a transparent data access policy. This policy clearly delineates how data can be accessed, detailing specific conditions and any potential restrictions or legal mandates that might impact data availability. By providing such guidance, the project aims to ensure that all stakeholders have a clear understanding of how to engage with and utilize the available data resources.

Transparent Data Access Policy: The transparent data access policy of the NOUS project specifies the terms under which data can be accessed, including any usage restrictions imposed by data sensitivity or security requirements. This policy is designed to be straightforward and easily understandable, ensuring that both technical and non-technical stakeholders are well-informed of their rights and responsibilities regarding data use.

Standardized Protocols and Persistent Identifiers (PIDs): To facilitate seamless and secure data access, the strategy emphasizes the adoption of standardized protocols that are open-source and designed for universal application, ensuring data security and integrity. The assignment of Persistent Identifiers (PIDs) to datasets guarantees their long-term discoverability and accessibility, addressing potential challenges related to data location or ownership changes over time.

Metadata Accessibility: Ensuring metadata accessibility is a critical aspect of the NOUS strategy. Accessible metadata allows users to understand data contexts and access conditions, even if the data itself is under certain restrictions. This approach is complemented by the implementation of standardized authentication and authorization procedures, which balance secure access with the need to avoid undue restrictions.

Comprehensive Documentation and Long-term Preservation: The provision of comprehensive documentation is identified as essential for empowering users to effectively access and utilize the data. This documentation includes detailed information on datasets, access procedures, and data formats. Additionally, the strategy commits to supporting data accessibility over time, recognizing the importance of strategies for long-term preservation and data migration to accommodate technological evolution.

Machine-Readability and Interoperability: In recognition of the increasing reliance on automated tools and systems, the strategy advocates for making data machine-readable wherever possible. This facilitates automated data processing and enhances the efficiency of data utilization. Additionally, the use of standard vocabularies and data formats is encouraged to improve interoperability with other datasets and systems.

Adherence to FAIR Principles: Lastly, the strategy considers the importance of demonstrating adherence to FAIR principles through accessibility audits and certifications. This external validation underscores the NOUS project's commitment to data accessibility and serves as a benchmark for best practices in data management.

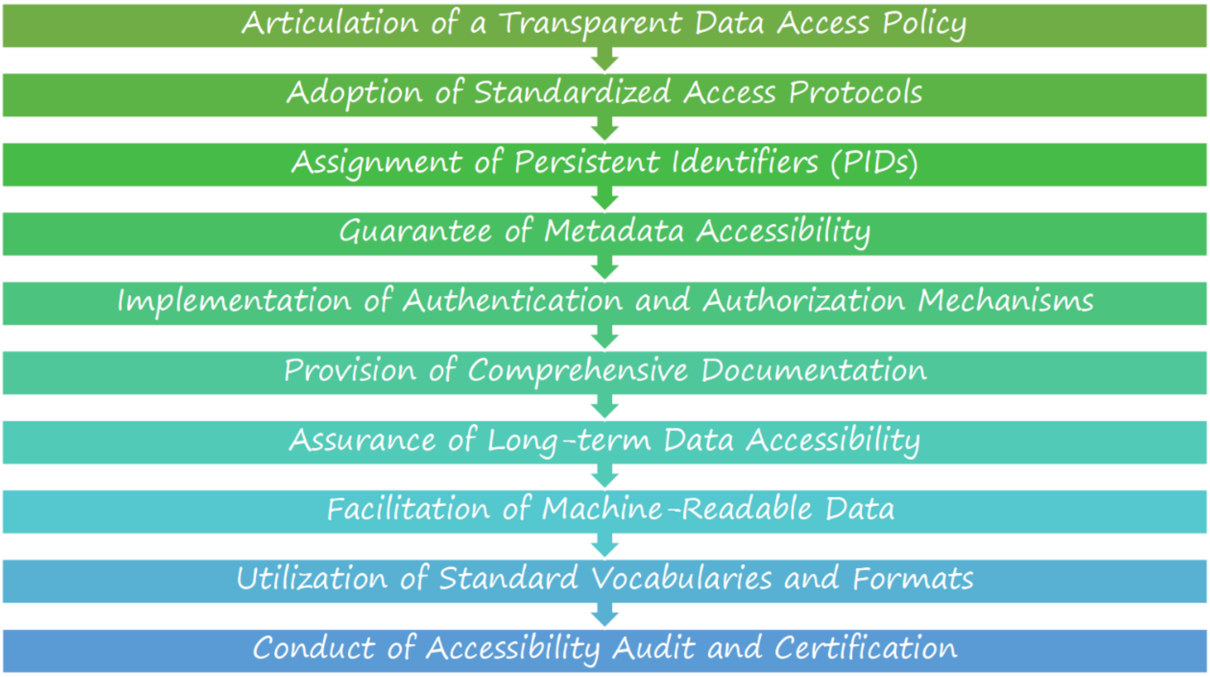
Articulation of a Transparent Data Access Policy
Adoption of Standardized Access Protocols
Assignment of Persistent Identifiers (PIDs)
Guarantee of Metadata Accessibility
Implementation of Authentication and Authorization Mechanisms
Provision of Comprehensive Documentation
Assurance of Long-term Data Accessibility
Facilitation of Machine-Readable Data
Utilization of Standard Vocabularies and Formats
Conduct of Accessibility Audit and Certification

FIGURE 1: NOUS PROJECT'S DATA ACCESSIBILITY STRATEGY

4.4 MAKING DATA ACCESSIBLE

The NOUS project provides a thorough examination of cutting-edge technologies and methodologies pivotal to enhancing data accessibility within the realms of advanced European digital infrastructures.

Implementation of Accessibility Tools:

- NOUS Web Platform: A key component for making data accessible is the development of the NOUS web platform. This platform allows users to input, manage, and access data related to their specific cybersecurity and computational research, facilitating the sharing and dissemination of findings and resources.

- NOUS Collaboration Platform: The project also creates an online collaboration platform, which serves as a central repository for project outputs, data, and other relevant information. This workspace is designed to be easily accessible to project participants and the wider community, ensuring that data and findings are readily available and can be utilized in real-time.

By rendering data more amenable to both human and machine access, these methodologies align with the overarching objectives of the FAIR principles, fostering the unencumbered sharing and reuse of research data across the digital and scientific communities. The foundation of these strategies on established security protocols and modern web standards ensures straightforward implementation, which could catalyze its adoption across a broad spectrum of data repositories within the European digital ecosystem.

Furthermore, our focal point of this DMP is its alignment with the Accessibility dimension of the FAIR (Findable, Accessible, Interoperable, and Reusable) data principles, through an in-depth analysis of two core strategies adapted to the specific needs of the NOUS project: "Advanced Access Control" and the "NOUS Data Access Gateway.”

1. Advanced Access Control Advanced Access Control emerges as a strategy to improve the discoverability and accessibility of data resources. It accomplishes this by providing standardized, secure access protocols that utilize modern authentication and authorization mechanisms. This modality enables both machines and researchers to autonomously navigate the data landscape of the NOUS project and identify resources without the necessity for manual intervention. Through the use of role-based access controls embedded within API gateways, the NOUS project ensures that data repositories enhance their accessibility and interoperability, providing users and automated systems with explicit pathways to traverse between datasets, metadata, and ancillary resources.

2. NOUS Data Access Gateway (NDAG) The introduction of the NOUS Data Access Gateway (NDAG) proposes an innovative methodology for bridging the gap from a resource's Persistent Identifier (PID) to the actual dataset. NDAG is engineered to surmount the challenges inherent in accessing data dispersed across myriad platforms and storage solutions, with a particular focus on those within the NOUS architecture. This solution employs implicit knowledge regarding the diverse data storage solutions deployed by consortium partners, thereby enabling automatic redirection from a PID to the requisite metadata and data resources. The NDAG service endeavors to simplify access to data by abstracting the complexities associated with specific access protocols and data formats used across different components of the NOUS project.

Both the Advanced Access Control and NOUS Data Access Gateway underscore the imperative of machine-actionable data accessibility, acknowledging the pivotal role of automation and security in managing and protecting escalating volumes of research data. Nonetheless, the fruition of these strategies is contingent upon a steadfast commitment to standards and collaborative endeavors across the consortium to ensure seamless interoperability. Moreover, the endurance of such initiatives hinges on active community participation and the cultivation of a shared infrastructure conducive to the FAIR principles.

To further enhance the accessibility of data, the NOUS project aligns with the broader initiatives of the European Data Strategy by integrating with European data spaces. This ensures that the data managed within the NOUS project can be easily accessed and utilized by other researchers and innovators across Europe, facilitating cross-sector and interdisciplinary collaborations that are essential for the digital transformation of the continent.

Acknowledging the complexity of the technologies involved, the NOUS project includes targeted training sessions and capacity-building workshops to ensure that all project participants and stakeholders are well-equipped to use the NOUS platform effectively. This educational component is vital for ensuring that users can fully utilize the data, contributing to the overall accessibility and applicability of the project outcomes.

4.4.1 CLEAR GUIDELINES ON DATA ACCESS

Open Access Publishing: Wherever possible, the NOUS project commits to publishing its findings and data in open-access formats, ensuring that researchers, policymakers, and the public can access the information without restrictions. This commitment is aligned with the project's goal to enhance European digital sovereignty and promote transparent scientific inquiry.

Sensitive Data Handling: For data classified as sensitive, particularly those involving critical infrastructure or personal data, the NOUS project outlines specific protocols to ensure that such data is accessible only to authorized personnel. This includes the use of secure data sharing platforms and restricted access repositories, employing advanced encryption methods and secure authentication mechanisms to protect data integrity and confidentiality.

4.4.2 USE OF STANDARDIZED DATA FORMATS AND METADATA

Metadata Provision: To facilitate easy access to data, the NOUS project emphasizes the importance of accompanying all datasets with comprehensive metadata. This metadata includes detailed information on how to access the data and any conditions or restrictions on its use. The metadata also describes the data's origin, context, quality, and structure, ensuring that it can be effectively integrated and used within the NOUS architecture and beyond.

Standardized Formats: Data within the NOUS project is stored and shared in standardized formats that are widely recognized and easily accessible across different platforms and software. This standardization facilitates easier access by researchers, technologists, and other stakeholders, ensuring interoperability and reducing barriers to data usage.

4.4.3 STAKEHOLDER ENGAGEMENT AND DISSEMINATION ACTIVITIES

Communication Strategy: Through its dissemination and communication strategy, the NOUS project ensures that stakeholders are well-informed about how to access the data and findings. This involves the use of various communication channels, including the project's website, social media, newsletters, and virtual workshops. The strategy is designed to engage a broad audience, from academic and industrial partners to policymakers and the general public.

Training and Workshops: The project includes training sessions and workshops aimed at stakeholders to familiarize them with the data access protocols and platforms. These educational activities ensure that all stakeholders can effectively use and benefit from the project's outputs. The workshops also serve as a feedback mechanism, allowing stakeholders to provide input on data management practices and suggest improvements.

4.5 MAKING DATA INTEROPERABLE

The NOUS project aligns with best practices for data accessibility and interoperability, proposing a refined strategy to ensure that the project's data management system adheres to the highest standards of data utilization and sharing. This strategy encompasses key components tailored to enhance the project's operational efficacy and collaborative potential, particularly in the context of integrating advanced computational technologies and cybersecurity measures. WP4's emphasis on compliance and standardization is integral to this strategy, which encompasses key components tailored to enhance the project's operational efficacy and collaborative potential.

Standard Data Formats for Aggregation and Analysis: In the NOUS project, the adoption of open, standardized data formats such as CSV, JSON, and XML is crucial for effectively aggregating and analyzing data across diverse technological domains. These formats facilitate the seamless integration and processing of data from disparate sources, including quantum computing outputs and blockchain logs, enabling efficient and coherent data analyses.

Common Data Models and Schemas for Structured Aggregation: Implementing widely recognized data models and schemas, such as the Dublin Core for metadata and schema.org for structured data on the web, ensures that data aggregated within the project is consistently structured. This uniform structuring is vital for integrating data from different computational platforms and extracting meaningful insights related to digital sovereignty and cybersecurity.

Controlled Vocabularies and Ontologies for Semantic Consistency: The use of controlled vocabularies and ontologies standardizes the terminology within the aggregated data, ensuring semantic consistency across the datasets. This is essential for interoperable analysis of complex fields related to advanced digital technologies, where precision in terminology significantly impacts data interpretation.

Metadata Standards for Enhanced Data Description: Adopting metadata standards relevant to the digital and technological sciences, such as the Data Documentation Initiative (DDI), enriches the data with comprehensive descriptions. This enhanced metadata facilitates the integration of datasets with other data, boosting interoperability and reusability for future analyses within and beyond the project.

Data Exchange Protocols for Seamless Data Sharing: Employing data exchange protocols like OAI-PMH and RESTful APIs enables the seamless sharing of aggregated data within the project and with external partners. These protocols are crucial for the fluid exchange of data, supporting collaborative analyses and the synthesis of findings related to digital technologies and security practices.

Data Quality Assurance for Reliable Analysis: Ensuring the high quality of aggregated data through validation checks and regular data cleaning underpins the reliability of analyses conducted. High-quality data is fundamental for accurately identifying and addressing security vulnerabilities and enhancing digital infrastructure, contributing to the project’s overall objectives.

Persistent Identifiers for Data Traceability: The application of Persistent Identifiers (PIDs), such as DOIs, facilitates the long-term traceability of datasets. PIDs are crucial for linking datasets and ensuring that references to data remain valid over time, supporting the FAIR principle of reusability.

Collaboration and Standard Development for Interoperability: Active participation in the development of data standards and collaboration with other stakeholders in the digital and cybersecurity communities ensures that the data management practices align with globally supported standards. This collaboration is key to addressing interoperability challenges and enhancing the reusability of data.

By integrating these strategic components, the NOUS project not only aligns with the FAIR principles but also optimizes its approach to managing and utilizing data. This strategic alignment ensures that data aggregated and analyzed within the project is interoperable, reusable, and poised to contribute significantly to the advancement of European digital sovereignty and security.

4.6 INCREASE DATA RE-USE

The NOUS project places a high priority on enhancing data re-use, reflecting the broader scientific and technological community's focus on accelerating innovation, optimizing research investments, and maintaining the integrity of scientific findings through reproducibility. To support this objective, we have conducted a literature review and found practices and perceptions across disciplines, technical and ethical hurdles, and strategies to foster data re-use.

Research by Tenopir et al. (2020) delves into the practices and perceptions of scientists worldwide concerning data management, sharing, and re-use. Despite a generally positive stance towards data sharing and re-use, actual practices frequently lag, with data often stored on personal or departmental servers, hindering long-term preservation and access. This underscores the necessity for dedicated data managers, accessible repositories, and educational initiatives to cultivate sound data management practices.

Hemphill et al. (2022) explored factors influencing data re-use, highlighting the pivotal role of thorough curation and dedicated funding for dataset access and preservation in facilitating re-use. This indicates a direct correlation between investment in data curation, preservation, and the frequency of data re-use.

The literature review by Bote and Termens (2019) examines the technical and ethical challenges associated with data re-use, including issues of interoperability among repositories and concerns around data privacy. Emphasizing the role of standards in validating data re-use and the enhancement of metadata, the review calls for improvements to facilitate dataset discovery.

Pronk (2019) introduced a mathematical model to quantify the efficiency gains derived from data sharing within the scientific community. The findings posit that data sharing, when properly supported, can significantly enhance community efficiency, underscoring the benefits of facilitating and encouraging data sharing and re-use.

Wang, Duan, and Liang (2021) dissected the data re-use process into initiation, exploration and collection, and repurposing stages, illuminating the iterative nature of data re-use. Key takeaways include the importance of trust in data sources and identifying barriers to effective data re-use.

Epperson et al. (2022) investigated data sharing and re-use strategies among data scientists in software teams, identifying common obstacles such as lack of incentives and the challenge of making data science code modular. The study advocates for the development of tools tailored to mitigate these barriers, thereby enhancing data re-use.

Investment in Data Curation and Preservation: Research underscores the critical role of thorough data curation and dedicated funding for dataset access and preservation in facilitating data re-use. In the NOUS project, substantial investments are made to ensure that data is well-curated, properly maintained, and preserved, enhancing the likelihood of its re-use in future projects and studies. This involves the establishment of accessible repositories and the employment of dedicated data managers to oversee these processes.

Standardization of Data Formats and Metadata: The NOUS project commits to using standardized data formats such as CSV, JSON, and XML, and enriching datasets with comprehensive metadata. This metadata includes detailed descriptions of the data's context, quality, and collection methodologies, significantly improving dataset discoverability and interoperability. Standardized formats ensure compatibility with a wide range of software tools, facilitating easier integration and analysis.

Open Access and Licensing: Aligning with the principle of open access, the NOUS project ensures that all datasets are deposited in reputable open access repositories, maximizing their visibility and availability to the global research community. The adoption of clear, permissive licensing, such as Creative Commons licenses, clarifies usage rights while protecting the creators' intellectual property, thus encouraging more widespread data re-use.

Enhanced Documentation and Training: To further support data re-use, the NOUS project provides detailed documentation for each dataset, including user guides and manuals that elucidate effective data interpretation and utilization. Additionally, training materials, webinars, and workshops are developed to educate stakeholders on the data's significance, potential applications, and best practices for re-use.

APIs and Data Retrieval Tools: A commitment to data standards and the implementation of APIs for accessible data retrieval are pivotal for enhancing data interoperability and integration with existing workflows. This allows for the seamless exchange of data across platforms and disciplines, enabling more robust and collaborative research endeavors.

Community Engagement and Feedback Mechanisms: The project recognizes the importance of active community engagement and establishes feedback mechanisms to capture user experiences, challenges, and suggestions. Platforms such as collaborative forums and web applications foster a vibrant community of contributors, sharing data, analyses, and findings to enrich the ecosystem of reusable data.

Case Studies and Success Stories: By publishing case studies and success stories of data re-use, the NOUS project not only demonstrates the datasets' value but also inspires further exploration and utilization. These narratives serve as practical examples of how data can be leveraged to achieve significant outcomes in digital security and technological advancements.

Long-term Preservation and Curation: Ensuring the long-term preservation and curation of data is a key priority, with efforts directed towards sustainable storage solutions and regular data updates to maintain relevance and utility. This approach guarantees that the datasets remain a valuable resource for future research initiatives and technological development.

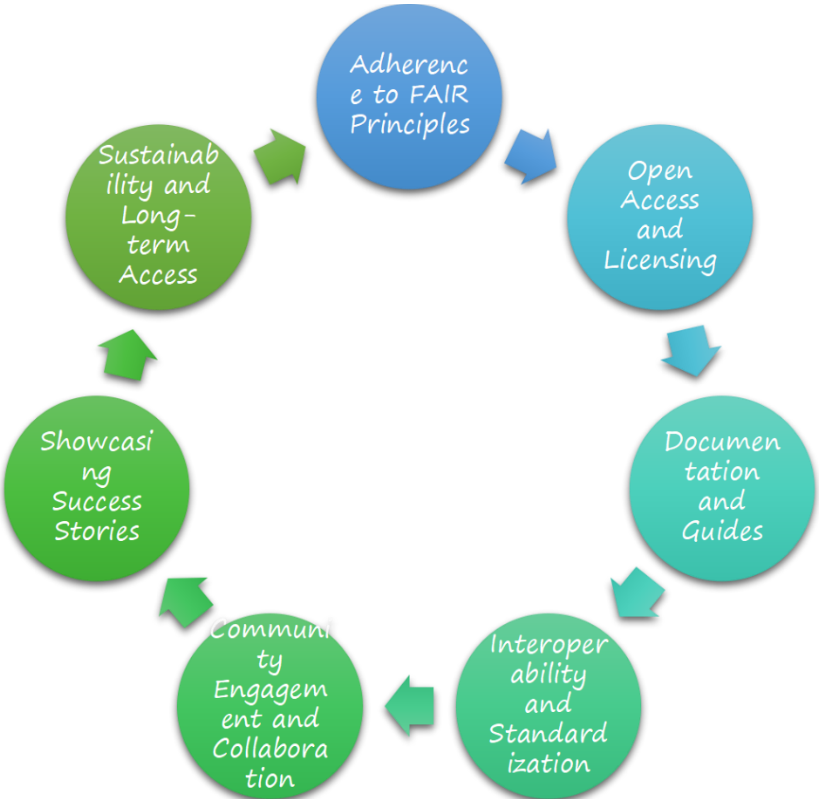
Adherence to FAIR Principles
Open Access and Licensing
Documentation and Guides
Interoperability and Standardization
Community Engagement and Collaboration
Showcasing Success Stories
Sustainability and Long-term Access

FIGURE 2: DATA RE-USE PROCESS

5 OTHER RESEARCH OUTPUTS

Within the framework of preparing the Data Management Plan for the NOUS project, the "Other Research Outputs" survey provided insights into consortium members' strategies for handling and disseminating both traditional and non-traditional datasets. NOUS project has developed a comprehensive strategy for the management and dissemination of both digital and physical research outputs. This strategy is tailored to align with the overarching goals of advancing digital sovereignty and cybersecurity within Europe and is grounded in the responses provided by the NOUS project participants.

5.1 STRATEGIES FOR OUTPUT SHARING AND MANAGEMENT

NOUS project participants have highlighted various approaches to the management of research outputs. The approaches to sharing and managing these digital outputs are informed by a blend of internal guidelines, legal standards, and ethical principles, that is aligned with D8.1. For instance:

- Entities like, FUNDACION INSTITUTO INTERNACIONAL DE INVESTIGACION EN INTELIGENCIA ARTIFICIAL Y CIENCIAS DE LA COMPUTACION (AIR) and AETHON, emphasize that research outputs will be open wherever possible, without compromising exploitable opportunities. This approach ensures that outputs are accessible while still protecting intellectual property and potential commercialization interests.

- NATIONAL CENTER FOR SCIENTIFIC RESEARCH "DEMOKRITOS" (NCSR "D") and POLITECNICO DI TORINO (POLITO) plan to publish research outputs in conferences and journals, preferably open access, to maximize dissemination and re-use. Participants like IOTAM INTERNET OF THINGS APPLICATIONS AND MULTI LAYER DEVELOPMENT LTD (ITML CY) mention that where possible, codebases will be available on public domains like GitHub.
- DIMOSIA EPICHEIRISI ILEKTRISMOU ANONYMI ETAIREIA (PPC) notes the importance of security measures before making datasets public to avoid data leaks and protect the interests of the data owner and consortium.
- As per F6S NETWORK IRELAND LIMITED (F6S IE), each research output should have an exploitation plan that defines its management in both digital and physical terms. This ensures that outputs are not only disseminated widely but are also integrated into future projects or products depending on their Technology Readiness Level (TRL).
- INESC TEC mentions the use of a Zenodo community for sharing research outputs, where deposits are managed and approved by designated data curators, emphasizing the structured and community-driven approach to data re-use.

The emphasis on digital data across participant responses reflects a broader movement towards the digitization of research management practices. This trend acknowledges the instrumental role of digital tools in streamlining the organization, examination, and distribution of research findings. Additionally, the collective dedication to abiding by internal policies, GDPR requirements, and ethical standards in the handling and sharing of research outputs denotes a multifaceted strategy that prioritizes legal adherence, data privacy, and the intrinsic value of research outputs to external audiences.

By proactively planning for the stewardship and dissemination of digital research outputs, mindful of legal and ethical imperatives, consortium members exhibit a forward-thinking approach to responsible research conduct. This methodology not only enhances the utility of research findings within the academic and scientific domains but also contributes to the broader objective of promoting research practices that are transparent, accessible, and consonant with societal expectations and regulatory frameworks.

6 ALLOCATIONS OF RESOURCES

The financial implications for data storage and upkeep won't necessitate additional funding after the project concludes. It's crucial to recognize the significance of the data, given its alignment with the pressing demands of the energy sector and consumer needs. Consequently, the data derived from this project are anticipated to have an immediate impact in the ensuing years, although their relevance may diminish as current challenges are addressed or supplanted by new priorities.

The duty of managing the project's document repository falls to the Project Coordinator, whereas the Scientific Coordinator is tasked with guaranteeing the integrity of all scientific data produced. Meanwhile, each partner is responsible for ensuring the data they produce can be recovered. As previously mentioned, the costs related to adhering to the FAIR principles for the data—making it Findable, Accessible, Interoperable, and Reusable—remain uncertain at this juncture, given that they will be influenced by the total amount of data generated.

From the survey conducted, most of the participants have already assigned a responsible person for Data Management in each entity and have indicated data long-term preservation methods, as presented in the table below.

TABLE 1: ALLOCATION OF RESOURCES

PARTNER ENTITY	RESPONSIBLE FOR DATA MANAGEMENT	DATA LONG-TERM PRESERVATION METHODS
FUNDACION INSTITUTO INTERNATIONAL DE INVESTIGACIONEN INTELIGENSIA ARTIFICIAL Y CIENCIAS DE LA COMPUTACION (AIR)	ENRIQUE MESONERO & ALBERTO MONTERO	AS AGREED IN THE CONDITIONS AS A EUROPEAN PROJECT, THE DOCUMENTATION OF THIS PROJECT MUST BE PRESERVED.
AETHON ENGINEERING SINGLE MEMBER PC (AETHON)	ALEX PAPACHARALAMPOUS	WITH SUSTAINABLE, RELIABLE STORAGE SOLUTIONS WITH REGULAR BACKUPS, MIGRATION STRATEGIES FOR DATA STORED ON OBSOLETE TECHNOLOGIES.
UNIVERSIDAD DE SALAMANCA (USAL)	DAVID SANCHEZ SANCHEZ	AS LONG AS THIS PROJECT LASTS.
NATIONAL CENTER FOR SCIENTIFIC RESEARCH "DEMOKRITOS" (NCSR "D")	PANAGIOTIS KROKIDAS AND CHRISTOFOROS REKATSINAS	
CONSORZIO INTERUNIVERSITARIO PER L' OTTIMIZZAZIONE E LEA RICERCA OPERATIVA (ICOOR)		
POLITECNICO DI TORINO (POLITO)	MARCO GHIRARDI	USING THE BEST TECHNOLOGIES PROVIDED BY THE (LONG TERM) DATA MANAGEMENT INDUSTRY
UNIVESITA DEGLI STUDI DI MODENA E REGGIO EMILIA (UNIMORE)	MARCO MAMEI	WE DO NOT HAVE PLANS FOR LONG TERM PRESERVATION OF DATA (OTHER THAN PUSHING

		DATA TO EXTERNAL SERVICES - E.G., GITHUB, NATURE SDATA)
TELECOM ITALIA SPA (TIM)		
NETCOMPANY - INTRASOFT SA (NET-INTRA)	LAW OFFICE “PISTIOLIS – TRIANTAFYLLOS & ASSOCIATES”, CONTACT PERSON: NIKOLAOS ZELIOS	NET-INTRA FOLLOWS THE CONTRACTUAL OBLIGATIONS THAT EMERGE FROM THE GA AND IT WILL RETAIN THE DATA ONLY FOR THE PROJECT PURPOSE
IOTAM INTERNET OF THINGS APPLICATIONS AND MULTI LAYER DEVELOPMENT LTD (ITML CY)	DIMITRA SIAILI	STORED IN TRUSTED REPOSITORY UNTIL THE END OF THE PROJECT
UNIVERSITA DI PISA (UNIPI)	UNIPI DPO	DEPENDS ON THE DATA. DETAILS STILL TO BE DEFINED
UNPARALLEL INNOVATION LDA (UNPARALLEL)		
KATHOLIEKE UNIVERSITEIT LEUVEN (KU LEUVEN)		
HEWLETT - PACKARD HELLAS ETAIREIA PERIORISMENIS EFTHINIS (HPE HELLAS)		
ECLIPSE FOUNDATION EUROPE GMBH (ECL)	DAVID REMON	DATA WILL BE STORED IN AN PUBLIC REPOSITORY, GITLAB, ENSURING THE LONG-TERM PRESERVATION
F6S NETWORK IRELAND LIMITED (F6S IE)	SOFIYA SAVOVA	LONG-TERM PRESERVATION WILL BE ENSURED THROUGH RELIABLE STORAGE PLATFORMS, REGULAR BACKUP AND A SUSTAINABLE ATTITUDE TOWARDS DATA MANAGEMENT FOR THE NOUS PROJECT.

DIMOSIA EPICHEIRISI ILEKTRISMOU ANONYMI ETAIREIA (PPC)	PPC'S CORRESPONDING DEPARTMENT	PPC'S PRIVATE CLOUD REPOSITORY
ARCTUR RACUNALNISKI INZENIRING DOO (ARCTUR D.O.O.)	DMO	STORED ON OUR SYSTEMS, SYSTEMS ARE WELL-SECURED. WE ARE OPERATING UNDER THE AUSPIECES OF ISO27001 CERTIFICATE.
AEGIS IT RESEARCH GMBH (AEGIS)	EVANGELOS RAPTIS	STORED IN TRUSTED REPOSITORY.
CS GROUP - FRANCE (CS GROUP - FRANCE)	DMO	STORED ON OUR SYSTEMS WITH REGULAR BACKUP.
INESC TEC - INSTITUTO DE ENGENHARIADE SISTEMAS E COMPUTADORES, TECNOLOGIA E CIENCIA (INESC TEC)	INESC TEC HAS A DATA STEWARD PROVIDING FAIR SUPPORT FOR I&D PROJECTS. A RESPONSIBLE PARTY FOR DAILY MANAGEMENT OF DATA MAY BE DEFINED AT PROJECT LEVEL, OR THE DATA CREATORS ARE RESPONSIBLE FOR THE MANAGEMENT OF THEIR DATA.	DATASETS DEPOSITED IN THE INESC TEC DATA REPOSITORY ARE PRESERVED INDEFINITELY.

7 DATA SECURITY

In the framework of enhancing research integrity and ensuring robust data protection, the NOUS project implements stringent measures to safeguard personal data. Recognizing the sensitive nature of the data involved, particularly in areas related to cybersecurity and digital infrastructure, all datasets containing personal information undergo a thorough anonymization process. This ensures that personal data collected during the project's lifecycle is handled with the utmost care for privacy and confidentiality.

The collection of personal data is tightly regulated and confined primarily to the initial phases of project activities. Crucially, the participation of individuals is contingent upon their voluntary and informed consent, which is obtained explicitly granting permission for the use of their personal information within the project's scope.

To maintain the confidentiality of personal identities, the NOUS project employs a system of anonymous coding. This system involves assigning unique codes to each participant's data, effectively masking their real identities. The linkage between these codes and the actual names of the participants is held in strict confidence, accessible only to authorized project partners. These partners bear the responsibility for ensuring that such sensitive records are stored in a secure environment, safeguarded from unauthorized access.

In scenarios where the dissemination of applications or data to external evaluators is necessary, the anonymized coding system facilitates this process without compromising personal privacy. Should there be a need to share data with collaborators outside the European Union, the project commits to obtaining the necessary clearances from the designated Data Protection Authorities, unless the data is transferred to a country already deemed by the EU to have adequate data protection measures in place.

Project partners are required to provide the European Commission with documented evidence of all approvals or notifications related to the processing of personal and sensitive data, as endorsed by the relevant institutional Data Protection Offices. This commitment ensures transparency and compliance with regulatory requirements.

To further bolster data security, personal information and sensitive data are encrypted during storage and transmission, minimizing the risk of data breaches or unauthorized access. The project's data protection strategy is meticulously designed to align with the General Data Protection Regulation (GDPR), which sets forth stringent guidelines for the processing of personal data and the free movement of such data within the EU.

By adhering to these regulatory frameworks and implementing advanced security measures, the NOUS project underscores its dedication to the ethical handling, privacy, and security of personal data throughout its duration. This comprehensive approach not only protects individual rights but also enhances the trustworthiness and reliability of the project's research outputs.

7.1 DATA SECURITY MEASURES

The "Data Security" survey conducted among NOUS consortium members has yielded significant insights into their strategies for ensuring data security, which are crucial for safeguarding sensitive information and maintaining trust in digital interactions. The responses emphasize various aspects of data security, demonstrating a robust commitment to secure and responsible data management.

7.1.1 HYBRID STORAGE SOLUTIONS

Several NOUS participants indicated a preference for hybrid storage solutions, combining on-premises storage with cloud-based solutions to leverage both security and scalability:

POLITECNICO DI TORINO (POLITO) discusses using cloud services under the governance of the EU's Cybersecurity Act, highlighting the strategic integration of advanced cybersecurity measures.

UNIVERSIDAD DE SALAMANCA (USAL) and DIMOSIA EPICHEIRISI ILEKTRISMOU ANONYMI ETAIREIA (PPC) employ strategies that emphasize secure on-premises storage, supplemented by cloud solutions to enhance accessibility and reliability.

7.1.2 UTILIZATION OF SECURE SOFTWARE SOLUTIONS

The adoption of secure software solutions such as SharePoint and other enterprise-level services underscores a commitment to robust data security frameworks:

FUNDACION INSTITUTO INTERNACIONAL DE INVESTIGACION EN INTELIGENCIA ARTIFICIAL Y CIENCIAS DE LA COMPUTACION (AIR) and AETHON ENGINEERING SINGLE MEMBER PC (AETHON) implement SharePoint for secure storage, which is managed to ensure that access is strictly controlled, and data is encrypted during transfer.

7.1.3 IMPLEMENTATION OF INTERNAL DATA PROTECTION MEASURES

The development of bespoke data protection protocols is highlighted by several members, which address unique security challenges:

ECLIPSE FOUNDATION EUROPE GMBH (ECL) and F6S NETWORK IRELAND LIMITED (F6S IE) have established specific security measures, including the use of GitLab for managing access controls and ensuring data encryption, along with comprehensive privacy by design strategies.

7.1.4 FRAMEWORK FOR DATA PROTECTION

Integrating robust data security measures within the NOUS project is critical for safeguarding the integrity and confidentiality of data throughout its lifecycle. This comprehensive approach encompasses:

Encryption: Implementing encryption for data both at rest and in transit forms a fundamental layer of protection, rendering data unintelligible to unauthorized parties. For NOUS, encryption ensures the secure handling of sensitive research data, protecting it from potential breaches during transmission and storage.

Access Control: Stringent access controls enforced by several entities ensure that only authorized personnel can access sensitive information. Access control mechanisms are critical in defining and enforcing who can view or use the project's data. By establishing robust authentication methods and precise user role definitions, the project minimizes the risk of unauthorized data access, ensuring that only verified users can access sensitive information.

Advanced Security Measures: Deployment of firewalls is crucial for establishing a secure network perimeter around the project's digital infrastructure. This helps prevent unauthorized access and cyberattacks, protecting both the data and the systems used for project activities.

Antivirus and anti-malware software: Protecting project assets from malware requires the deployment of comprehensive antivirus solutions. Regular updates and scans will defend against the latest threats, a necessity given the evolving nature of cybersecurity risks.

Data back-up and Recovery: Establishing a rigorous data backup and recovery plan ensures the project's resilience against data loss scenarios. Secure and regularly tested backups, as underlined by most of the participants, support the project's continuity, safeguarding against data loss from technical failures or cyber incidents.

Regular Security Audits and Compliance Checks: Conducting periodic security audits verifies the effectiveness of the project’s security measures. These assessments, coupled with compliance checks, ensure that the project adheres to relevant data protection standards and regulations.

Overall, the collective responses from NOUS consortium members highlight a sophisticated and layered approach to data security. By integrating hybrid storage models, utilizing enterprise-grade software solutions, and formulating specific internal data protection protocols, the consortium demonstrates its commitment to leveraging technological advancements for secure data storage and management. This strategic approach not only ensures the integrity and confidentiality of the project's data but also facilitates efficient data management and collaborative research activities, aligning operations with the project’s objectives within a secure and protected environment.

8 ETHICS

8.1 ETHICAL ASPECTS

To ensure adherence to all legal and ethical standards, the NOUS project has developed a comprehensive strategy guided by the Ethics Requirements, outlined in the deliverables of Work Package 4. This strategy includes a bespoke monitoring process for the project's progression, utilizing a privacy-by-design framework through a methodological plan rooted in a "Socio-legal Approach." This approach addresses privacy and data protection concerns by adopting a risk management perspective, in accordance with the principles set forth in the General Data Protection Regulation (GDPR).

8.2 NOUS ETHICS

Within the NOUS project, a series of protocols have been developed to ensure the privacy of all participating end-users is adequately protected. The consortium overseeing the project will carefully control access to information, imposing necessary restrictions. Essential measures include:

- Maintaining ethical standards and guidelines that align with those set by Horizon Europe, regardless of the location of the NOUS demonstrations.
- Providing all participants with clear explanations of the project's aims and the objectives of the study in an understandable format.
- Highlighting the voluntary basis of participation in the study.
- Informing participants about their privacy rights, the possible effects on their lives, and the privacy protection strategies in place, including data anonymization and secure storage techniques.
- Outlining the expected duration and commitment required from participants for any activities.
- Detailing participants' rights to withdraw from the study at any time and request the deletion of their personal data.
- Offering access to contact information for key project stakeholders.

The consortium commits to upholding ethical standards through ongoing reporting mechanisms. Should the project entail human involvement, participants will be briefed on privacy, confidentiality, and adherence to both national and EU legislation. Participants will receive understandable

Information Sheets and Informed Consent Forms, elucidating the voluntary nature of participation, potential risks and benefits, and protocols for handling incidental findings.

Participants will have the opportunity to pose questions, receive clear responses, and retract their participation and data at any point without any adverse effects.

Throughout the project, only anonymized or aggregated data, devoid of any individual identifiers, will be utilized for workshops and events. The same data privacy principles will apply during the dissemination phase. Should personal data processing become necessary under certain conditions, a designated Data Protection Officer will be assigned by the responsible partner to guarantee GDPR compliance and secure prior authorization for data processing.

Authorization from the relevant authority in the partner's country will be required in accordance with European and national laws. For the use of specific platforms like Twitter, LinkedIn, Facebook, Google Cloud, etc., that may involve personal data processing, a Data Processing Addendum/Agreement will be secured by the responsible partner.

8.3 NOUS COMPLIANCE

The NOUS consortium is acutely aware that the project's activities may lead to concerns regarding ethics, fundamental rights, privacy, and data protection. As such, there is a firm commitment to comply with the most stringent ethical, fundamental rights, and legal standards established both within the European Union and on an international scale. This commitment entails concerted efforts to ensure that the project's proposals are in harmony with principal ethical guidelines and fundamental rights as delineated in key documents, including the EU Charter on Fundamental Rights, the European Convention for the Protection of Human Rights and Fundamental Freedoms, and the General Data Protection Regulation (GDPR).

Through this comprehensive approach, the NOUS consortium underscores its dedication to upholding the highest ethical and legal standards. This commitment not only facilitates the responsible conduct of research but also ensures that the project's outputs are achieved in a manner that respects the ethical considerations and fundamental rights central to its objectives.

9 CONCLUSIONS

The Data Management Plan (DMP) of the NOUS project underscores its pivotal role in pioneering FAIR (Findable, Accessible, Interoperable, and Reusable) data management practices tailored for the evolving landscape of European digital sovereignty and advanced cybersecurity measures. The project's accomplishments are marked by the formulation of detailed data management plans, the creation of accessible online platforms and repositories, and the execution of strategies aimed at enhancing data interoperability and reusability. Central to its mission, NOUS has upheld stringent ethical and security standards, prioritizing privacy and data protection throughout its lifecycle.

The project has successfully achieved its initial objectives, significantly improving data management, stewardship, and reusability within the framework of digital innovation by adhering to FAIR principles. Through its structured deliverables and adaptive strategies, NOUS has navigated the complexities of evolving technological and regulatory landscapes.

The implications of NOUS's work extend far beyond its immediate objectives, setting a benchmark for future digital infrastructure projects by promoting efficient data utilization, enhancing cybersecurity measures, and fostering innovation in R&D efficiency. These achievements are instrumental in driving forward the agenda of European digital sovereignty, contributing to economic growth and technological advancement.

Lessons gleaned from the project highlight the criticality of thorough data management planning, the benefits of collaborative stakeholder engagement, and the necessity for adaptability in response to technological advancements and regulatory changes. Challenges in data privacy, security, and interoperability were met with targeted responses, underscoring the project's proactive approach to problem-solving.

Looking ahead, future research and development could explore enhancing data interoperability standards, creating innovative data privacy and security solutions, and developing sophisticated analytical tools that utilize FAIR data. Addressing these areas will be crucial for maintaining the momentum of Europe's digital evolution.

For policymakers, the project emphasizes the importance of creating regulatory environments that support FAIR data practices, promote open access to research data, and secure robust data protection. Incentivizing the adoption of cutting-edge data management and analysis technologies could further align governmental and industrial practices with European digital sovereignty goals.

In order to ensure the continued success and sustainability of the data management practices established by the NOUS project, continuous monitoring activities will be embedded, using internal trackers and templates that partners will be required to periodically fill out. The aforementioned practice will help in tracking compliance, identifying any deviations, and making necessary adjustments in real-time. Furthermore, partners will submit regular reports on their data management activities, challenges faced, and compliance status. These reports will be reviewed to ensure adherence to the DMP and to address any issues promptly. Lastly, next steps include organizing webinars and training sessions to support consortium members in tackling the challenges identified in their responses to the survey, as well as ensuring that metadata and datasets are uploaded to public repositories such as ZENODO, obtaining DOIs to enhance data findability and accessibility.

Ultimately, the NOUS project stands as a landmark contribution to the fields of digital infrastructure and cybersecurity, establishing new precedents for data management, accessibility, and utility. Its legacy promises not only to enhance the efficiency and innovation of current R&D endeavours but also to lay the groundwork for a future characterized by collaborative, innovative, and secure digital practices. This concise conclusion encapsulates the project's comprehensive achievements and forward-looking aspirations, reflecting the deliverable's intent for a meaningful and informative closure.

10 REFERENCES

1. Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3. https://doi.org/10.1038/sdata.2016.18

2. Vlijmen, H., Bouwman, J., Karkkainen, L., & Liberatore, F. (2020). The Need of Industry to Go FAIR. *Data Intelligence*, 2(1), 276-284. https://doi.org/10.1162/dint_a_00050

3. Wise, J., de Barron, A. G., Splendiani, A., Fluck, J., Hunter, S., & ten Hoopen, P. (2019). Implementation and relevance of FAIR data principles in biopharmaceutical R&D. *Drug Discovery Today*, 24(4), 933-938. https://doi.org/10.1016/j.drudis.2019.01.008

4. Broeder, D., Elbers, W., Gawor, M., Concordia, C., Larrousse, N., & Uytvanck, D. (2022). Towards FAIR Data Access. *Research Ideas and Outcomes*, 8. https://doi.org/10.3897/rio.8.e94386

5. Tenopir, C., Rice, N., Allard, S., Baird, L., Borycz, J., Christian, L., Grant, B., Olendorf, R., & Sandusky, R. (2020). Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. *PLoS ONE*, 15(3), e0229003. https://doi.org/10.1371/journal.pone.0229003

6. Hemphill, L., Pienta, A., Lafia, S., Akmon, D., & Bleckley, D. (2022). How do properties of data, their curation, and their funding relate to reuse?. *Journal of the Association for Information Science and Technolog*y, 73, 1432-1444. https://doi.org/10.1002/asi.24646

7. Bote, J., & Termens, M. (2019). Reusing Data: Technical and Ethical Challenges. *DESIDOC Journal of Library & Information Technology*, 39(6), 329-337. https://doi.org/10.14429/djlit.39.06.14807

8. Pronk, T. (2019). The Time Efficiency Gain in Sharing and Reuse of Research Data. *Data Science Journal*, 18, 10. https://doi.org/10.5334/DSJ-2019-010

Wang, X., Duan, Q., & Liang, M. (2021). Understanding the process of data reuse: An extensive review. *Journal of the Association for Information Science and Technology*, 72, 1161-1182. https://doi.org/10.1002/asi.24483.

11 APPENDIX A

NOUS DMP SURVEY

The D1.3 DMP survey, conducted among NOUS consortium members, offers a structured overview of data management intentions and practices across various aspects of research and development projects.

The survey is organized into distinct sections—Allocation of Resources, Ethics, Data Summary, FAIR Data, Other Research Outputs, and Data Security—aiming to capture the consortium's strategies for managing, sharing, and securing data responsibly and in line with ethical, legal, and operational standards throughout the project lifecycle.

11.1 DMP SURVEY RESPONSES

11.1.1 DATA SUMMARY

PARTNER	**Q1.1: What existing data will you re-use, and for what purpose?**
AIR	Within the scope of our project, we will reuse administrative and tracking data templates.
AETHON	Re-use existing data for metrics, performance benchmarks, and user data from prior projects to enhance system design, optimize performance, and ensure robust user data management
USAL	USAL does not directly reuse data; our role is to maintain the SharePoint server where project data is stored and ensure that permissions and access are correctly configured for use by the AIR coordinator.
NCSRD	
ICCOR	
POLITO	1) AUTOMOTIVE SMART AREA data to sustain Use Case #1 2) ENERGY PREDICTION AND ENERGY DATA LIFECYCLE MANAGEMENT data used to sustain Use Case #2. 3) CRISIS MANAGEMENT AND CIVIL PROTECTION to sustain Use Case #3. 4) SCIENTIFIC DATA HPC STORAGE AND AI ANALYTICS to sustain the Use Case #4

UNIMORE	Mobility data in the MASA use case, and distributed systems performance to access deployed services
TIM	
NET-INTRA	NET-INTRA is not a data provider - so no existing data will be brought in
ITML CY	Open accessible data for conducting experimentations; Possibly Pilot data; for training
UNIPI	Open accessible data for conducting experimentations; logs and performance data; Pilot data;
UNP	
KUL	None/Not applicable
HPE	
ECL	The code developed by the partners in the project as open source will be hosted at Gitlab, a public repository, under a Eclipse instance called the Eclipse Research Labs
F6S	F6S will not reuse any existing data.
PPC	Production data from PPC's power plants, customer, supplier, production, and energy data.
ARCTUR	None at the moment.
AEGIS	None at the moment.
CS GROUP - FRANCE	Re-use existing data from prior projects (when opened) and opened data: - for demonstration purposes, - to support of the overall system design and development - for evaluation.
INESC TEC	Human-Centered AI existing dataset, as the ones available at https://human-centered.ai/, would be used for a wider context to infer best practices for human in the loop for data-agile economy. This activity is part of T7.5 that is expected to start in M18. Further details about the datasets to be used will be available after the task kickoff.

PARTNER	Q1.2: What types and formats of data will you generate or re-use?
AIR	Videos of the presentations and discussions from monthly meetings. Text documents containing meeting minutes and attendance lists. Images and graphics presented during meetings. All of this data will primarily be maintained in multimedia formats (MP4, JPEG) and documents (PDF, DOCX).
AETHON	System logs, user activity logs, performance data, metadata, JSON and XML, etc.
USAL	As stewards of the SharePoint server, we do not generate or reuse data. Our role focuses on technical management and security of the platform.
NCSRD	
ICCOR	
POLITO	1) To apply NOUS services to improve connected vehicles perception relying on roadside cameras and other cars' cameras. 2) Leverages production data from electrical power plants as well as energy data. 3) Geospatial and situational information on a need-to-know basis between multiple stakeholders in several interconnected command centers and first - responders on the field. data for surveillance of the environment and critical infrastructures, data for management of crises offering, data to improve situational awareness and operation management capacities. 4) The types of data that will be examined, generated, and handled within this use case include: molecular simulation data, MOF design functional parameters data, nano-metrology data (including synthetic and real data from surface scans), structural material analysis data (such as e.g., damage data in multi-layered advanced materials), textual data relating to one or more of the above domains (e.g., related publications, dataset descriptions, etc.).
UNIMORE	Data logs, performance data, metadata in JSON and CSV
TIM	
NET-INTRA	Log data in JSON format
ITML CY	Data logs, propably in JSON and CSV format
UNIPI	Data logs, performance data;
UNP	
KUL	None/Not applicable

HPE	
ECL	None at the moment.
F6S	**Event registration and participation datasets:** Participant registration for an event will be done using an online registration form or in person at the event. The information collected will be the same independently of the registration is on or offline. Information collected will be the minimum needed to generate statistics regarding participation at the event. When relevant, consent for usage of participants' images and audio in dissemination actions will also be collected. Participant registration will be collected for every event run in the project. Information related to participants' registration will be stored on the project's documentation repository to enable long-term storage and access to data by consortium members. **Webinars and other video recordings datasets:** The project will organise and run one or more webinars that may be recorded for and manage a project website and social media channels to raise awareness about the project, engage stakeholders, and disseminate information and results. Webinars and other video recordings will be stored on the project's repository and/or public platforms. Webinars and other activities will be recorded with video in the following cases: - Participants have been informed and are able to withdraw consent to be recorded for the purpose of the webinar/ activity. - The recording is necessary to fulfill the objectives of the project. - Recording is in the public interest or is in the interests of the recorder unless those interests are overridden by the interests of the participants that require the protection of personal data. **Project website and social media datasets:** The project website will be set up to provide statistics that will be collected using a web analytics service. Only anonymized data will be collected that can not be tracked back to a specific individual. Social media platforms include their own analytics services that will be used to collect and provide information about user engagement on these platforms. In order to allow for a more thorough and varied analysis, data will remain stored on the analytics platforms. If relevant, data may be exported to the project's documentation repository to enable access by consortium members. Data will be stored separately, and divided into the website and relevant social media platforms. **External communication with third-party datasets:** The project partners will engage and communicate with third parties throughout the project on one or more of the planned activities. Communication and Information exchange between the partners and third parties will be carried out through one or more different platforms and may include e-mail, the F6S platform, or other communication platforms. The communications and information exchanged will primarily be stored on the platforms used for this purpose.
PPC	TBD
ARCTUR	None at the moment.
AEGIS	None at the moment, but e.g. system logs, network traffic data, after M18.

CS GROUP - FRANCE	- System logs, user activity logs, analysis, images , 3d models, video, gis data, ... - Numerous standard formats (JSON, PNG, ...) and internal formats
INESC TEC	None at the moment. Further details about the datasets will be available after the task T7.5 kickoff on M18.

PARTNER	Q1.3: What is the expected size of the data that you intend to generate or re-use?
AIR	It is anticipated that the total size of the generated and reused data will be approximately 50-100 GB annually, depending on the frequency and duration of meetings, as well as the amount of graphical material presented.
AETHON	TBD
USAL	Not applicable, as USAL does not generate or reuse project-specific data.
NCSRD	
ICCOR	
POLITO	While it is not possible athis point of the project to estimate the data the following is just an estimate thant can be refined furtther during the the project progress 1) N Terabytes with N larger than 40 2) M Terabytes with M larger than 80 3) W Terabytes with W larger than 40 4) Q Terabytes with Q larger than 80
UNIMORE	TBD (<1TB)
TIM	
NET-INTRA	KByte-MByte
ITML CY	TBD
UNIPI	TBD (<1TB)
UNP	

KUL	None/Not applicable
HPE	
ECL	None at the moment.
F6S	TBD
PPC	N/A
ARCTUR	None at the moment.
AEGIS	N/A
CS GROUP - FRANCE	from few MB to several GB
INESC TEC	None at the moment. Further details about the datasets will be available after the task T7.5 kickoff on M18.

PARTNER	**Q1.4: What is the origin/provenance of the data either generated or re-used?**
AIR	The reused data comes from internal project records, which include multimedia material and documentation generated by consortium partners during monthly meetings. This content is compiled and initially stored in the USAL SharePoint system, serving as the central repository for the project, ensuring controlled and organized access.
AETHON	Most of the data will be obtain from pilot partners.
USAL	We do not directly manage the origin of the data; we only provide technical support and infrastructure for its secure storage.
NCSRD	
ICCOR	

POLITO	1) Generated by cars /veichles traffic. 2) data that are publicly available through the web sites of the power production organization participating NOUS, Data are combined with past meteorological data and weather forecast data that are specialized wind and solar farms, from weather forecast public web sites. 3) A. First, the digital twin of the environment that can span over Large regions, information known about this area, which contains massive geospatial datasets including 3D terrain, aerial imagery, cadastral information, BIM models, without any limitation in terms of surface and resolution, for open space, urban and indoor environments. B. Second, live sensor data about the ongoing situation such as video streams from surveillance cameras or RPV/UAV, temperature or hygrometry probes, weather data, earth observation data. C. Third, tactical information (commands, situation reports, pictures, videos, audio recordings) exchanged between the various first responders and command chains of the multiple organizations involved in the management of a crisis. 4) The NCSRD team contributes to these data with the natural sciences and technology aspect, while the IIT part contributes its leading role to HPC (EuroCC and EuroCC 2 Project), data management and AI expertise (contribution to the AI on-demand platform of the EU and multiple related efforts), as well as multi-disciplinary related previous projects (e.g. DARE). All these are finely complemented by the ahead DIH understanding of how this expertise can be transferred to an industrial innovation setting.
UNIMORE	Data will be obtained from pilot applications
TIM	
NET-INTRA	Records of events occuring within the system
ITML CY	AI model parameters and metadata will be exchanged with the project's AI tools. Pilot Data
UNIPI	Open data, Pilot data;
UNP	
KUL	None/Not applicable
HPE	
ECL	Code developed by the project partners
F6S	Data generated by third parties (participants, subscribers, users) and project team members (consortium).
PPC	PPC data

ARCTUR	
AEGIS	From the project's technical assets (e.g. servers, computers, edge devices)
CS GROUP - FRANCE	-Internal (created by CS) -Open data (such as open street map for instance) -From the NOUS projet
INESC TEC	We will use existing open datasets as well as data collected from the project pilots.

PARTNER	**Q1.5: To whom might your data be useful outside your project?**
AIR	The data could be of interest to other research groups and projects working in similar thematic areas, especially those focused on collaboration dynamics and management of large-scale projects. Additionally, the documented results and methodologies could benefit educational institutions and organizations seeking to improve their project management practices and inter-institutional collaboration.
AETHON	Security Analysts, Cloud service providers
USAL	Not applicable, as we do not directly handle data that may be useful outside the project.
NCSRD	HPC experts, HPC users, LLM developers
ICCOR	
POLITO	1) Public bodies, Road Safety analysis centers, Suppliers of safety systems, Car manufacturers, Automotive Component producers, Suppliers of on-board services for urban and extra-urban infrastructures, and research centers. 2) Power Supplier Customers, Power Suppliers, Power Productors. 3) Public Officials in charge of civil protections , Military organisations , Red Cross, EU Citiziens. 4)The Use Case #4 modeling will describe the system from the viewpoint of the people or entities (the actors) who will interact with the system. A complete description of how these actors will use the system For Scientific Data Hpc Storage and AI Analytics to perform their tasks and achieve their objectives.
UNIMORE	Cloud service providers, edge computing analysis

TIM	
NET-INTRA	None
ITML CY	Not applicable
UNIPI	Cloud service providers, edge computing analysis
UNP	
KUL	None/Not applicable
HPE	
ECL	Developer community
F6S	EU institutions, cloud service providers, edge computing analyzers
PPC	Competitors
ARCTUR	National Competence Center?
AEGIS	Tasks related to cybersecurity of other research projects
CS GROUP - FRANCE	N/A
INESC TEC	None at the moment. Further details about the usage of data is expected after the task kickoff on M18.

11.1.2 FAIR DATA

PARTNER	**Q2.1: Will data be identified?**
AIR	
AETHON	TBD. In any case the data management process will comply with GDPR and other privacy regulations.

USAL	
NCSRD	
ICCOR	
POLITO	TBD. In any case the data management process will comply with GDPR and other privacy regulations.
UNIMORE	No the data is fully anonymized and compliant with GDPR
TIM	
NET-INTRA	Not decided yet - due to the project timeline
ITML CY	No. Always in compliance with GDPR.
UNIPI	TBD. In any case the data management process will comply with GDPR and other privacy regulations.
UNP	
KUL	
HPE	
ECL	N/A
F6S	The data produced and used by the NOUS project will be largely discoverable with metadata, identifiable and in some cases, indexable/findable using a persistent and unique actor key. The project will evaluate using the unique actor key across the multiple platforms that will be used in the delivery of the project to enhance findability.
PPC	Data referring to customers will be GDPR compliant, and we will apply seudo-anonymization of sensitive information (e.g., serial numbers, unique numbers, public IP addresses).
ARCTUR	
AEGIS	TBD
CS GROUP - FRANCE	N/A

INESC TEC	INESC TEC datasets that are made publicly available through the institutional repository are minted with a DOI and citation information generated in DataCite. Publication of data is supported by the INESC TEC data steward. The DOI will be created for our datasets even if the datasets are made available in services other than the institutional repository.

PARTNER	**Q2.2: Will rich metadata be provided to allow discovery? If yes, what metadata standards will be followed?**
AIR	
AETHON	TBD. Research from legal and privacy tasks will be utilized.
USAL	
NCSRD	
ICCOR	
POLITO	Yes rich metadata will be provided. Possible metadata standard will be (the list is not exahustive , it is only a snapshot) : CWM : for Datawarehousing, (The purpose of the CWM metamodel is to enable easy exchange of warehouse metadata and business intelligence in distributed heterogeneous environments. ISO 19115: Geographic data, (The ISO 19115:2003 Geographic information — Metadata standard defines how to describe geographical information and associated services, including contents, spatial-temporal purchases, data quality, access and rights to use.). SAML : Shibboleth has been evolved by Internet2/MACE. It provides a method of distributed authentication and authorization for participating HTTP(S) based applications. (Security Assertion Markup Language is an XML-based open standard data format for exchanging authentication and authorization data between parties. As example schema can be used the Advancing open standards for the information society - OASIS.
UNIMORE	Data formart will be custom, but easily interpretable
TIM	
NET-INTRA	Not decided yet - due to the project timeline
ITML CY	TBD

UNIPI	TBD. Research from legal and privacy tasks will be utilized.
UNP	
KUL	
HPE	
ECL	N/A
F6S	TBD. In case rich metadata is used, all metadata standards will be applied.
PPC	TBD
ARCTUR	
AEGIS	TBD
CS GROUP - FRANCE	N/A
INESC TEC	By norm, INESC TEC requires minimal metadata, as part of a dataset deposit form, to be filled in by the data creators. This minimal metadata include administrative, descriptive, semantic, temporal and spatial information (if necessary), based on the Dublin Core standard. Customizable metadata fields may be created if required during the deposit of datasets, mediated by the INESC TEC data steward.

PARTNER	**Q2.3: Will search keywords be provided in the metadata to optimize discovery and potential re-use?**
AIR	
AETHON	TBD
USAL	
NCSRD	
ICCOR	

POLITO	The (meta)data will be recorded and indexed in a long-storage device allowing them to be searched and easily retrieved. by means of search engines, publicly available which will involve consistent ranking trees and keywords and key-concepts and the long-term data preservation strategy as one of the objective of NOUS Cloud.
UNIMORE	TBD
TIM	
NET-INTRA	Not decided yet - due to the project timeline
ITML CY	TBD
UNIPI	TBD
UNP	
KUL	
HPE	
ECL	N/A
F6S	Yes
PPC	Yes, TBD according the datasets provided.
ARCTUR	
AEGIS	TBD
CS GROUP - FRANCE	N/A
INESC TEC	Yes, all datasets will be made available with detailed subject metadata to enable FAIR.

PARTNER	Q2.4: Will metadata be indexable?
AIR	
AETHON	TBD
USAL	
NCSRD	
ICCOR	
POLITO	To be FAIR the NOUS data must be easy to find or retrieve by everyone, humans and machines. To do this, the (meta)data which will accompany NOUS data will have a unique and durable identifier. Data used in NOUS will be described with rich metadata including in particular and explicitly an identifier of the data they describe, unique and durable.
UNIMORE	TBD
TIM	
NET-INTRA	Not decided yet - due to the project timeline
ITML CY	TBD
UNIPI	TBD
UNP	
KUL	
HPE	
ECL	N/A
F6S	Yes.
PPC	TBD
ARCTUR	
AEGIS	TBD

CS GROUP - FRANCE	N/A
INESC TEC	Yes.

PARTNER	**Q2.5: Will the data be deposited in a trusted repository?**
AIR	The data will be stored in a SharePoint repository designated for the project, which is managed by our organization and created/maintained by USAL. SharePoint is a recognized platform that offers robust security controls and compliance, ensuring data integrity and protection. However, it is a private system rather than a traditional public repository, so access is restricted to authorized users by invitation.
AETHON	YES
USAL	While the SharePoint server we maintain is secure, it is an internal infrastructure designed for controlled access by AIR and does not constitute an openly accessible repository.
NCSRD	
ICCOR	
POLITO	Yes we confirm that data will be stored in trusted reppsitort
UNIMORE	TBD
TIM	
NET-INTRA	Data will be hosted in HETZNER (GDPR compliant & DPA with NET-INTRA will be in place)
ITML CY	YES
UNIPI	The data will be stored in a SharePoint repository designated for the project, which is managed by our organization and created/maintained by USAL. SharePoint is a recognized platform that offers robust security controls and compliance, ensuring data integrity and protection. However, it is a private system rather than a traditional public repository, so access is restricted to authorized users by invitation.
UNP	
KUL	

HPE	
ECL	The code developed by the partners in the project as open source will be hosted at Gitlab, a public repository, under a Eclipse instance called the Eclipse Research Labs
F6S	Yes
PPC	PPC's private cloud repository
ARCTUR	
AEGIS	Yes
CS GROUP - FRANCE	N/A
INESC TEC	INESC TEC has an institutional data repository registered on Re3data.org. Moreover, a INESC TEC community is available on Zenodo. The approach is flexible, and other services may be considered for data publication in order to comply with dataset requirements. The publication in other services can be a complement to the availability of data in the institutional repository.

PARTNER	**Will all data be made openly available, or are there restrictions? Explain the reasons.**
AIR	Not all data will be openly accessible due to confidentiality policies and non-disclosure agreements protecting the information shared within the consortium. Restrictions are designed to safeguard partners' intellectual property rights and the privacy of sensitive information. Authorized users can access the data through a secure SharePoint platform, upon invitation, ensuring that only parties with relevant permissions have access to relevant information.
AETHON	Use case partners data will be protected.For the rest, TBD.
USAL	Data in SharePoint is subject to access restrictions established by AIR. USAL only ensures that permissions and access are configured correctly according to the coordinator's directives.
NCSRD	
ICCOR	
POLITO	NOUS data will not necessarily open data. They will in all cases be recoverable by their identifier using a standard communication protocol (open, free, and of universal use), and in all cases the related metadata will be available under known conditions

	thanks to clear licenses (e.g. Creative Commons), and clearly visible; if an authentication and authorization procedure protocol is necessary (e.g.: precise identification of the person consulting, passage by a committee for granting consultation rights) this condition must also be clearly visible.
UNIMORE	TBD
TIM	
NET-INTRA	Not decided yet - due to the project timeline
ITML CY	Itml will make data as openly available as possible
UNIPI	TBD
UNP	
KUL	
HPE	
ECL	When it comes to the open-source code, it will be openly available
F6S	Not all the data gathered throughout the NOUS project will be made available. Public data will be in accordance with GDPR regulations, NOUS' Data Management Plan and NOUS' Grant Agreement.
PPC	Access to data provided by PPC must be limited to partners that participate in the pilots of PPC and relevant tasks, on a need-to-know basis. Bilateral agreements may be addressed depending on Fair and Reasonable conditions (to be specified).
ARCTUR	
AEGIS	TBD
CS GROUP - FRANCE	Data will most probably not be made openly available. Data generated is in relation with our use case (crisis management) and is not meant to be shared.
INESC TEC	An open as possible approach will be adopted. However, there are possible restrictions to be addressed to share Human-centered AI data.

PARTNER	**Q2.6: Will the data be made freely available in the public domain? What licensing will be used?**

AIR	Given the nature of the project and confidentiality restrictions, the data will not be published in the public domain. Access will be limited to consortium members and specific authorized parties under controlled conditions. Regarding the license, internal project data will not be subject to an open-use license but will be governed by internal agreements specifying terms of use and distribution among partners. This allows for effective management of usage and distribution rights, tailored to the project's specific needs and restrictions.
AETHON	TBD
USAL	No, access to the data is limited to authorized partners. USAL does not determine licensing policies.
NCSRD	TBD
ICCOR	
POLITO	The (meta)data will be recorded and indexed in a long-storage device allowing them to be searched and easily retrieved. by means of search engines, publicly available which will involve consistent ranking trees and keywords and key-concept andthe long-term data preservation strategy as one of the objective of NOUS Cloud.
UNIMORE	NO
TIM	
NET-INTRA	Not decided yet - due to the project timeline
ITML CY	Where possible, codebases can be available on public domains e.g. github & open sources ML frameworks
UNIPI	TBD
UNP	
KUL	
HPE	
ECL	An open-source license will be applied, normally a weak copyleft or a permissive license. The specific license will be agreed as part of the project, but normally would be MIT, EPL, APACHE2.0, etc.
F6S	Public data will be in accordance with GDPR regulations, NOUS’ Data Management Plan and NOUS’ Grant Agreement.

PPC	The results obtained by processing the data provided by PPC must be exploited jointly among the partners that participated in the joint effort. Bilateral agreements may be provisioned to cover specifics of that exploitation with Fair and Reasonable conditions (to be specified).
ARCTUR	
AEGIS	TBD
CS GROUP - FRANCE	No
INESC TEC	As long as no ethical or other requirements are identified, the approach to be adopted, according to INESC TEC's FAIR practices, is to make the data available under Creative Commons Attribution-ShareAlike 4.0 International.

11.1.3 OTHER RESEARCH OUTPUTS

PARTNER	**Q3.1: How will the management of research outputs be considered and planned for, both digital and physical?**
AIR	
AETHON	Will be defined in every development plan of each task.
USAL	N/A
NCSRD	Research outputs will be published in conference and journals (preferibly open access)
ICCOR	
POLITO	To be defined when more information on technology structure of NOUS will be clear ; e.g: research outputs will be published in conference and reviews and other available media.
UNIMORE	Research outputs will be published in conference and journals (preferibly open access)
TIM	
NET-INTRA	Not applicable
ITML CY	Research outputs will be published in conference and journals (preferibly open access)

UNIPI	TBD
UNP	
KUL	
HPE	
ECL	
F6S	Each research output should have an exploitation plan that defines its management in digital and physical terms.
PPC	For publications IEEE DataPort, Zenodo
ARCTUR	
AEGIS	TBD
CS GROUP - FRANCE	Research outputs are usually considered for exploitation in other project or product integration depending on the achieved TRL.
INESC TEC	INESC TEC has a Zenodo community for the purpose of sharing other research outputs.

PARTNER	**Q3.2: How will these outputs be managed and shared or made available for re-use?**
AIR	Research outputs will be open if possible. Without compromising exploitable opportunities.
AETHON	Research outputs will be open if possible. Without compromising exploitable opportunities.
USAL	N/A
NCSRD	
ICCOR	
POLITO	To be defined when more information on technology structure of NOUS will be clear ; in any case research outputs will be open . opportunities

UNIMORE	Research outputs will be open if possible. Without compromising exploitable opportunities.
TIM	
NET-INTRA	Not applicable
ITML CY	Where possible, codebases can be available on public domains e.g. github & open sources ML frameworks
UNIPI	TBD
UNP	
KUL	
HPE	
ECL	Research outputs will be open if possible. Without compromising exploitable opportunities. Open-source code developed by the partners in the project will be made available under an open source license, which will be agreed by the project partners
F6S	The generated research outputs will be made available according to their dissemination level, defined in the Grant Agreement and deeper analyzed in the exploitation deliverable for the NOUS project.
PPC	Before making any dataset public, security measures will take place to avoid data leaks and possible loss of interest for the data owner and the consortium. In particular, the dataset will be shared only via the project's official communication channels, i.e., the NOUS SharePoint.
ARCTUR	
AEGIS	TBD
CS GROUP - FRANCE	N/A
INESC TEC	Researchers are responsible for making the deposit. The designated data curator of the community has to approve the deposit. In each deposit, Funding Award information has to be filled in, with reference to the Project (when it applies).

11.1.4 ALLOCATION OF RESOURCES

PARTNER	Q4.1: Who will be responsible for data management in your entity?
AIR	Enrique Mesonero and Alberto Montero
AETHON	Alex Papacharalampous
USAL	David Sánchez Sánchez
NCSRD	Panagiotis Krokidas and Christoforos Rekatsinas
ICCOR	
POLITO	Marco Ghirardi
UNIMORE	Marco Mamei
TIM	
NET-INTRA	Law office “PISTIOLIS – TRIANTAFYLLOS & ASSOCIATES”, Contact Person: Nikolaos Zelios, E-mail: privacy@netcompany.com
ITML CY	Dimitra Siaili
UNIPI	UNIPI DPO
UNP	
KUL	
HPE	
ECL	David Remon
F6S	Each partner in the NOUS project is responsible for the application of the DMP for the data it contributes to the project repository and data that is received, stored, modified, or deleted on a platform the project uses that is such partner’s responsibility. On behalf of F6S, Sofiya Savova will be responsible for data management.
PPC	PPC's corresponding department
ARCTUR	DMO
AEGIS	Evangelos Raptis

CS GROUP - FRANCE	DMO
INESC TEC	INESC TEC has a data steward providing FAIR support for I&D projects. A responsible party for daily management of data may be defined at project level, or the data creators are responsible for the management of their data.

PARTNER	**Q4.2: How will long-term preservation be ensured?**
AIR	As agreed in the conditions as a European project, the documentation of this project must be preserved.
AETHON	With sustainable, reliable storage solutions with regular backups, migration strategies for data stored on obsolete technologies.
USAL	As long as this project lasts.
NCSRD	
ICCOR	
POLITO	Using the best technologies provided by the (Long term) Data Management Industry
UNIMORE	We do not have plans for long term preservation of data (other than pushing data to external services - e.g., github, nature sdata)
TIM	
NET-INTRA	NET-INTRA follows the contractual obligations that emerge from the GA and it will retain the data only for the project purpose
ITML CY	Stored in trusted repository until the end of the project
UNIPI	Depends on the data. Details still to be defined
UNP	
KUL	
HPE	

ECL	Data will be stored in an public repository, Gitlab, ensuring the long-term preservation
F6S	Long-term preservation will be ensured through reliable storage platforms, regular backup and a sustainable attitude towards data management for the NOUS project.
PPC	PPC's private cloud repository
ARCTUR	Stored on our systems, systems are well-secured. We are operating under the auspieces of ISO27001 Certificate.
AEGIS	Stored in trusted repository.
CS GROUP - FRANCE	Stored on our systems with regular backup.
INESC TEC	Datasets deposited in the INESC TEC data repository are preserved indefinitely.

11.1.5 ETHICS

PARTNER	**Q5.1: Are there any ethics or legal issues that can impact data sharing?**
AIR	
AETHON	Everything will comply with the General Data Protection Regulation (GDPR) for personal data, as well as other relevant laws and ethical guidelines specific to the nature of the data
USAL	
NCSRD	
ICCOR	
POLITO	The answer will be best defined when the corpus of data will be clear and close to definitive set in order to verify which ethic issues could arose ; In any case, it will be ensured that all stakeholders will befully informed about the extent and purpose of data sharing and long-term preservation, and consent will be obtained in a way to be compliant with standards.
UNIMORE	

TIM	
NET-INTRA	n/a
ITML CY	no
UNIPI	
UNP	
KUL	
HPE	
ECL	The use of an open source license allows the project consortium to manage legal issues related to the code developed in the project which is oepn to re-use
F6S	While no ethics issues apply to the NOUS project, as stated in the Description of Action and the Grant Agreement, ethics will be taken into consideration in the way data is collected, stored and regarding who can visualize and use it. The project will work to ensure that the management of personal data is compliant with GDPR and other applicable legal frameworks related to personal data protection. During the project, each partner will consider the standards, treaties and laws regarding data protection and privacy in both EU and national level legislation.
PPC	No
ARCTUR	NO
AEGIS	No
CS GROUP - FRANCE	No
INESC TEC	There may be ethical constraints with Human-Centred AI data that will require an assessment of issues regarding data sharing.

PARTNER	**Q5.2: Will informed consent for data sharing and long-term preservation be included?**
AIR	Yes, informed consent will be an integral part of our data management process. Before collecting and using any data within the project, we will ensure that all consortium partners are fully informed and agree to the data sharing and long-term preservation policies.

AETHON	In any case, it will be ensured that all participants are fully informed about the extent and purpose of data sharing and long-term preservation, and consent will be obtained in a manner that is transparent and compliant with legal standards.
USAL	USAL only ensures the technical infrastructure.
NCSRD	Yes according to AI Act
ICCOR	
POLITO	Yes if necessary; In any case, and with a certain extent it will be ensured that all stakeholders will be fully informed about the aim of data sharing and long-term storage, and agreement will be asked in a manner that is transparent and compliant with standards.
UNIMORE	The data will not be refereed to individuals
TIM	
NET-INTRA	n/a
ITML CY	n/a
UNIPI	
UNP	
KUL	
HPE	
ECL	Yes, if necessary. Regarding open source developed code, the developer needs to sign an agreement
F6S	Yes
PPC	Yes
ARCTUR	YES
AEGIS	Yes
CS GROUP - FRANCE	Usually done at project level

INESC TEC	The Data Protection Officer at INESC TEC provides all the necessary support to deal with ethical and legal issues, with already available templates, as is the case with informed consent.

11.1.6 DATA SECURITY

PARTNER	**What provisions for data security are in place, including data recovery, secure storage/archiving, and transfer of sensitive data?**
AIR	In our project, data security measures include regular internal backups and secure storage in a SharePoint managed by USAL, which is only accessed by invitation. Additionally, transfers of sensitive data are conducted through encrypted channels, ensuring the integrity and confidentiality of the information
AETHON	Regular backups, distributed storages for recovery, and secure protocols for data transfer, while access controls, encryption at rest, and ongoing security training ensure the safeguarding of sensitive information
USAL	USAL implements robust security measures, such as regular backups of the SharePoint server, encrypted storage, and secure data transfer protocols. We ensure that the platform is secure and that only authorized users have appropriate access.
NCSRD	Regular backups, use of secure protocols for data transfer, access controls, encryption. Data will be stored in the Zenodo repository, which is an open repository, for EU-funded projects
ICCOR	
POLITO	The data will be under the **Cybersecurity Act strengthens of the EU Agency for cybersecurity (ENISA)** which establishes a cybersecurity certification framework for products and services. in particular for the next future Data will be under the future (Amendment 18/04/23) adoption of European certification schemes for 'managed security services' covering areas such as incident response, penetration testing, security audits and consultancy. Certification will be the key to ensure high level of quality and reliability of these highly critical and sensitive cybersecurity services which will assist companies and organisations to prevent, detect, respond to or recover from incidents.
UNIMORE	Regular backups, use of secure protocols for data transfer, access controls, encryption.
TIM	

NET-INTRA	Backups for data recovery, end-to-end encryption on transfer, about storage to be discussed, authentication and authorization of access
ITML CY	System authentication/authorization & backups for data recovery, use of secure protocols for data transfer
UNIPI	Regular backups, use of secure protocols for data transfer, access controls, encryption.
UNP	
KUL	
HPE	
ECL	The data security provisions in the public repository are provided by Gitlab, going from access control to encryption and 2 Factor Authentication.
F6S	F6S is aware of its data security obligations and responsibilities. There is a dedicated focus on privacy by design and the rights of individuals engaged in the project. F6S will take responsibility for the platforms which it contracts and/or operates for the project. The project will maintain records on each platform that each organization carries responsibility for under the project. Each partner will review each of the platforms they use to deliver the project with regard to data security, data encryption, data retention, secure access, secure transfer, and the security of storage.
PPC	PPC's private cloud repository, ISO IEC 27001, ISO/IEC 27010:2015
ARCTUR	We are operating under the auspices of ISO27001 Certificate and are obligated to operate under our security policies.
AEGIS	Regular backups, use of secure protocols for data transfer, access controls, encryption.
CS GROUP - FRANCE	Regular backups, use of secure protocols for data transfer, access controls, encryption.
INESC TEC	Data is stored in the institutional network drive (Nextcloud) and published in the institutional data repository (when possible). In both services, short-term backups are performed, and data can also be stored on a long-term encrypted tape archive. Nextcloud is hosted on a datacenter at INESC TEC with biometric access control, restricted to IT staff. The drive denys unauthorized persons access to data, prevents unauthorized reading, copying or removal of data. Traceability is ensured via activity logs. The data repository is openly available, but some datasets may be deposited only in private mode, for data preservation purposes.