PILLAR ROBOTS

Purposeful Intrinsically motivated
Lifelong Learning Autonomous Robots

D12.2
Data Management Plan

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement no. 101070381.

1. Change Control

1.1. Document Properties

Name deliverable	**Data Management Plan**
Deliverable No.	**12.2**
Work Package title	**Management**
Author/s	**BOTIFOLL, Lluís**
Contributor/s	**-**
Reviewer/	**DURO, Richard**
Editor/s	**DURO, Richard**
Dissemination Level	**PU**

1.2. Revision History

Version	Date	Change	Modified by
0.1	01-10-2022	Draft Table of Contents based on the "Guide for Researchers. How to comply with Horizon Europe mandate for Research Data Management"	Ll. Botifoll
0.2	06-03-2023	First complete draft	Ll. Botifoll
0.3	15-03-2023	Final full draft submitted for internal review	R. Duro
0.4	21-03-2023	Comments and contributions received by other partners	-
0.5	25-03-2023	External version approved	R. Duro
1.0	31-03-2023	Final submission to EC	R. Duro

2. Contents

3. Executive summary

Deliverable D12.2 'Data Management Plan' describes the way in which data collection and processing will be managed within the project. PILLAR-Robots will collect and must subsequently manage large amounts of data.

This document presents a summary of all initial datasets identified, the protocols that will be followed to open / public data to ensure data as Findable, Accessible, Interoperable and Reusable (FAIR). It also discusses about the allocation of resources and costs associated to the maintenance of the infrastructure involved in data management. It goes further to explain the security features related to handling of data which will be implemented and refers to the compliance of the data management in PILLAR-Robots with the European ethical standards.

It is intended for internal use only and it will be mainly employed by project participants responsible for, or in any way involved in, the data collection and data handling processes.

Furthermore, this data management plan will be updated as the project progresses, since not all data or potential uses are clear from the beginning. New versions of the DMP will be created whenever significant changes occur in the project such as the inclusion of new data sets, modifications in consortium policies or external factors.

4. List of Abbreviations

Abbreviation	Explanation
CA	Consortium Agreement
CC	Creative Commons
DMP	Data Management Plan
FAIR	Findable, Accessible, Interoperable, Reusable
GA	Grant Agreement
GDPR	General Data Protection Regulation
IPR	Intellectual Property Rights
SEP	Standard Ethics Protocol

5. Introduction

This deliverable is inspired in the Horizon Europe DMP template[1]. The contents of the document are developed based on several crucial documents such as the Description of Action (DoA) and the deliverable D12.1 Project Quality Handbook.

1 The current Programme Guide includes a Horizon Europe DMP template (as of May 5, 2021). https://www.openaire.eu/images/Guides/HORIZON_EUROPE_Data-Management-Plan-Template.pdf

This DMP describes the data management life cycle for all data to be collected, processed and/or generated by the project. The DMP has the following structure:

- **Data Summary** – Presents the purpose of data within project, data types and formats, the data sources, also presenting the storage platforms that will be used by the Consortium;
- **FAIR Data** – Details the protocols that will be applied to open / public data to ensure data as Findable, Accessible, Interoperable and Reusable (FAIR);
- **Allocation of resources** – Refers to the costs involved in maintaining the infrastructure and the resources involved in data management;
- **Data security** - Describes the integrated security features related to the handling (storage, transfer and recovery) of data and the Data Manager responsibilities;
- **Ethics** – Details the compliance of the data management with European ethical standards and directives in protecting personal data and data security.

6. Data summary

Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.

The re-use of data has been considered but discarded because of the non-existence of adapted and available data to the needs of the project. No re-use of any existing data is foreseen at the present stage.

What types and formats of data will the project generate or re-use?

Data collection activities in PILLAR-Robots are still being planned; thus, a complete list of planned datasets is not yet available. The DMP will be periodically updated as data collection plans are refined.

To develop of all the project modules and AI-based characteristics, labelled and unlabelled data is needed for training purposes. For the proper training and validation of the autonomous prototypes, realistic datasets are essential. The project consortium will use and generate the following datasets:

1. Experimental data on performance. Information related to demonstrations activities:
    a. Audios of natural language;
    b. Images of gazes;
    c. Images of pointing gestures;
    d. Written language;
2. Questionnaires; personal and industrial data;
3. Content on webpages, statistics from Google. Analytics and email addresses from newsletter subscribers.
4. Scientific articles elaborated by the project partners.

The format of data would follow industry standards as much as possible.

The PILLAR-Robots project will use widely accepted formats for data generation, such as text documents, spreadsheets, databases, audio files, pictures, and videos, however, some datasets may include different data formats. Details regarding the size of the datasets are not yet known but they will be updated by the end of the project.

An internal document will compile an overview of the type of data collected for each Work Package and relevant tasks and this will serve as an input for the update of the DMP to be reported in D12.4 'Data Management Activity Report'. We will take into consideration the FAIR principle, which will be addressed with a level of detail appropriate to the project.

What is the purpose of the data generation or re-use and its relation to the objectives of the project?

PILLAR-Robots will create an integrated repository to store and manage the data collected from the different sources generated. The overall purpose for data collection is to help the project to develop a solution to operationalize the concept of Purpose within open-ended learning autonomous robots. In particular, the data generated is crucial for the success of the testing and validation project phases.

What is the expected size of the data that you intend to generate or re-use?

Data gathered by the partners and its activities should represent approx. 25 TB during the project.

What is the origin/provenance of the data, either generated or re-used?

Data will be acquired mainly during the tests and simulations in the lab (WP6), in the field validation (WP7-9) and stakeholders (WP10-11). Data from standardized training sets and other public repositories related to the topics of the project may be used for internal lab testing of algorithms. However, no data is collected from other sources.

A sample for the preliminary dataset structure is presented below:

Type	Publication date	DD.MM.YY
	Data source	
	Author (anonymized)	Generated ID
	Message	Text
	Language	
	Method of data extraction	Keyword/s
	Link to the original source	
	Type of content	

To whom might your data be useful ('data utility'), outside your project?

Other researchers in similar open-ended learning autonomous robots projects.

7. FAIR data

7.1. Making data findable, including provisions for metadata

PILLAR-Robots project aims to adhere to the protocols to ensure FAIR data to make the project data Findable, Accessible, Interoperable and Reusable (FAIR). However, the project will also generate datasets, or parts of datasets, that cannot be shared due to certain restriction regarding data privacy sensitivity.

Will data be identified by a persistent identifier? Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how. Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use? Will metadata be offered in such a way that it can be harvested and indexed?

Data will be categorized according to type and purpose according to each partner. All partners will document their data in a different way (using dedicated software, IP management systems) Data will be stored in line with standards of good scientific work and ethical guidelines, and according to the relevant project task and will be archived with a clearly identifiable filename. Metadata will also be included where appropriate. Datasets will be made available as follows:

1. REQUIRED: An archive of the data itself.
2. REQUIRED: A description of what the data is and how to use it.
3. OPTIONAL (AS REQUIRED): Instruction of use, including any code and how to use it.

The description of the procedures to generate data is associated to a dataset (i.e., collection of data).

At this stage of the project, the specific typology and total number of variables in a single dataset table cannot be defined a-priori. The procedures for the identification of data are defined as follows:

- Each dataset is initially assigned to a unique ID, automatically generated through a Universally Unique Identifier (UUID) application;
- Each dataset is also associated to a Digital Object Identifier (DOI). The service is provided by the DOI[2] community through a request to a local Registration Agency (RA).

The use of a DOI guarantees, at the same time, unique identification of the single dataset and the possibility of automatic data web retrieval. After this step, the dataset is univocally associated to an identifier. The implementation of the data description depends on the typology of datum considered. In most cases, a text description is appropriate. In this case, data are described by compiling a form (data description template), available to all users.

[2] https://www.doi.org

Naming conventions for datasets generated in the PILLAR-Robots project have not been set so far.

The dataset names should offer information regarding:

- Dataset (DS), (clearly showing what is being uploaded is a dataset);
- The source of the data;
- The Data Controller (i.e., name of the partner responsible for the dataset);
- A brief description of the content;
- The data number generated by the research metadata list in TEAMS.

The exact naming conventions and associated explanations will be detailed in D12.4 'Data Management Activity Report'.

All datasets that will be uploaded to Zenodo or equivalent platform will be provided with DOI versioning. This allows to revise papers and datasets while adhering to data citation principles (allows editing and updating of already published datasets and makes it possible to cite a specific version of an upload).

No relevant metadata standards available in the project field. One or more metadata files will be generated for each dataset. The metadata are identified by the same unique ID of the related dataset, with a different suffix/extension. Each metadata file will be uploaded in a standardized format, depending on the dataset considered. Appropriate templates will be available for download to all partners in the collaborative platform.

Keywords are provided in the metadata to optimize the possibility for discovery and re-use. The project will define a limited list of general keywords that should apply to all public datasets, as well as public deliverables and scientific publications. Such a list of general keywords will be provided in future revisions of this document.

7.2. Making data accessible

Repository: Will the data be deposited in a trusted repository? Have you explored appropriate arrangements with the identified repository where your data will be deposited? Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?

Three repositories will be employed:

A. Internal tool for online collaboration and file sharing: TEAMS

PILLAR-Robots project uses TEAMS and Microsoft Cloud as the project internal tool for online collaboration and file sharing (details can be found in D12.1 'Project Quality Handbook'). The access

to the TEAMS platform is limited to project partners, the access being granted based on an invitation-only procedure.

TEAMS uses a two-factor authentication system. This prevents credentials from being used without a second factor and mitigates the impact of compromised passwords. The cloud also applies in-transit encryption and does not authenticate connections over HTTP but redirects to HTTPS. At rest, data is encrypted at the disk level using BitLocker encryption and at the file level using keys.

B. Zenodo Repository

In PILLAR-Robots we will employ the catch-all research data repository Zenodo. The Zenodo repository will be used to make the research findable in accordance with the Open Access mandate[3]. Uploads are made easily and uniquely citeable, as Zenodo assigns every published record a Digital Object Identifier (DOI). The DOI is a top-level and mandatory field in the metadata. Datasets published in Zenodo will be registered with the following default metadata:

- DOI;
- Version number;
- Bibliographic information;
- Keywords;
- Abstract/description;
- Associated project and community;
- Associated publications and reports;
- Grant information;
- Access and licensing info;
- Language.

C. GibHub Repository.

A GitHub repository will be used for hosting the source code of the PILLAR-Robots architecture and modules. Following the git protocol, the code will be hosted in a secure manner in terms of availability and access. The repository will be structured in branches and each branch will store historical data, ensuring thus the availability of the code. Access to the repository will be achieved using GitHub accounts and two-factor authentication. The technical teams will receive an invitation by email to join the project repository. Other features covered by the GitHub are related to:

- Context switching using branches.
- Role-based code lines using dedicated branches for specific environments and specific scenarios or end-users.
- Feature-based workflow using a branch for each feature and thus achieving an elevated level of granularization of the codebase.
- Disposable experimentation by using dedicated branches for experiments, which can later be deleted.

[3] how-to-comply-with-horizon-europe-mandate-for-publications (openaire.eu)

Data: Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement. If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible. Will the data be accessible through a free and standardized access protocol? If there are restrictions on use, how will access be provided to the data, both during and after the end of the project? How will the identity of the person accessing the data be ascertained? Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?

Generated data will be publicly available at the end of the project, so that they can be re-used by interested parties. Data generated will be multimodal (synchronized / correlated); this would be particularly attractive to external researchers.

By default, and as a first case-study of data management, only data related to publications will be made openly available in general, the Management Committee / General Assembly will decide on a case-by-case basis which data can be released to avoid issues related to IPR protection or access. All data and metadata files will be uploaded onto a cloud storage and sharing facility specifically dedicated to PILLAR-Robots project.

Metadata:
Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data? EU Grants: Data Management Template (HE):V1.0 – 05.05.2021 5 How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available? Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)?

Public metadata and datasets will be made available to users once publications are finalized through the PILLAR-Robots' website and through the OpenAIRE sharing web platform. In particular, relevant metadata and dataset will be uploaded by the involved researchers to the Zenodo5 platform, compiling project-related information. This will enable automatic data extraction from the OpenAIRE platform, thus ensuring accessibility through a standard platform for Open Data access.

7.3. Making data interoperable

What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you

follow community-endorsed interoperability best practices? Which ones? In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them? Will your data include qualified references to other data (e.g. other data from your project, or datasets from previous research)?

Since data will be stored in the most common formats, it is reasonable to expect that data could be re-used with a good level of interoperability.

All public data will be made available in standard/open formats compliant with commercial/open software to allow as much as data exchange between researchers and institutions. Standard vocabulary for metadata description will be used in case this will not be possible a mapping of more common ontologies will be provided. In this case, specific technical contribution from specialists in semantics and logics will be considered.

7.4. Increase data re-use

How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)? Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement? Will the data produced in the project be useable by third parties, in particular after the end of the project? Will the provenance of the data be thoroughly documented using the appropriate standards?

The public data will be made available according to Open Licenses such as Creative Common Licences. The data will be available for re-use upon decision of the Management Committee / General Assembly, to avoid issues related to IPR protection or access. Once the data are made openly available, they will remain open. Within the strategy of the development of the DMP, the datasets that will be firstly available are those reported in publications originated from the consortium, thus intrinsically created for being reused. However, specific agreements with the Editors of scientific/technological journals will be considered and provided. The data quality is assured by each partner, having responsibility of the datasets. The tools necessary for describing and identifying the dataset and for preparing the metafiles will be provided by the Management Committee / General Assembly in strict collaboration with the Coordinator and the Innovation Manager. When a public deliverable is approved and accepted by the European Commission, datasets associated with the deliverable will be made available. Public datasets will be available, and thus remain reusable, via Zenodo. All confidential data will be deleted by the end of the project.

Describe all relevant data quality assurance processes. Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects related to the allocation of resources, data security and ethical aspects.

Project beneficiaries will follow the data naming conventions, detailed previously, in addition to following the 5 criteria, detailed below, to ensure data quality:

- Accuracy: stored data should be an accurate reflection of what is described.
- Relevancy: the data should meet the requirements for the intended use.
- Completeness: the data should not have missing values or miss data records.
- Timeliness: the data should be up to date.
- Consistency: the data should be of an expected data format for the data type and be cross referable with the similar results, if any.

8. Allocation of resources

What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.) ? How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)

The PILLAR-Robots project will use Zenodo, a free of change data repository, which ensures the preservation of public data on a long-term basis. The costs of data management activities in PILLAR-Robots are limited to project management costs and will thus be covered by allocated resources in the project budget. Additionally, the GitHub costs are covered from the allocated resources in the project budget.

Who will be responsible for data management in your project?

The Data Manager, Lluís Botifoll (UDC) will have overall responsibility for data management in PILLAR-Robots and is the Data Protection head of the project. The people responsible for Zenodo, TEAMS and Github are as follows:

- Zenodo – Lluís Botifoll (UDC)
- TEAMS – Richard Duro (UDC)
- GitHub – José Antonio Becerra (UDC)

Support is also provided by the Data Protection Officers (DPOs) of the project partners that will handle any personal data within PILLAR-Robots project. The project coordinator Richard Duro (UDC) will support the Data Manager in carrying out his work as needed.

How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?

These aspects will be available to project partners once the project finishes. What will be saved, and its duration, will be reviewed throughout the project, and updated accordingly in D12.4 'Data management activity report' to be delivered in M45.

9. Data security

What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)? Will the data be safely stored in trusted repositories for long term preservation and curation?

Data security must be set in relation to each of the repositories that will be used.

A. PILLAR-Robots TEAMS

A private working space has been established for the project on UDC's TEAMS platform. The PILLAR-Robots TEAMS is a secure access point and it is accessible only by the personnel employed by partners directly and specifically involved in the project. The user management will be done by UDC. A dedicated folder for research datasets will be set up, thus allowing for even stricter access control if needed. The following security settings apply for the PILLAR-Robots TEAMS platform:

- Secure storage: Documents and elements are stored in Microsoft's TEAMS solution;
- Secure access: Restricted to project members (specific persons) only. Further restrictions on specific folders can be imposed on a person-by-person basis as needed;
- Secure transfer: Data is protected during transfer between partners and TEAMS platform through encryption with SSL/TLS;
- Secure manipulation of data: Possible manipulation of data is registered and/or prevented through threat management, security monitoring and file/data integrity;
- Secure Backup: daily automatic backup is done, also on Microsoft's TEAMS solution.
- All project data will be stored for 5 years, unless otherwise agreed in contracts and data processing agreements.

B. Zenodo Repository

Zenodo has specified the following policies related to longevity:

- Versions: Data files are versioned. Records are not versioned. The uploaded data is archived as a Submission Information Package. Derivatives of data files are generated, but original content is never modified. Records can be retracted from public view; however, the data files and records are preserved;
- Replicas: All data files are stored in the CERN Data Centres, primarily Geneva, with replicas in Budapest. Data files are kept in multiple replicas in a distributed file system, which is backed up to tape on a nightly basis;
- Retention period: Items will be retained for the lifetime of the repository. The host laboratory of Zenodo CERN, has defined a lifetime for the repository of the next 20 years minimum;
- Functional preservation: Zenodo makes no promises of usability and understandability of deposited objects over time;
- File preservation: Data files and metadata are backed up nightly and replicated into multiple copies in the online system;
- Fixity and authenticity: All data files are stored along with an MD5 checksum of the file content. Files are regularly checked against their checksums to assure that file content remains constant;

- Succession plans: In case of closure of the repository, a guarantee has been made from Zenodo to migrate all content to suitable alternative institutional and/or subject based repositories.

C. GitHub Repository

GitHub data security will be assured by the GitHub's integrated security policies:

- Two factor authentication for all the invited team members of technical partners.
- Dependable alerts and security updates provided by GitHub after scanning the libraries used in the codebase.
- Dependency graphs will provide a better understanding of all the dependencies of the codebase.
- Repository history using branches, commits, pull requests and tags ensure a high-level of availability.

10. Ethical aspects

Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA). Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data?

PILLAR-Robots contains three main areas for ethics consideration, which have been identified during the proposal preparation, the review process and the grant preparation phase.

10.1. Human Research Ethics

PILLAR-Robots will recruit human participants in order to carry out research activities in WPs 7, 8, 9 and 10. Only adults who are members of stakeholders' organisations with legal decision capacity and autonomy will be allowed to participate in the activities and will have to sign a detailed consent form which will be tailored to their specific needs or conditions if needed. There is an exception in WP8 where the validation will involve several groups of minors. Specific consent informed measures will be adopted in this WP and a particular task of monitoring and reporting will be in place (D8.9: Informed consent requirements report). Furthermore, all partners have a mutual duty of care to each other and to maintain the project's autonomy. They also have a duty of care to participants in ensuring that they are not put at risk of harm as a result of their participation. Informed consent Legal and ethical principles and precepts having to do with informed consent will be examined. Whatever their health status, human participants will be recruited according to the ethical guidelines published by the European Commission regarding Horizon Europe projects, which means that no discrimination based on protected attributes will be conducted.

Participants will be informed about the project, the type of data that will be processed, the way their contributions will be used, and by whom. Participation in the project will be fully voluntary and all participants will be provided with an opportunity to ask questions and receive clear answers prior to making fully informed decisions about their participation. Participants will have the right to withdraw and terminate their participation in the research at any time and will be reminded about this right before participating. The studies will be registered according to relevant national and organisational research ethics regulations.

10.2. Protection of Personal Data: data handling and processing

There are two areas where personal data collection has been considered. These areas are in WP7-9, as well as WP10. Where possible, interviews will be transcribed rather than recorded using audio or visual means. Any data collected will only be used for the purpose for which it was collected and will be stored (as a minimum) on a password protected computer. Access to data will always be restricted to only those people in the project beneficiaries with a direct need to access it for the purposes of achieving the project results. Participants of the surveys will be informed of who shall have access to the results and for what purpose. Personal data produced within the framework of the project will be stored by the partner in charge of the task concerned and not shared to the rest of partners or third parties.

10.3. Artificial intelligence

The project will follow the definition of AI prepared by the EC's High Level Expert Group that is contained in the document "A definition of AI: Main capabilities and disciplines. Definition developed for the purpose of the AI HLEG's deliverables" (published the 8th April 2019). PILLAR-Robots will follow the provisions contained in the "Ethics Guidelines for Trustworthy Artificial Intelligence" and the "The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment" prepared by the EC' High-Level Expert Group on AI that establish a set of requirements/considerations that AI systems should meet in order to be deemed trustworthy. These key requirements/considerations will be monitored together by the project Data Manager (DM), the Ethics Advisor and the Equality & Ethics Committee (EEC) in M15 through D1.0 – OEI - Requirement No. 1' and updated in a final report in D12.3 - Trustworthy AI Assessment) that will be delivered at the end of the project.

The ethical standards, guidelines of Horizon Europe and EU regulations will be rigorously applied, regardless of the country in which the research is carried out. The EU regulations to be applied are:

- Directive (EU) 2016/680[4] on the protection of natural persons with regard to the processing of personal data by competent authorities for the purpose of prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA.
- General Data Protection Regulation (Regulation (EU) 2016/679[5] of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regards to the

[4] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016L0680
[5] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

processing of personal data and on the free movement of such data and repealing Directive 95/46/EC).

- DIRECTIVE 2013/40/EU[6] of the European Parliament and of the Council of 12 August 2013 on attacks against information systems and replacing Council Framework Decision 2005/222/JHA.

Moreover, PILLAR-Robots will follow international standards and requirements for research with human subjects. The Nuremberg Code (1947), the Declaration of Helsinki (1964), and The Belmont Report (1979) provide key references in this regard. The Nuremberg Code underlined the need for guaranteeing and respecting the voluntary nature of human participation in research and pointed out the requirement of establishing mechanisms for Informed Consent, also ensuring people involved in research can withdraw from it at any time. This code also underlined that researcher must ensure the welfare and protect the interests of participants. With this aim in mind, researchers must establish in advance mitigation measures for addressing any risk of harm for them. The Declaration of Helsinki followed the same approach.

Conversely, the Belmont Report developed four key ethical principles to be considered when carrying out research activities that will be respected through PILLAR-Robots activities:

- **respect for people:** research subjects must be treated so as to protect their safety, respect their autonomy and ensure their consent on an informed basis;
- **beneficence:** possible benefits for the participants will be maximised while possible harm or risk will be minimised;
- **justice:** any benefits and burdens derived from research must be balanced;
- **competence**: the limitations and boundaries of the researchers' competence must be recognised and made explicit.

In this regard, this DMP provides preliminary information regarding how to tackle the main ethical implication of data treatment in PILLAR-Robots, which is the reuse of personal data. This process will have to follow principles and requirements stated in the GDPR. Respect for participants and concern for their wellbeing will remain a top priority. Personal information will never be disclosed unless informed consent has been given for disclosure or there is other legal and ethical ground for it. All ethical aspects are covered by ethics deliverables that will be provided following the project deliverable schedule. The DMP will be updated accordingly.

10.4. Information letter and informed consent procedures

Deliverable D12.4 'Data Management Activity Report' will detail the procedures and criteria that will be used to identify/recruit research participants, also will detail the informed consent procedures related to research activities involving personal data collection. The consent form will include information about the type of data that will be collected, how it will be used, stored, transferred, and destroyed. A template of the consent form will be provided.

[6] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013L0040&rid=4

10.5. Alignment with the GDPR

All planned datasets, by default and by design, will comply with the general principles for the processing of personal data as described in the GDPR article 5. This includes data: 1) lawfulness, fairness, and transparency, 2) purpose limitation, 3) minimisation, 4) accuracy, 5) storage limitation and 6) integrity and confidentiality, as well as 7) accountability.

11. Other issues

Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?

We do not make use of other national/funder/sectorial/departmental procedures.

12. Conclusion

Throughout the duration of the PILLAR-Robots project, this DMP will be updated to adapt to any unforeseen circumstances that might arise, and to provide more information on topics such as:

- Datasets (types, formats, sizes);
- Naming conventions;
- Approach to search keywords;
- Alignment of datasets with principles for the processing of personal data in GDPR article 5.

The last updated version of this DMP will constitute D12.4 ‘Data Management Activity Report’ and available at the end of the project.

www.pillar-robots.eu

Funded by the European Union