This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement No 101120823

MANiBOT

Advancing the physical intelligence and performance of roBOTs towards human-like bi-manual objects MANipulation

D1.4. Data Management Plan v1

WP number and title	WP1 – Project Management
Lead Beneficiary	AUTH
Contributor(s)	All Partners
Deliverable type	DMP
Planned delivery date	30/04/24
Last Update	30/04/24
Dissemination level	PUB

Disclaimer

This document reflects only the author's view. Responsibility for the information and views expressed therein lies entirely with the authors. The European Commission is not responsible for any use that may be made of the information it contains. The MANiBOT Consortium consists of the following partners:

Participant No	Participant organisation name	Short Name	Country
1	ETHNIKO KENTRO EREVNAS KAI TECHNOLOGIKIS ANAPTYXIS	CERTH	GR
2	ARISTOTELIO PANEPISTIMIO THESSALONIKIS	AUTH	GR
3	TECHNISCHE UNIVERSITAT DARMSTADT	TUDa	DE
4	SCUOLA SUPERIORE DI STUDI UNIVERSITARI E DI PERFEZIONAMENTO S ANNA	SSSA	IT
5	UNIVERSIDAD DE BURGOS	UBU	ES
6	TECHNISCHE UNIVERSITAET WIEN	TUW	AT
7	TWI ELLAS ASTIKI MI KERDOSKOPIKI ETAIREIA	THL	GR
8	CENTRALNY INSTYTUT OCHRONY PRACY - PANSTWOWY INSTYTUT BADAWCZY	CIOP	PL
9	ASEA BROWN BOVERI SA	ABB	ES
10	FRAPORT ETAIRIA DIACHEIRISIS TON PERIFEREIAKON AERODROMION TIS ELLADAS ANONYMI ETAIREIA	FG	GR
11	SCHWARZ DIGITAL GMBH & CO. KG	SDI	DE
12	DIAMANTIS MASOUTIS AE SUPER MARKET	MASOUTIS	GR
13	UNIVERSITY OF BRISTOL	UoB	UK

Document Authors

Partner	Author
AUTH	Zoe Doulgeri

Document History

Version	Date	Status	Description
0.1	20/03/24	Draft	Table of Contents
0.2	06/04/24	Draft	First Draft Initial Version
0.3	17/04/24	Draft	First Draft Complete
0.4	26/04/24	Draft	Addressed the internal reviewers' comments
1.0	30/04/24	Final	Ready for submission

Definitions, Acronyms and Abbreviations

Acronyms and Abbreviations	Description
DMP	Data Management Plan
EU	European Union
VAT	Value Added Tax
M	Month
D	Deliverable
FAIR	Findable, Accessible, Interoperable, Reusable
EC	European Commission
HE	Horizon Europe
CE	Conformité Européenne
DoA	Description of Action

Executive Summary

MANiBOT seeks to revolutionise the robotics landscape by enhancing robot's handling skills through the development of novel perception, manipulation and cognition methods while focusing on the needs for baggage handling at airports and shelves' restocking at supermarkets. Towards this goal, data have a crucial role. They will be gathered and used for the understanding of the users' requirements so that the project successfully addresses their needs, the efficient training of the methods that will be developed as well as for disseminating the project.

The present deliverable deals with the strategy that MANiBOT project will follow to exploit its data towards the Open Science framework. It outlines the manner the data will be handled during the implementation of the project and after its completion. The role of data manager is dedicated to AUTH, responsible to provide the guidelines for collecting managing, and sharing the research data, as well as for the day-to-day quality assessment. All partners will have an active role for the data collections and preservation to be available and to be cited in publications in order to follow Open Access principles.

Table of Contents

1 Introduction

1.1 Scope of the deliverable

Horizon Europe established the creation of a Data Management Plan (DMP) as a mandatory process in all research projects. Data Management and the procedures regarding privacy, data protection, and ethical challenges are an integral part of the research process for MANiBOT and this document outlines the manner the data will be handled during the implementation of the project and after its completion. For the purpose of drawing up our DMP, we are using the recommended template provided by the European Commission (EC) in order to manage the process and to arrive at a robust plan.

The purpose of this early version of the plan (by Month 6) is to provide a data plan to guide the MANiBOT project partners in relation to the creation, storage, sharing, discovery and destruction of data created over the course of the project. It provides guidance to the project partners and delivers processes and procedures for handling and managing data and refers to current legislative requirements and best practice guidelines, including those relating to GDPR compliance.

1.1.1 Intended readership

The present DMP report is a public (PU) document. Its readership is considered to be the European Commission, the MANiBOT Project Officer, the partners involved in the MANiBOT Consortium, beneficiaries of other European funded projects, and the general public.

1.2 Relation to other Activities and Deliverables

The current deliverable will be used as a point of reference for MANiBOT partners, the components they are developing, and the data they produce and use throughout the project. This report aims to support the work included in MANiBOT by standardizing the exchange of data and additionally facilitate future research activities after the completion of the project. As a result, this deliverable is related with most of the forthcoming MANiBOT deliverables, which involve the use of data for experimentation, exploitation or dissemination reasons. The deliverable is also highly related to D1.2 “Data Management Plan & Ethics” (M6) and its revision D1.3 “Data Management Plan & Ethics-v2” (M42) that act complementary to the DMP in the sense that they refer to specific details of the project (e.g. specific datasets, consent forms) that may contain sensitive information. The DMP’s revisions D1.5 and D1.6 are planned for M24 and M42.

1.3 Structure of the deliverable

The deliverable is structured as reported below:

Chapter 1 – Introduction – Provides the purpose and scope of this document, as well as its connection to other project deliverables.

Chapter 2 – Data Summary – Introduces the reader to the types of data used in this project, the conventions used, categorizes them in terms of access and offers information on their utility.

Chapter 3 – FAIR Data – Explains the project data management under the FAIR data guiding principles.

Chapter 4 – Other Research Outputs – Mentions the remaining project research outputs

Chapter 5 – Allocation of Resources – Examines the effort allocated on this task.

Chapter 6 – Data Security – Offers information on the security aspects related to MANiBOT data.

Chapter 7 – Ethics – Considers the ethical aspects of data collection and management.

Chapter 8 – Other Issues – Reports other issues connected to data management.

2 Data Summary

MANiBOT will mainly rely on newly produced data. The generation and collection of data is a crucial dimension of the execution of the project in order to meet the objectives in terms of machine learning performance, as the data processing outputs will have to be carefully collected and shared within the consortium. The scientific research data will then be easily discoverable, accessible, assessable and intelligible and useable beyond the original purpose for it which it was collected and interoperable to specific quality standards. To define the data categories expected to obtain throughout the MANiBOT project, it is necessary to understand the modalities of data which are explained consequently:

Data
"Data" is a set of values or observations that can be processed, analyzed, and used to draw conclusions or support decision-making. It can come in many forms, such as numbers, text or images. As for the MANiBOT project, data are all the information and knowledge that is collected and possibly created in the scope of MANiBOT. Concerning the more technical data, a neutral format will be used that is not tied to any specialized software to be handled. That involves both the technical data and the user-provided data that will be addressed appropriately, with full respect to EU regulations.

Dataset
A dataset is a collection of data that is organized and formatted in a specific way, so that it can be processed and analyzed. Datasets can come in various forms, such as tables, spreadsheets, or databases.
This task that spans the whole duration of the project will oversee: (1) the compliance of MANiBOT data collection and usage to user privacy regulations, (2) the preservation of the confidentially of personal data, (3) the achievement of a high-level security regarding potential risks.

Metadata
Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. It can include descriptive information such as title, abstract, author, and keywords, as well as administrative information such as creation date, file type, and size.

2.1 Types, origin and formats of generated data

The data has two main origins: the state of the art and the project itself. At the early stages of the project, MANiBOT will reuse available and existing data from the scientific and industrial community to facilitate testing and decision making on the adoption of different, competing technologies. On the other hand, during the main research and innovation activities of the project it will generate data around the development of its objective, such as image sequences, annotations on the content of images, text documents, spreadsheets, laboratory notebooks, diaries, reports and similar.
Personal Data (PD): partners 'personnel involved in the project, EC officers, external people key stakeholders in the industry (attendants to project events, contacts in dissemination and communication events, experts).
Apart from this, the MANiBOT project will generate, collect and process multiple types of data that are shortly categorized as follows:

Financial and administrative data (FAD): partners 'data (beneficiary name, address, VAT number), costs, effort.

Research Data (RD): Data resulting from scientific research activities, including, among others, designs, industrial specifications, process parameters, testing results, materials and samples characterization results, monitoring data or performance data. Parts of these data will be made available and fully open to use by anyone. The software to process each of the above-mentioned data will not be made available, as the intellectual property will remain to the partner that has developed it.

Experiment Data (ExD): Real data taken from experiments conducted in the supermarket and airport test sites.

Exploitation Data (ED): markets characterization data (sizes, trends, competitors), standards, patents, existing technologies, regulations, industrial partners 'commercial goals and strategies, prices, investments, sales projections. Events and publications database.

Dissemination Data (DD): all material that is allowed to be shared on the MANiBOT website https://manibot-project.eu/ , social media (Twitter / X: https://twitter.com/MANiBOT_project LinkedIn: https://www.linkedin.com/showcase/manibot Youtube: https://www.youtube.com/@MANiBOTProject) and at virtual and in person events and exhibition as virtual and printed material, including videos.

Data Format. The formats used for the EU reporting, dissemination and communication data will be assessed according to the current needs, the long and short-term analysis, the storage and the corresponding sharing purposes. Further to the above, for scientific data the following data formats are MANiBOT's preferred choices to enable sharing and long-term validity of the data:

- **JSON**: JSON is a simple file format that is very easy for any programming language to read. Its simplicity means that it is generally easier for computers to process than others, such as XML.

- XML: XML is a widely-used format for data exchange because it gives good opportunities to keep the structure in the data and the way files are built on and allows developers to write parts of the documentation in with the data without interfering with the reading of them. Among others, MANiBOT aims to use XML as a data type for exporting and making available the content of the Database managing the data collected throughout the project's lifecycle. This will be verified in the coming months, and if not possible, a different approach will be followed.
- **JPEG & PNG images**: types of raster image file format that uses compression to store data, and has the file extension: .jpg; .png

- **ENVI files**: The ENVI image format is a flat-binary raster file with an accompanying ASCII header file where the data are stored as a binary stream of bytes.

- **H5 files**: The Hierarchical Data Format version 5 (HDF5), is an open-source file format that supports large, complex, heterogeneous data. This uses a 'directory-like' structure to store computational models and is frequently used in machine learning applications.
- **Spreadsheets**: Many authorities have information left in the spreadsheet, for example Microsoft Excel. This data can often be used immediately with the correct descriptions of what the different columns mean. However, in some cases there can be macros and formulas in spreadsheets, which may be somewhat more cumbersome to handle. It is therefore advisable to document such calculations next to the spreadsheet, since it is generally more accessible for users to read.

- **Comma Separated Values (CSV):** CSV files can be a very useful format because it is compact and thus suitable to transfer large sets of data with the same structure. However, the format is so spartan that data are often useless without documentation since it can be almost impossible to guess the significance of the different columns. It is therefore particularly important for the comma-separated formats that documentation or meta-information of the individual fields is provided and is sufficient and accurate.

- **Text Documents**: Classic documents in formats like RTF, ODF, OOXML, or PDF are sufficient to show certain kinds of documents, e.g., deliverables, reports, etc. Templates may be used whenever possible, so that displayed data can be re-used.

- **Plain Text (TXT):** Plain text documents (.txt) are chosen because they are very easy to read and process via plain text parsers. They generally exclude structural metadata.

- **HTML**: Nowadays much data is available in HTML format on various sites. This may well be sufficient if the data is very stable and limited in scope. In some cases, it could be preferable to have data in a form

easier to download and manipulate, but as it is cheap and easy to refer to a page on a website, it might be a good starting point in the display of data. Typically, it would be most appropriate to use tables in HTML documents to hold data, and then it is important that the various data fields are displayed and are given IDs which make it easy to find and manipulate data.

- **Web Services**: For data that changes frequently, and where each pull is of limited size, it is very relevant to expose data through web services. There are several ways to create a web service, but some of the most used are SOAP and REST.
- **Proprietary formats**: Some dedicated systems, etc. have their own data formats that they can save or export data in. It can sometimes be enough to expose data in such a format – especially if it is expected that further use would be in a similar system as that which they come from.

2.2 Naming Convention

The data generated and collected in MANiBOT as a result of a specific activity will be compiled in a Dataset that should be internally uniquely identifiable. Dataset naming includes a two-digit identifier, followed by the acronym of the consortium partner responsible for the collection of said dataset, the type of collected data as well as its name, followed by another two-digit number identifying the version of the dataset. Hence, the general form would be DSXX.PARTNERNAME.DATATYPE.NAME.VXX. For example:

DS01.AUTH.HUMAN_DEMONSTRATIONS.BIMANUAL_DEMONSTRATIONS.V01

MANiBOT will provide descriptions of the collected and generated data, to facilitate their identifiability and management. The data descriptions will specify the data name, purpose, type, format, origin, size, utility, privacy principles, metadata, accessibility, storage locations and responsible person.

2.3 Categorization of data access

Considering the openness level, each produced dataset will be categorized as follows:
Confidential (CO): Access to these data will be limited to the Consortium members and the EC. In this category fall datasets regarding sensitive information. In particular, what will be excluded from data dissemination is:
i confidential data, especially the one in preparation for future protection by patent or constituent to a secret know how,
ii personal data,
iii classified data, or data made confidential by law,
iv more broadly any data whose disclosure could jeopardize the strategy (including the promotion) or the economic intelligence policy of an organization.
Restricted (RE): If needed for research or exploitation purposes, access can be granted to identified external entities only if confidentiality obligations, similar to those assumed by the partners and the EC, are in force for such entities (Consortium Agreement, section 10). Legitimate interests of the partners will be considered before granting such access.
Open (OP): data under this category will be freely and publicly available by making them available through an open repository and/or within a publication.
Other (OTH): If a partner wishes to have a different type of openness level for a specific dataset, it will be indicated in that particular dataset description.

2.4 Expected size of the data

The size of the data produced during the project should not be more than 750GB, as it will mainly be combination data of images, texts and numeric i.e., textual interaction data, game data, categories etc. Nevertheless, the indicated size is our current rough estimate and could be modified on the final DMP to be delivered at M42 of the project.

2.5 Data Utility

First, the data will be used by the consortium members to achieve the objectives of the project and then a complementary role of the data will be to support future or related research activities. The consortium members are encouraged to maximize data re-use through permissive licenses that allow others to freely access, mine, exploit, reproduce and disseminate previously collected and generated data. They are also encouraged to make the data freely available, except for restricted exceptions, and to use public repositories. It is expected that the data will be used by the scientific community and industries working in the field.
Regarding the use of the data by parties outside the project consortium particular attention will be paid to the data sharing of confidential information and the IPR agreements included in the Grant Agreement Article 16. Depending on the category of data type (see section 2.3), certain data will be open for public use instantly or available after an embargo period.

3 FAIR data

According to the EC rules and templates, FAIR is a set of guiding principles to make data Findable, Accessible, Interoperable, and Reusable. MANiBOT project is completely compatible with these guidelines and this chapter describes these rules.

3.1 Making data findable, including provisions for metadata

Data produced will be easily discoverable with bibliographic metadata including the terms 'European Union' and 'Horizon Europe', using the number of the grant or acronym of the project. Parts of the collected datasets will be published as open data in relevant open repositories where a dedicated project space will be established, e.g. OpenAIRE's Zenodo (see section 3.2). Provisions and actions that are to be taken to ensure the discoverability of MANiBOT data include:

- Accompanying datasets with properly structured and accurate metadata.
- Providing proper documentation identifying their content and potential uses.
- Making data identifiable by using standard identification mechanisms and persistent and unique identifiers (e.g. Digital Object Identifiers (DOI)), where applicable.
- Advertising them in global search engines.
- Publishing papers and reports with references to them.
- Ensuring accessibility of the hosting infrastructure (at least 99.9%).

Furthermore, the data will be accompanied with metadata, providing context and additional information about the datasets. These provisions will make the data more findable and will increase the impact of MANiBOT. All metadata will be documented and organized in order to improve data accessibility. The planned metadata will include information on:

- The context of data collection: aims, objectives and hypotheses.
- Data collection methods: data collection protocol, sampling design, instruments, hardware and software used, data scale and resolution, temporal coverage and geographic coverage.
- Dataset structure of data files, cases, relationships between files.
- Data sources used.
- Data validation, checking, proofing, cleaning and other quality assurance procedures carried out.
- Modifications made to data over time since their original creation and identification of different versions of datasets.
- Information on data confidentiality, access and use conditions, where applicable.

At data-level, datasets should also be documented with:

- Names, labels and descriptions for variables, records and their values.
- Explanation of codes and classification schemes used.
- Codes of, and reasons for, missing values (if any).
- Derived data created after collection, with a description of the algorithm used to create them.
- Weighting and grossing variables created data listing with descriptions for cases, individuals or items studied.

Moreover, all publications coming from MANiBOT project will bring the following statement: 'This work has received funding from the European Union's Horizon Framework Programme for Research and Innovation under grant agreement no 101120823'.

3.2 Making data accessible

Three main repositories are used in MANiBOT for data storage. These repositories are the MANiBOT website, the Google Drive cloud service for storage of high-volume data, the NextCloud based project repository and the Zenodo platform for reaching users outside the project. Additional internal communication tools and channels that might be used by the consortium partners are independent and are not considered as repositories of the project.

3.2.1 MANiBOT Website

The MANiBOT website (https://manibot-project.eu/) will be a channel for the communication of project achievements and results. It will contain information about the project, including public project details, news, events, dissemination and communication materials (e.g., flyers, posters, papers) as well as public deliverables. The website will also provide access to several public project datasets that will be made available in the Zenodo repository.

3.2.2 Google Drive

Google Drive is a cloud-based storage and collaboration service developed by Google. It allows users to store and share files such as documents, images, and videos, and access them from any device with an internet connection. To comply with the GDPR, Google has implemented various measures such as providing users with specific data protection terms and privacy policy, and offering tools for data processing agreements, data protection impact assessments, and breach notification. Google has provided users with tools to manage their data, including the ability to delete or export their data from Google Drive. Users also have the right to access, correct, and delete their personal data in accordance with the GDPR. Overall, Google Drive is subject to the GDPR's strict data protection requirements, and users can rest assured that their personal data is being handled in compliance with the regulation.

MANiBOT partners can store large datasets to Google cloud repositories. This will be a safe, password protected service that will be accessible only by authorized members of the consortium. The Google Drive is aimed at large files and datasets that are cumbersome to be stored in the NextCloud repository (see below). These data will not contain sensitive information or any kind of personal data. All partner will have their own local folders for sharing their data and will be personally responsible for their management. The datasets that will be available for use after the end of the project will have been registered in Zenodo (see below).

3.2.3 NextCloud project repository

NextCloud is a self-hosted open-source platform for file sharing and collaboration. It allows users to store, share, and collaborate on files securely and easily, with control over their data. NextCloud is designed to be GDPR compliant, which means it provides a range of features and functionalities that enable users to manage their data in accordance with the European Union's General Data Protection Regulation (GDPR), as summarized below:

- Data protection and encryption: NextCloud offers robust encryption features that ensure data privacy and security. All data stored on NextCloud is encrypted, both in transit and at rest, which means that it cannot be accessed by unauthorized individuals.
- Data retention and deletion: NextCloud allows users to control how long their data is retained and when it is deleted. This feature ensures that personal data is not kept for longer than necessary, reducing the risk of data breaches.
- Data portability and access: NextCloud enables users to export and transfer their data easily, giving them greater control over their personal information. Users can also access their data from anywhere, using any device, as long as they have an internet connection.

- Privacy by design: NextCloud is designed with privacy in mind, ensuring that all data is stored securely and that the platform complies with GDPR regulations.

Overall, NextCloud is a powerful platform that offers a range of features that support MANiBOT in achieving GDPR compliance, providing users with greater control over their personal information and ensuring that the stored data is protected.

CERTH has provided a cloud repository where all project related communication files and supporting materials will be accumulated to facilitate collaboration among partners. This repository is a safe, password protected repository that will only have internal use between the authenticated and authorized MANiBOT members. Therefore, the repository can be used by all partners to store and manage data. The MANiBOT repository is managed by CERTH and will be available during and five years after the completion of the project.

3.2.4 Zenodo

MANiBOT will also provide publicly available datasets, using the Zenodo platform (https://zenodo.org/), which has been widely used by EU projects for data sharing. Zenodo is a general-purpose open-access repository operated by CERN and funded within OpenAIRE project. Zenodo is an easy to use and innovative service that enables researchers, EU projects and research institutions to share and showcase multidisciplinary research results (data and publications) that are not part of existing institutional or subject based repositories. Namely, Zenodo enables users to:

- Easily share the long tail of small data sets in a wide variety of formats, including text, spreadsheets, audio, video, and images across all fields of science.
- Display and curate research results, get credited by making the research results citable, and integrate them into existing reporting lines to funding agencies like the EC.
- Easily access and reuse shared research results.
- Define the different licenses and access levels that will be provided.

Furthermore, Zenodo assigns a Digital Object Identifier (DOI) to all publicly available uploads, in order to make content easily and uniquely citable. This repository also makes use of the OAIPMH protocol (Open Archives Initiative Protocol for Metadata Harvesting) to facilitate the content search through the use of defined metadata. It uses a JSON schema as the internal representation of metadata and offers export to other formats.
In addition, the short- and long-term storage of the research data will be secured since they are stored safely in the cloud infrastructure of CERN, which uses digital preservation strategies to storage multiple online replicas and to back up the files (Data files and metadata are bac ked up on a nightly basis).
Therefore, this repository fulfils the main requirements imposed by the EC for data sharing, archiving and preservation of the data generated in HE projects.

3.2.5 Other platforms

Furthermore, MANiBOT data will be made available in trusted online repository such as:

- **Grant Management Services Platform** is managed by the European Commission. In this site, administrative and financial data will be shared with the EC.
- **Social media** tools will be used to advertise and disseminate the project's results to the open public and the presence of MANiBOT and its partners at public exhibitions and events.

3.3 Making data interoperable

Interoperability refers to open access data and thus, the data MANiBOT project is going to deliver and share through the previously indicated (see section 3.2.4) open-source other repositories. The data produced in the project are interoperable because they will be used in two different activities, training machine learning models and defining the project user requirements.

MANiBOT activities assume the creation of metadata to facilitate data exchange and re-use and the consortium is committed to follow the Data Management Plan to enhance the interoperability of data. The guidelines provide the principles and requirements for technical and quality management system documentation in the regulatory context of CE marking the MANiBOT multicomponent system, including product qualification and risk classification. To ensure the interoperability of published datasets, the project will design structure and publish the data using digital data models based on widely used standards and datatypes as indicated in section 2.1.

The metadata produced should use vocabularies that follow FAIR principles and for certain terms reference will be made to open, external vocabularies, e.g., license (Open Definition) and grants (OpenAIRE). Additionally, metadata will include qualified references to other metadata, and each referred external piece of metadata is qualified by a resolvable URL.

3.4 Increase data re-use

MANiBOT has developed a website that reports and results decided in the GA can be easily accessible by the public. Other repositories and platforms will also link to the project as it continues to develop, which will be referenced in the project website. Reports that bring the PUBLIC type label, publications and exploitation activities generated by the MANiBOT project will be unlimited to share, thus a CC BY 4.0 license is sufficient. This license lets others distribute, remix, adapt, and build upon all materials, even commercially, as long as they credit the project for the original creation. Furthermore, the coordination of the project will decide the way of re-usage of the data generated in the project. Definitions and methods will be applied in accordance with the data presentation and thus file headers, comments and/or data sheets will be implemented.

4 Other research outputs

Many of software packages stemming from the MANiBOT research efforts will be available for public usage after the completion of the project. Additionally, pre-trained deep neural network models will be made available for luggage as well as supermarket items detection. A wide variety of sensors is going to be developed and utilized by the Consortium partners in both research efforts, as well as other projects. The final integrated robotic platform will be also available for further exploitation. Depending on the necessities of each work package and the consequent workplan, additional software maybe necessary and it would be updated in the final version of the DMP.

All data collected and processed are in accordance with Data Protection Legislation and comply with Grant Agreement regulations.

5 Allocation of resources

Data management plan and processes will be under the supervision of Zoe Doulgeri (AUTH) as leading member of the AUTH team in the MANiBOT project and is responsible of the DMP task. Additionally, future Zenodo repository cost could be associated which is not applicable currently as it is a costless repository. Furthermore, there is no cost devoted to data management plan or any other costs concerning the management tools. Besides, there is no license fee for data management software. The data will be preserved for the timeframe of four years after the end of the project, as stated in the DoA.

6 Data security

Since data are stored in a shared cloud repository, the NextCloud platform will guarantee backups of data. Furthermore, the Zenodo repository is a dedicated tool creating regular backup of the data.
MANiBOT partners will ensure data security following the below listed guidelines: Provide access to authorized personnel only, store data in at least two separate locations to avoid data loss, limit the use of USB flash drives, label files in a systematic way in order to ensure consistency of the final dataset. Long-term data preservation will be ensured by partner institution's data repositories.
The MANiBOT website has implemented a cookie policy to be compliant to the present GDPR regulations. The statistics gathered from the social media is typically giving personal information useful for market analysis (e. employer company) but is already anonymized by the platforms themselves.

7 Ethics

Any ethical or legal issues that can have an impact on data sharing can be discussed in the context of the ethics review. The MANiBOT partners are to comply with the ethical principles as set out in Article 14 the Grant Agreement, which, among other, states that all activities must be carried out in line with the highest ethical standards and the applicable EU, international and national law on ethical principles and the beneficiaries must commit to and ensure the respect of basic EU values (such as respect for human dignity, freedom, democracy, equality, the rule of law and human rights, including the rights of minorities).

Following the above it is necessary to clearly state that the innovations to be developed in the project regard object manipulation by robotic systems. As such it will not process sensitive or any other kind of human-related data. This greatly minimizes the ethical issues that may be relevant to the project and the collection/storage of data. The only kind of personal data collected by this project regard the user requirement definition, which will be kept to the minimum required for the project.

To proceed with the development of AI methods, as well as for the definition of user requirements, human participants (volunteers) will be asked to produce demonstration datasets of object manipulation primitives, and to complete questionnaires. MANiBOT aims to ask the minimum possible information from the recruited volunteers. In particular, the demonstration datasets will not collect or use any user personal information. For the case of use case definition, demographic data will be collected from participants in the FG and MASOUTIS premises, including age, gender, educational level, position within the company and years of professional expertise, The collected data will pass through an embargo period for validation of no issues in data collection, and will be pre-processed to make them fully anonymous before being used in the project.

8 Other issues

No other issues to report at this time.