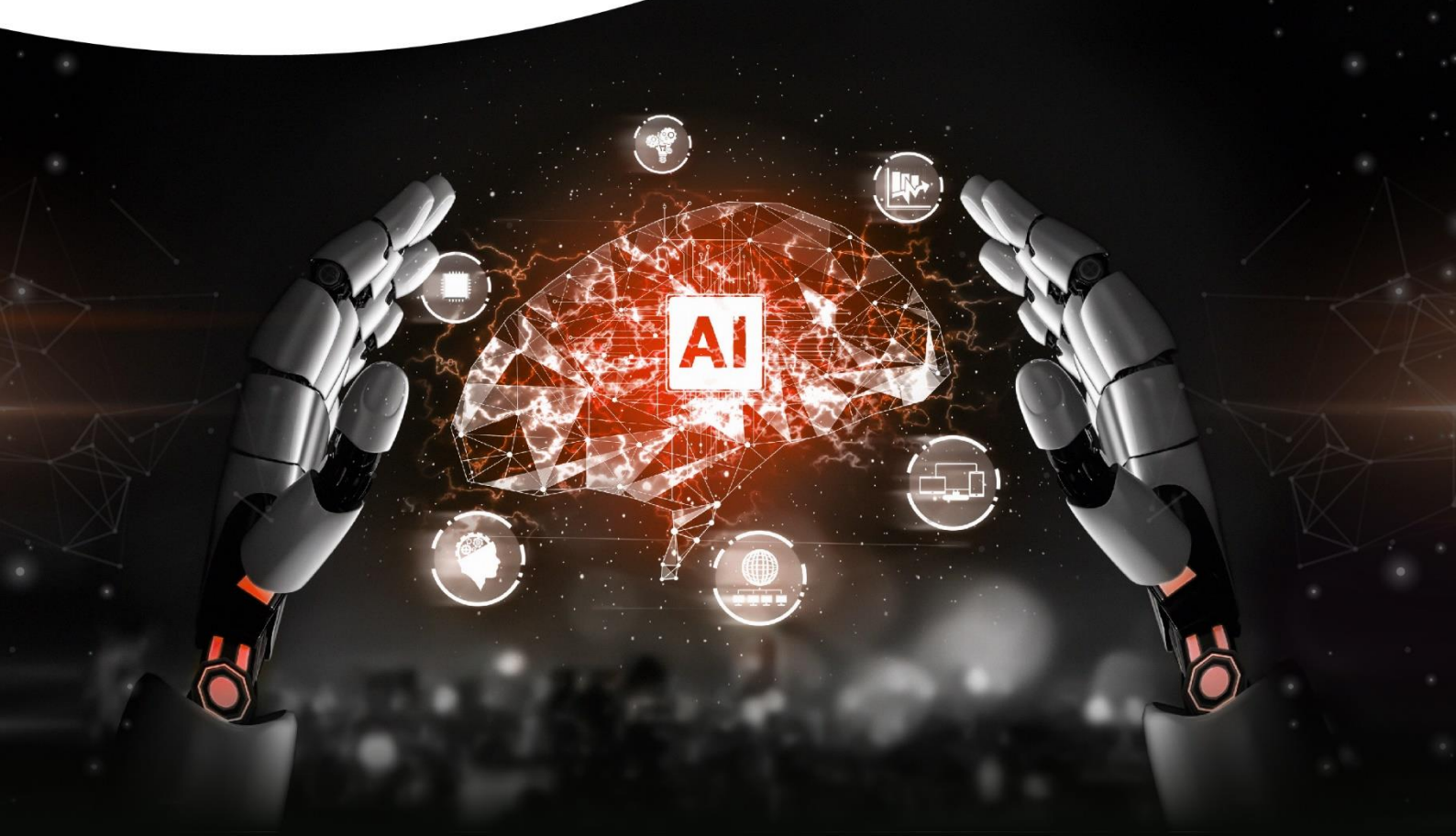
AGIMUS

INNOVATIVE ROBOTICS FOR AGILE PRODUCTION

D8.2 Data Management Plan First Version



PROJECT ACRONYM: AGIMUS

PROGRAMME: Horizon Europe
GRANT AGREEMENT: No 101070165
TYPE OF ACTION: Horizon Research & Innovation Actions
START DATE: 1 October 2022
DURATION: 48 months

Funded by the European Union

Document Information

Issue date: **Issued by:**	30/12/2022 Q-PLAN
Due date: **Work package leader:**	31/12/2022 CNRS
Dissemination level:	PU (Public)

Document History

Version	Date	Modifications made by
0.1	16/11/2022	First draft available by Q-PLAN
0.2	02/12/2022	Input from partners collected
0.3	12/12/2022	Updated draft circulated by Q-PLAN for review
0.4	19/12/2022	Peer reviewed by INRIA and KLEEMANN
1.0	19/12/2022	Final version available by Q-PLAN for submission
1.1	30/12/2022	Final version available using revised template – Q-PLAN

Authors

First Name	Last Name	Beneficiary
Nicolas	Mansard	CNRS
Florent	Lamiraux	CNRS
Justin	Carpentier	INRIA
Vladimir	Petrik	CTU
Alessandro	Di Fava	PAL
Narcis	Miguel Baños	PAL
Pierre	Fernbach	TOWARD
Maximilien	Naveau	TOWARD
Michail	Mitsouridis	Q-PLAN
Vangelis	Genitsaris	Q-PLAN
Apostolos	Tsolakis	Q-PLAN

Theofilos	Mastos	KLEEMANN

In case you want any additional information, or you want to consult with the authors of this document, please send your inquiries to: mitsouridis@gplan-intl.gr

Quality Reviewers

First Name	Last Name	Beneficiary
Pierre-Guillaume	Raverdy	INRIA
Theofilos	Mastos	KLEEMANN

Executive Summary

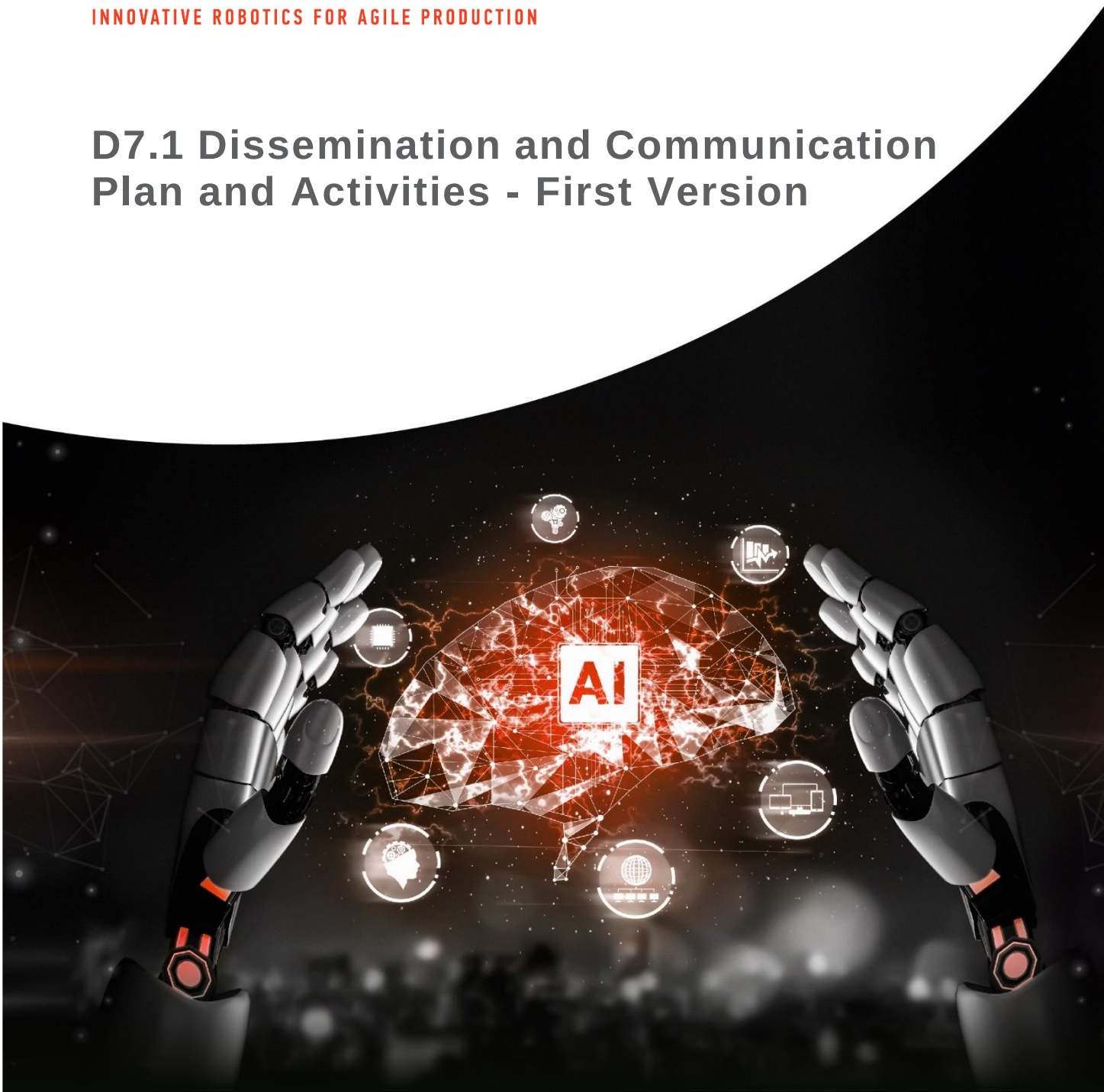
This document constitutes the initial version of the Data Management Plan (DMP) and has been elaborated as a deliverable (D8.2) in the framework of the AGIMUS project.

AGIMUS aims to deliver an open-source breakthrough innovation in AI-powered agile production, introducing solutions that push the limits of perception, planning, and control in robotics, enabling general-purpose robots to be quick to setup, autonomous and to easily adapt to changes in the manufacturing process.

To achieve such agile production, AGIMUS leverages on a combination of (i) an advanced task and motion planner that can learn from online available video demonstrations; (ii) optimal control policies obtained from progresses in trajectory optimization and reinforcement learning rooted on efficient differentiable physics simulations of the manufacturing process; as well as (iii) advanced perception algorithms able to handle objects and situations unseen during initial training. Along the way, developments in torque-control

INNOVATIVE ROBOTICS FOR AGILE PRODUCTION

D7.1 Dissemination and Communication Plan and Activities - First Version



PROJECT ACRONYM: AGIMUS

PROGRAMME: Horizon Europe

hardware design and implementation will support further pushing the limits of autonomy, including specific attention to computational energy efficiency and robot-to-cloud communication.

The AGIMUS solutions and their impact will be demonstrated and thoroughly stress-tested in 3 testing zones, as well as 3 industrial pilots in Europe, under numerous diverse real-world case studies and scenarios (different tools, environments, processes, etc.). In every step, and from the very beginning, AGIMUS will go beyond current norms and involve a wide range of stakeholders, starting from the production line itself, to identify the essential ethical-by-design principles and guidelines that can maximise acceptance and impact.

In this context, the initial version of the project's DMP sets out the overall methodological principles pertaining to the management of the data that will be collected and/or generated in the framework of AGIMUS, safeguarding sound and ethical data management for the entire duration of the project. Moreover, it provides a first, yet still meaningful overview of AGIMUS's data, as identified in this early stage of the project, along with information on the methodology pertaining to their management as well as making them Findable, Accessible, Interoperable and Re-usable (FAIR).

The initial version of the DMP is the first of three versions of the AGIMUS's Data Management Plan to be produced in the course of the project and will serve as living document (D8.2 Data Management Plan – First Version delivered in M3 will be updated to D8.3 Data Management Plan – Second Version in M24 and ultimately fixed as D8.4 Data Management Plan – Final Version in M48). Along these lines, the DMP will be updated and further elaborated during the project to reflect an accurate, up-to-date and ultimately comprehensive plan for managing the data that will be collected, and/or generated by the project across the entire life cycle, both during and after the completion of AGIMUS.

Table of Contents

List of Figures

List of Tables

List of Terms and Definitions

Table 1: Terms and Definitions

Abbreviation	Definition
ADRA	AI, Data and Robotics Association
AI	Artificial Intelligence
CERN	European Organization for Nuclear Research / Conseil Européen pour la Recherche Nucléaire
CESSDA	Consortium of European Social Science Data Archives
DCMI	Dublin Core Metadata Initiative
DMP	Data Management Plan
DOI	Digital Object Identifier
EAB	Ethics Advisory Board
EFFRA	European Factories of the Future Research Association
EU	European Union
FAIR	Findable, Accessible, Interoperable, Re-usable
GDPR	General Data Protection Regulation
IAB	Industrial Advisory Board
KPI	Key Performance Indicator
PID	Persistent Identifier
SME	Small Medium Enterprise
URL	Uniform Resource Locator

1. Introduction

The current document represents the initial version of the Data Management Plan (DMP) of the AGIMUS project which has received funding from the European Union's Horizon Europe Research and Innovation funding programme under Grant Agreement No 101070165.

AGIMUS aims to deliver an open-source breakthrough innovation in AI-powered agile production, introducing solutions that push the limits of perception, planning, and control in robotics, enabling general-purpose robots to be quick to setup, autonomous and to easily adapt to changes in the manufacturing process.

To achieve such agile production, AGIMUS leverages on cutting-edge technologies and goes beyond the state-of-the-art to equip current mobile manipulators with a combination of (i) an advanced task and motion planner that can learn from online available video demonstrations; (ii) optimal control policies obtained from advances in reinforcement learning based on efficient differentiable physics simulations of the manufacturing process; as well as (iii) advanced perception algorithms able to handle objects and situations unseen during initial training. Along the way, optimization of energy efficiency and the use of 5G technology will support further pushing the limits of autonomy.

The AGIMUS solutions and their impact will be demonstrated and thoroughly stress-tested in 3 testing zones, as well as 3 industrial pilots in Europe, under numerous diverse real-world case studies and scenarios (different tools, environments, processes, etc.). In every step, and from the very beginning, AGIMUS will go beyond current norms and involve a wide range of stakeholders, starting from the production line itself, to identify the essential ethical-by-design principles and guidelines that can maximise acceptance and impact.

To this end, the consortium of AGIMUS brings together a complementary and interdisciplinary group of 9 partners across 4 different countries within the EU, as presented in Table 2.

Table 2: AGIMUS Consortium

Partner Role[1]	Partner No	Partner Name	Partner Short Name	Country
CO	1	CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE	CNRS	France
BE	2	CESKE VYSOKE UCENI TECHNICKE V PRAZE	CTU	Czechia
BE	3	INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET AUTOMATIQUE	INRIA	France
BE	4	PAL ROBOTICS SL	PAL	Spain

[1] CO: Coordinator, BE: Beneficiary

Partner Role[1]	Partner No	Partner Name	Partner Short Name	Country
BE	5	TOWARD SAS	TOWARD	France
BE	6	Q-PLAN INTERNATIONAL ADVISORS PC	Q-PLAN	Greece
BE	7	AIRBUS	AIRBUS	France
BE	8	KLEEMANN HELLAS SA	KLEEMANN	Greece
BE	9	THIMM OBALY KS	THIMM	Czechia

All partners of AGIMUS' consortium adhere to sound data management principles in order to ensure that the meaningful data collected, processed and / or generated throughout the duration of the project are well-managed, archived and preserved, in line with the structure and guidelines of the Horizon Europe Data Management Plan Template.

Along these lines, this initial version of the DMP aims to achieve the following objectives:

- Describe the data management lifecycle for the data to be collected and/or generated in the framework of AGIMUS, serving as the key element for good data management.
- Outline the methodology employed to ensure the sound management of the data collected, and/or generated as well as to make them Findable, Accessible, Interoperable and Re-usable (FAIR).
- Provide information on the data that will be collected and/or generated and the way in which they will be handled during and after the end of the project, along with the standards applied to this end.
- Provide details on how the data will be made openly accessible and searchable to interested stakeholders as well as their curation and preservation.
- Present information on the resources to be allocated so as to make data FAIR clearly identifying responsibilities pertaining to data management, while addressing data security and ethical aspects.

With the above in mind, this initial version of **the DMP is structured in 8 distinct chapters**, as follows:

- **Chapter 1** provides introductory information about the DMP, the context in which it has been elaborated as well as its objectives and structure.
- **Chapter 2** presents a summary of the data to be collected/generated during the activities of AGIMUS including the purpose for its collection/generation and its types and formats. Additionally, it outlines its origin, expected volume and the stakeholders that may find them useful.
- **Chapter 3** describes the methodology that is applied in the framework of AGIMUS in order to ensure the effective management of data across the entire lifecycle, making it FAIR.

- **Chapter 4** plans for the FAIR management of other research outputs (digital or physical) that may be generated or re-used by the beneficiaries throughout the project.
- **Chapter 5** estimates the resources required for making the project's data FAIR, while also identifying data management responsibilities.
- **Chapter 6** outlines the data security strategy applied within the context of AGIMUS along with the respective secure storage solutions employed.
- **Chapter 7** addresses ethical aspects as well as other relevant considerations pertaining to the data collected/generated during the implementation of the project.
- **Chapter 8** concludes on the next steps foreseen in the framework of the project with respect to its data management plan.

Annexed in the document are (i) the project's Privacy Policy (Annex I), the templates for the (ii) Informed Consent Form (Annex II), and (iii) the Data Subject Request Form (Annex III).

Note that **the DMP is not a fixed document**. It will evolve during the lifespan of the project and will be **further elaborated and updated at least twice more throughout the duration of AGIMUS** (i.e., as D8.3 in M24 and D8.4 in M48). Additional ad hoc updates may be realised (if necessary), in order to include new data, better detail and/or reflect changes in the methodology or other aspects relevant to their management (such as costs for making data FAIR, size of data, etc.), changes in consortium policies and plans or other potential external factors. Q-PLAN is responsible for the elaboration of the DMP and with the support of all partners will update and enrich it when required.

2. Data Summary

AGIMUS will collect/generate meaningful non-sensitive data that do not fall into any special categories of personal data as described within the General Data Protection Regulation (GDPR). These data may be quantitative, qualitative or a blend of those in nature and will be analysed from a range of methodological perspectives with a view to producing insights that will successfully feed AGIMUS' activities, enable us to deliver evidence-based results and ultimately achieve the objectives of the project. With that in mind, the second chapter of the Data Management Plan (DMP) starts by explaining the purpose for which these data will be collected/generated and how they relate with AGIMUS. It proceeds by describing the different types and formats of these data as well as their origin and expected volume, before concluding with an overview of potential stakeholders for whom they may prove useful for re-use.

2.1 Purpose of data collection / generation and its relation to the objectives of the project

In order to successfully meet its objectives and ensure the production of evidence-based results, AGIMUS entails several activities during which data will be collected/generated. The purpose for which these data are collected/generated is interrelated with the objective of the activity during which they are produced.

In particular, these activities along with their objectives in the framework of AGIMUS are as follows:

1. **Identification of needs, challenges, opportunities, data and platforms for agile production**, to understand the current European landscape, and in more detail for the AGIMUS industrial pilot regions, to ultimately feed into the requirements and specification of the AGIMUS framework. Moreover, activities for retrieving specifications and standards from other related projects, initiatives, and platforms are also foreseen in order to support the development of the AGIMUS framework, as well as the stress-testing and industrial pilot use cases.
2. **Identification of ethical, trustworthy, and secure design guidelines, for next-generation robotics for agile production**, to also feed into the requirements and specification of the AGIMUS framework.
3. **Technical and non-Technical specifications during the co-design (technical) workshops**, to complement the design of the AGIMUS system architecture. A diverse group of key stakeholders will engage in co-designing the building blocks of the AGIMUS architecture. Ideas, relevant insights, and key takeaways will be collected and written up by means of notes and minutes as document files. Data will also be collected on the type of stakeholders engaged to better reflect on background and expertise of involved stakeholders.
4. **Motion dataset and sensor readings collected during development experiments,** including but not limited to videos of robot experiments, video of demonstration motion

by human actors, as well as sensor data from the robot and the surrounding environment.

5. **Object 6D pose estimation dataset for performance tracking and training.** This activity involves the collection of real-life and generation of virtual datasets for the purpose of neural networks training and for evaluation of the proposed improvements, *e.g.*, temporal smoothing or physical consistency.
6. **Specifications and deployment of the stress-testing experiments and benchmarking criteria in the AGIMUS testing zones**, which involves the definition of the implementation concept of industrial mock-up use cases and the generation of valuable data for re-training and improving the AGIMUS framework.
7. **Specifications, deployment, and evaluation of the AGIMUS industrial pilot use cases.** This activity involves the definition of the implementation concept of the industrial pilot use cases as well as the deployment of the AGIMUS framework along the upgraded Tiago2 robot in the three industrial pilots in France, Greece, and Czechia. The purpose of this activity is to test and fine-tune the project's solutions under real-life environments, considering overall performance, autonomy, safety, and energy efficiency requirements, among others.
8. **Business planning exercise** which includes **market analysis** (market size, market structure, prospects, and growth rate, etc.), **competition analysis** (number and profile of competitors, potential gap, etc.), as well as **financial planning** (cash flow analysis, costs and result forecast, etc.). This exercise will fuel the elaboration of the AGIMUS Business Plan which aims to ensure the effective and viable exploitation of the AGIMUS results after the end of the project.
9. **Monitoring and assessment of the dissemination and communication results of the project,** with a view to measuring the impact of the relevant activities, fine-tuning AGIMUS' strategy accordingly, as well as fulfilling its reporting requirements towards the Commission.
10. **Collaboration and Synergies with the ADRA and Related Networks and Initiatives**, in order to coordinate, develop and benefit from synergies with other relevant EU initiatives. This includes euROBIN, ELSA, ADRA-e, as well as other H2020 and Horizon Europe projects.
11. **Setup and Operation of AGIMUS' Advisory Boards** (Industrial & Ethics)**.**

The following section provides further details on the different types and formats of data collected / generated during the project's activities.

2.2 Types and formats of collected / generated data

AGIMUS is set to collect / generate data of various structures and formats. Along these lines, the data definition process used for this DMP is based on the source and the physical format of the data[2]. In particular, we define two main aspects: (i) the process under which the underlying data are created / captured which includes electronic texts documents, spreadsheets, questionnaires, images, videos, code, and transcripts, among others and (ii) the storage format of quantitative and qualitative data. Examples of this aspect include easily accessible formats, such as post scripts (e.g., pdf, xps, etc.), machine readable formats (xml, html, bag, etc.), spreadsheets, (e.g., xlsx, csv, etc.), text documents (e.g. docx, rtf, etc.), images (e.g., jpeg, png, etc.), videos (e.g., mp4, avi, etc.), code (e.g., cpp, py, lib, rosbag, etc.), compressed formats (e.g. rar, zip, etc.) or any other format required by the objectives and methodology of the activity within the framework of which is produced.

Under this framework, special attention will be paid in using open **formats**[3] (such as csv, pdf, zip, etc.) and / or **machine-readable formats**[4] (such as xml, json, rdf, html, bag, py, cpp, etc.) when possible, to enhance the **interoperability and re-use of data**. In doing so, we will be providing data that is **easily readable** and **freely usable in any software program** employed by third-parties interested in utilizing the data.

The type and formats of the data collected / generated in the context of AGIMUS can be divided into 3 categories, namely:

- data collected / generated by **direct input methods**;
- data collected / generated through **the AGIMUS framework, and it's use**; and
- data collected / generated from **dissemination, communication, and stakeholder engagement activities**.

These categories are described in more detail in the following subsections.

2.2.1 Data collected / generated through direct input methods

Direct input methods, under the scope of AGIMUS, involve methodologies for collecting data through desk research and interactions between consortium partners and external stakeholders, with the latter providing data to the former. Along these lines, external stakeholders undertake the role of a data

[2] Jakobsson, U., Braukmann, R., Lundgren M., Expert Tour Guide on Data Management. Retrieved from https://www.cessda.eu/Research-Infrastructure/Training/Expert-Tour-Guide-on-Data-Management/1.-Plan.

[3] According to the Open Data Handbook: "*An open format is a file format with no restrictions, monetary or otherwise, placed upon its use and can be fully processed with at least one free/open-source software tool and it is not encumbered by any copyrights, patents, trademarks or other restrictions so that anyone may use it*".

[4] According to the Open Data Handbook: "*Machine readable formats are file formats that can be automatically read and processed by a computer. Machine-readable data must be structured data*".

subject that is a natural person whose personal data is being processed[5]. In particular, the identification and selection of suitable data subjects are based on purposeful sampling according to which, external stakeholders are identified and selected by consortium partners based on their relevance with the AGIMUS framework (e.g., AI, Robotics, Agile Production, etc.) and the objectives of the respective activity for which data is collected. In this context, quantitative and qualitative data[6] will be collected / generated during AGIMUS:

- **Quantitative data** is numerical and acquired through counting or measuring. Examples of quantitative data are the yearly turnovers of a business, the hourly compensation of a worker, the number of SMEs in Europe, etc. This data may be represented by ordinal, interval or ratio scales and lend themselves to statistical manipulation.
- **Qualitative data,** sometimes referred to as categorical data, is data that can be arranged into categories based on physical traits, gender, colours or anything that does not have a number associated with it. Moreover, written documents, interviews, and various forms of in-field observation are all sources of qualitative data. Examples of qualitative data are the preferences of learning, skillsets, country of origin, etc.

Additional details with respect to the different types and formats of data that will be collected through direct input methods under the frame of AGIMUS are provided below.

Identification of needs, challenges, opportunities, data, and platforms for agile production

The data deriving by this activity fall under Task 1.1 (Q-PLAN) and will include desk research results and opinions of external experts and other European and/or International actors on the needs, challenges, opportunities, data and platforms for agile production. The opinions will originate from semi-structured interviews and online surveys and will include qualitative and quantitative data, stored in documents and spreadsheets (interview transcripts, notes, minutes, questionnaires, survey results) in .docx and .xlsx format.

Identification of ethical, trustworthy, and secure design guidelines, for next-generation robotics for agile production

The data deriving by this activity fall under Task 1.2 (INRIA) and will include opinions of external experts and other European and/or International actors on the guidelines required for delivering ethical, trustworthy, and secure by-design next-generation versatile robotics. Data collected will be aligned with the data from Dataset#1. The opinions will originate from semi-structured interviews and online surveys and will include qualitative and quantitative data, stored in documents and spreadsheets (interview transcripts, notes, minutes, questionnaires, survey results) in .docx and .xlsx format.

[5] Regulation (EU) 2016/679 of the European parliament and of the council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016R0679.

[6] Neuman, W. L. (2014). Social research methods: Qualitative and quantitative approaches. Boston: Pearson.

Technical and non-Technical specifications during the co-design (technical) workshops

The data deriving by this activity fall under Task 1.3 (TOWARD) and refers to data collected through the co-design technical workshops for co-designing the AGIMUS system architecture. The data will include clear requirements, opinions, diagrams, and other visual elements which will be documented in documents and spreadsheets, (notes, minutes, requirements, etc.) in .docx, .xlsx, or .bag format.

2.2.2 Data collected / generated by the AGIMUS framework

This section focuses on the data collected and their formats. Additional documentation, accompanying each data(set), such as a manifest listing each data in the dataset, or metadata to facilitate its usage can be found in Section 3.

Motion Solver Toolkit Benchmark Dataset

The data deriving by this activity fall under WP2 (INRIA) and refer to data collected for creating the baseline values for benchmarking the performance of (i) the efficient differentiable physics simulator, (ii) the trajectory optimiser, and (iii) the task-an-motion planner and optimiser. The data will mainly consist of several “problems requiring solving” and will be available in compressed files in .tar, .gz, or .zip format.

Sensor readings recorded during experiments (real or simulated)

The data deriving by this activity fall under WP3 (CNRS) and refer to data collected from sensors to test software developments. The data will include images taken by the robot camera, point clouds, joint angles and will be available in ROS packages in .rosbag format.

Motion dataset

The data deriving by this activity fall under Task 3.2 (CNRS) and refer to data collected by the motion planner (T2.3 and T3.4) or trajectory optimiser (T2.2) but also the motion dataset module (T3.2). the data will include trajectories, whereas the format will be defined during the implementation stage.

Videos of motions performed by human actors

The data of this dataset will be used to feed Task 3.3 (CNRS) and consist of online publicly available videos of human actors performing a series of actions (human demonstrations). The videos are expected to be in .mp4 format.

Dataset for 6D object pose estimation training and evaluation

The dataset will be used to train and evaluate object 6D pose estimation methods developed in Task 4.1. Data consists of images depicting the object(s) of interest, the corresponding ground truth poses of the objects, the information about temporal sequences of the images, and the geometrical description of the surrounding environment. File formats: .png, .jpg, .yaml, .urdf and .srdf.

Collection of data for learning sensor-to-state predictive model

The sensor readings of the robot joints and force/torque sensors will be collected on a real robot to train a sensor-to-state mapping for the purpose of model predictive control. The raw sensory readings will be stored in rosbag format.

Robot performance benchmarking and stress-testing

Generated data in this dataset derive from Task 5.1 and 5.2 (PAL). In particular, through Task 5.1, a list of goals and corresponding metrics that would allow to assess the successful performance of the components developed in the previous technical WP is expected. Mobile manipulator technologies would also be benchmarked, and the most adequate ones for the project would be integrated in an evolved version of the TIAGo robot. Concerning Task 5.2, on the one hand, these integrations and improvements should be done in the real robot; and on the other hand, the metrics defined in T5.1 should be used to assess the performance of the platform in the three testing zones of AGIMUS. The foreseen tests here are expected to be performed using the data of time series of readings of the sensors, encoders, etc. of the robot. All of these will be gathered in the form of ROS packages (i.e., rosbags in .bag format).

Visual Control Experiment Dataset

The data deriving by this activity fall under Task 5.3 (INRIA). Images and videos of robot with surrounding environment for model predictive control from visual inputs, based on several pick-and-place tasks in all testing zones. The images and videos are expected to be in .jpeg and .mp4 format, respectively.

Physical Interaction Experiment Dataset

The data deriving by this activity fall under Task 5.4 (CTU). The dataset will consist of the 3D mesh models of tools used in the experiment and a set of images and corresponding 6D poses generated for the visual perception training. File formats: .stl, .dae, .yaml, .urdf

Multi-step Task Planning Experiment Dataset

The data deriving by this activity fall under Task 5.4 (CNRS). The dataset will consist of videos of motions performed by human actors, including occluded visual feedback and challenging planning and interaction control scenarios, to evaluation the identification of optimal plans. The videos are expected to be in .mp4 format.

AIRBUS Use Cases Dataset

The data deriving by this activity fall under Task 6.1 and 6.2 (KLEEMANN) and specifically regarding the AIRBUS industrial pilot use cases. The dataset will consist of a set of sensor readings, images, and videos collected during the preparation and deployment of the industrial pilots, to better prepare and evaluate the AGIMUS solutions. The sensor readings are expected to be in .bag format, the images in .jpeg, and the videos in .mp4. Where possible CAD models of the equipment used for the use case(s) will also be included in the dataset, in .dwg or .pdf format.

As a general note, due to strict data privacy regulations in the AIRBUS industrial sites, it will be quite challenging to collect any visual imagery for the project. If this data is collected, it shall be done only after approvals of the associated personnel and only for internal consortium usage. In this event, the approval document and its guidelines for use shall also be reported and added to the AGIMUS database. Alternatively, the “mirror” industrial setup at CNRS will be used to generate the envisioned data.

KLEEMANN Use Cases Dataset

The data deriving by this activity fall under Task 6.1 and 6.2 (KLEEMANN) and specifically regarding the KLEEMANN industrial pilot use cases. The dataset will consist of a set of sensor readings, images, and videos collected during the preparation and deployment of the industrial pilots, to better prepare and evaluate the AGIMUS solutions. The sensor readings are expected to be in .bag format, the images in .jpeg, and the videos in .mp4. Where possible CAD models of the equipment used for the use case(s) will also be included in the dataset, in .dwg or .pdf format.

THIMM Use Cases Dataset

The data deriving by this activity fall under Task 6.1 and 6.2 (KLEEMANN) and specifically regarding the THIMM industrial pilot use cases. The dataset will consist of a set of sensor readings, images, and videos collected during the preparation and deployment of the industrial pilots, to better prepare and evaluate the AGIMUS solutions. The sensor readings are expected to be in .bag format, the images in .jpeg, and the videos in .mp4. Where possible CAD models of the equipment used for the use case(s) will also be included in the dataset, in .dwg or .pdf format.

Cross-pilot Comparative Assessment Dataset

The data deriving by this activity fall under Task 6.3 (AIRBUS) and refers to data collected through the evaluation and assessment activities of the industrial pilot use cases’ deployment. The data will include evaluation results, questionnaires, benchmarks, statistics, and other elements which will be documented in documents and spreadsheets (performance metrics, key insights, etc.) in .docx, .xlsx, or .bag format.

Business planning exercise

The data deriving by this activity fall under Task 7.3 includes market analysis (market size, market structure, prospects, and growth rate, etc.), competition analysis (number and profile of competitors,

potential gap, etc.), as well as financial planning (cash flow analysis, costs and result forecast, etc.). This exercise will fuel the elaboration of the AGIMUS Business Plan which aims to ensure the effective and viable exploitation of the AGIMUS results after the end of the project.

Setup and Operation of AGIMUS' Advisory Boards

The data regarding the setup of the AGIMUS IAB and EAB fall under Task 8.2 (INRIA) and include personal information of candidate AB members (e.g., name and surname, gender, organisation, position, email, country of placement, brief description of professional profile). They will be collected by the consortium partners, under the guidance of INRIA, following the process of retrieving information from online public sources (professional websites, etc). Each partner will be asked to identify at least 3-5 suitable relevant stakeholders from their own network and provide to the consortium information that is publicly available. The candidates will be assessed by a selection process and will be formally approached and invited to give their consent to become members of the AB. The data will be stored in text document, image, and spreadsheet format (.docx, .xlsx, .pdf, .jpeg).

Collaboration and Synergies with the ADRA and Related Networks and Initiatives

The data deriving by this activity fall under Task 7.5 (Q-PLAN) and regard photos, number of event participants and presentations, meeting minutes, past knowledge and good practises deriving from the partners' synergetic actions with sister projects and initiatives (e.g., euROBIN, ELSA, etc.), as well as other relevant EU initiatives and the collaboration with the ADRA. The synergies and how they will be forged to benefit the AGIMUS project activities will be decided, per case, based on discussions with representatives of these projects. The data will be stored in text documents, excel spreadsheets, presentation, and image files (.docx, .xlsx, .pdf, .ppt, .jpeg, .png).

2.2.3 Data collected / generated from dissemination, communication and stakeholder engagement activities

The data deriving by monitoring and assessment of the dissemination, communication results of the project and by the stakeholder engagement with a view to measuring the impact of the relevant activities, fall under Task 7.1 (Q-PLAN) and include (i). Website and social media accounts (i.e., YouTube, Facebook, LinkedIn, Twitter, at this point in the project) analytics; (ii). Data collected from project events; (iii) Data collected from dissemination and communication actions (e.g., participation to external events, participation in project workshops, etc.); and (iii). Newsletter subscriptions. The data will be identified through online analytics, utilising google analytics, SMAs, partners reporting, Mailchimp platform, etc. Q-PLAN is responsible for sending the necessary templates (.docx, .xlsx) to all partners, alongside with guidelines on how to fill them in, as well as for collecting input on an ad-hoc basis (i.e., each time a dissemination or stakeholder engagement action is performed). Q-PLAN is also responsible for preparing the necessary reports to evaluate the overall progress of dissemination and communication activities (measuring outcomes against pre-set KPIs) throughout

the lifespan of the project. The storage format of the data to be collected during the project's duration, includes .csv, .docx, .xlsx, .pdf, .ppt, .jpeg and .png files.

Website and social media analytics

Website analytics: These data will be collected/generated through a periodic monitoring of the project's website statistics with a view to measuring and assessing the performance and results of the project's website activity in terms of dissemination and communication. With that in mind, the data will be mostly quantitative in nature addressing key metrics (e.g., number of visitors, average page views per session, session duration, average time on page, country of visitors, occupation of visitors, etc.). All in all, the data will be stored in a spreadsheets (.xlsx or .csv) while at the same time the analysis of the results will be stored in a standard word document (.docx).

Social media statistics: These data will be collected/generated through a periodic monitoring of the project's social media statistics (i.e., Facebook, YouTube, Twitter, and LinkedIn) with a view to measuring and assessing the performance and results of the project's social media activity in terms of dissemination and communication. With that in mind, the data will be mostly quantitative in nature addressing the metrics reached on each channel (e.g., number of followers, tweets impressions on twitter, friends, likes on Facebook, number of people reached through posts, etc.). All in all, the data will be stored in a Microsoft excel file (.xlsx) while at the same time the analysis of the results will be stored in a standard word document (.docx).

Data collected from project events

These data will be collected during the implementation of the project through: (i) the different events (e.g. co-creation workshops, trainings, interviews, physical and virtual events, etc.) organised by AGIMUS (either alone or jointly with other projects or initiatives) consisting of the participants lists that will enclose demographic information about the participants; and (ii) the participation of AGIMUS partners in relevant third party events in order to reach out and engage stakeholders, thus collecting general information about the events attended and their outreach.

Along these lines, these data will be collected so as to keep track of the results of activities in events for stakeholder engagement and provide the opportunity to project partners to report on these activities. Moreover, these data will be updated every time a partner attends an event, or a partner organises an event. Finally, the data will be both quantitative and qualitative in nature and will be stored in a standard spreadsheet (.xlsx).

Data collected from dissemination and communication actions

These data will be collected through the periodic monitoring of the project's miscellaneous dissemination activities such as publications in relevant journals, posts in the blogs, etc. The data will consist of a spreadsheet (i.e., .xlsx or .csv) designed to keep track of any kind of communication and

dissemination activity, including, but not limited to, press releases, social media posts, website articles, interviews, events (conferences, meetings, workshops, etc.), other publications, e-mails, presentations, informal discussions, seminars, etc. The purpose of collecting these data is to assess the outreach and efficiency of the dissemination activities during the implementation of the project. For this purpose, a template will be shared with all partners to recommend activities to be performed and log the activities they performed. The template is provided also online so as the partners can directly update their input. Finally, all the data will be integrated in a single spreadsheet (.xlsx or .csv).

Newsletter subscriptions

A subscription form hosted in the project's web site will aid the collection of these data in which any interested stakeholder can freely provide their contact details in a dedicated sign-up form so as to receive the most up-to-date news and outcomes of the project. A newsletter will be sent to subscribers once per 6 months. With that in mind, these data will be collected so as interested stakeholders can be informed about AGIMUS's progress and upcoming events, as well as its Toolbox. The data will be comprised of a list of subscribers along with their personal information such as: (i) email address, (ii) first and last name, (iii) country, (iv) type of organisation, and (v) region. A copy of this contact list will be stored to MailChimp's server (http://mailchimp.com), which is used for e-mail campaigns and newsletters distribution. All personal information included in this contact list will be used and protected according to MailChimp's Privacy Policy. Subscribers can find details about how their personal information are managed on the respective privacy policy section of the AGIMUS website.

2.3 Origin of data and re-use of pre-existing data

In the context of AGIMUS, **new data** will be collected/generated by partners as well as external stakeholders participating in the activities of the project. With that in mind and aside consortium partners, **external groups of stakeholders from which new data will originate include**:

- Agile Production personnel that may serve as end-users of the AGIMUS solutions;
- Knowledge, technology, and innovation solution providers (e.g., within academic institutions, non-university public research organisations, technology organisations, high-tech SMEs and large enterprises, etc.);
- AI, Data, and Robotics Experts with scientific and industrial expertise;
- Relevant Initiatives (EU projects focusing on AI, Data, and Robotics networks and working groups);
- Policy makers at regional, national and EU level (related to AI, Data, Robotics, and Industrial Applications)

Moreover, **pre-existing data** will be utilised within the context of AGIMUS. In particular, outputs from EU-funded projects (e.g., MEMMO, DIH4AI, EVOLVED-5G, IMPACT, SPRING, OPTIMAI, etc.), national projects, institutions, Industrial Advisory Board members, and other relevant initiatives (e.g., ADRA and EFFRA) in a large extent will provide a solid basis for AGIMUS. The AGIMUS consortium

will strive to make the most of and advance the work and results of these projects and initiatives. Moreover, relevant data will be harvested from trusted selected online sources regarding the state of play on agile production in European industries, with focus on the consortium countries. Finally, consortium partners' internal knowledge, experience and expertise from their participation in other projects and initiatives will directly and indirectly support the implementation of activities throughout the project.

2.4 Expected size of data

AGIMUS entails a series of activities aiming at setting the stage for and ultimately facilitating the demand-driven evidence-based development, piloting, evaluation, validation and fine-tuning of the AGIMUS framework and value propositions. With that in mind, the table that follows presents the different activities implemented during the course of the project in which data is collected/generated, the types and formats of the data as well as the expected size of the data.

Table 3: Expected size of data

No	Name of activity	Data	Type of data	Format of data	Expected size of data[7]
1	Identification of needs, challenges, opportunities, data, and platforms for agile production	Desk Research mapping and stakeholders' opinions for versatile autonomous robotics for agile production	Interview transcripts, Notes, Minutes, Questionnaires, Spreadsheets	.docx/ .xlsx	< 50 MB
2	Identification of ethical, trustworthy, and secure design guidelines, for next-generation robotics for agile production	Desk Research mapping and stakeholders' opinions for ethical, trustworthy, and secure design guidelines for versatile autonomous robotics in agile production	Interview transcripts, Notes, Minutes, Questionnaires, Spreadsheets	.docx/ .xlsx	< 50 MB
3	Technical and non-Technical specifications during the co-design (technical) workshops	Technical and non-Technical specifications, opinions, requirements, and blueprints for the AGIMUS system architecture	Notes, Minutes, Designs	.docx, .bag	< 500 MB
4	Motion Solver Toolkit Benchmark Dataset	Problem data including the mobile manipulation platform model, the	Robot models, environment	.tar .gz .zip	1 GB

[7] The estimated expected size of the data is based on the adjusted size of data generated via similar activities of project partners in the past unless otherwise indicated.

No	Name of activity	Data	Type of data	Format of data	Expected size of data[7]
		production environment, the task(s) to be performed, and the (optimised) results.	descriptions, task description		
5	Sensor readings recorded during experiments (real or simulated)	Sensor readings from the robot recorded during experiments (real or simulated)	Images, point clouds, joint angles	.bag	100 GB
6	Motion Dataset	Collection of trajectories	Trajectories	TBD	100 GB
7	Videos of motions performed by human actors	Recording of human actor asked to perform given manipulation task	Videos	.mp4	15 GB
8	6D object pose estimation training and evaluation	Collection of real and synthetic images together with the geometrical description of corresponding environments	Images, poses, environment geometry	.tar.gz, .zip	10GB
9	Collection of data for learning sensor-to-state predictive model	Recording of robot state, including the force/torque signals, joints positions and velocities, etc.	Sensor readings	.bag	1GB
10	Specification and deployment of the stress-testing experiments and benchmarking criteria in the AGIMUS testing zones	Robot performance benchmarking and stress-testing	List of components, detailed comparative analysis, Rosbags, time series of sensor readings	.docx, .xlsx, .bag	2 GB
11		Visual Control Experiment Dataset	Images	.jpeg, .mp4	100 Gb
12		Physical Interaction Experiment Dataset	Images, poses, meshes	.tar.gz/. zip	10 Gb
13		Multi-step Task Planning Experiment Dataset	Videos of human actors	.mp4	15 Gb
14	Specifications, deployment, and evaluation of the AGIMUS industrial pilot use cases	AIRBUS Use Cases Dataset	Sensor readings, images, videos, CADs	.bag, .jpeg, .png .mp4, .dwg, .mp4	100 GB
15		KLEEMANN Use Cases Dataset	Sensor readings, images, videos, CADs	.bag, .jpeg, .png .mp4,	100 GB

No	Name of activity	Data	Type of data	Format of data	Expected size of data[7]
				.dwg, .mp4	
16		THIMM Use Cases Dataset	Sensor readings, images, videos, CADs	.bag, .jpeg, .png .mp4, .dwg, .mp4	100 GB
17		Cross-pilot Comparative Assessment Dataset	Notes, Questionnaires, Documents, Spreadsheets, pilot results	.docx, .xlsx, .bag	50 MB
18	Business planning exercise	Market analysis, Competition analysis, as well as financial planning	Notes, spreadsheets	.docx, .xlsx	20 MB
19	Monitoring and assessment of the dissemination, communication, and stakeholder engagement activities	Website and social media analytics	User generated/ machine generated	.csv, .docx, .xlsx, .pdf, .ppt, .jpeg, .png	20 MB
20		Data collected from project events	Spreadsheets	.xlsx	200 KB
21		Data collected from dissemination and communication actions	Spreadsheets, images	.csv, .xlsx, .pdf, jpeg, .png, .docx	200 KB
22		Newsletter subscriptions	Spreadsheets	.csv, .xlsx	100 KB
23	Collaboration and Synergies with ADRA and Related Networks and Initiatives	Data on the collaboration and synergies (type, purpose, organization/initiative/sister project, photos) activities	Photos, number of event participants, presentations, joint activities,	.docx, .xlsx, .pdf, .ppt, .jpeg, .png	200 MB
24	Setup and Operation of AGIMUS' Advisory Boards	IAB and EAB membership and activities	Notes, spreadsheets, images	.docx, .xlsx, .pdf, .jpeg	5 MB

2.5 Data utility

The stakeholders that may find meaningful utility for the data to be collected/generated by the project (both within as well as outside of AGIMUS' consortium) along with the benefits that could arise for them by utilizing these data, are concisely presented in the table that follows.

Table 4: Data Utility

Stakeholder group	Data utility
Academic institutions/ Research Centres	AGIMUS will provide the academic community and researchers with relevant datasets that could be utilised for further analysis and research in order to address a wide range of technical challenges in many fields, such as perception, trajectory optimisation, task-and-motion, learning by demonstration, among others.
Technology Providers	AGIMUS data can be utilised by technology companies and developers in the AI, Data, and Robotics domain, who are interested in developing their own applications for versatile robotics for agile industrial production. In particular, technology companies can take advantage of the proposed open-source solutions, as well as data collected/generated during the implementation of the project activities.
Industrial Entities	The data foreseen to be collected/generated in the frame of AGIMUS is expected to offer additional value to manufacturing industrial entities, such as the AGIMUS industrial pilots (early adopters), which are interested in optimising their processes and embrace technological innovation. The AGIMUS datasets can be used mainly as evidence for the benefits of the involved technologies for both the workforce (from executive personnel to shopfloor workers) and the business.
Policy Makers	The AGIMUS testing zone and industrial pilot use cases will produce quantifiable data which can fuel the design, implementation and deployment strategies of relevant safe, trustworthy and ethical AI-powered robotics in industrial environments. Policy makers can utilise the AGIMUS data to conduct informed decisions regarding the support policies of AI-powered versatile autonomous robotics in industrial environments, covering all s/w and h/w aspects from designing to deploying.
Standardisation - code-driven [open-source]	Blueprint models and libraries generated by AGIMUS could be re-used by those bodies in their standardisation effort, especially towards trustworthy AI.

3. FAIR data

The guidelines on Data Management Plans available in the Horizon Europe Data Management Plan Template of the Commission emphasise the importance of making the data produced by projects funded under Horizon Europe **Findable, Accessible, Interoperable as well as Reusable (FAIR)**, with a view to ensuring its sound management. This means using standards and metadata to make data discoverable, specifying data sharing procedures and which data will be open, allowing data exchange via open repositories as well as facilitating the reusability of the data. With that in mind, the following sections of the DMP lay out the methodology followed in the framework of AGIMUS with respect to making data findable, accessible, and interoperable as well as ensuring their preservation and open access, with a view to increasing its re-use.

3.1 Making data findable, including provisions for metadata

3.1.1 Data discoverability and identification mechanisms

AGIMUS places special emphasis in enhancing the discoverability of the data collected/generated or re-used during the course of its activities. **Open data produced during the implementation of the project will be locatable by means of a standard identification mechanism**. Indeed, AGIMUS will be able to assign globally resolvable **Persistent Identifiers (PIDs)** on any open data (more information on open data as well as the respective repositories we plan on employing in the context of the project are provided on section 3.2). An identifier is a unique identification code that is applied to a dataset, so that it can be unambiguously referenced. For example, a catalogue number is an identifier for a particular specimen and an ISBN code is an identifier for a particular book. PIDs are simply maintainable identifiers that allow for permanent reference to a digital object. In other words, PIDs are a way of giving digital resources, such as documents, images and data records, a unique and persistent reference number.

At the same time, data that are not open will be deposited in a searchable resource (i.e., the cloud web storage service of the project) and well-tailored identification mechanisms will be utilized as well, in the form of standard naming conventions that will safeguard their consistency and make them easily locatable for partners within the frame of the project. Along these lines, the following subsection provides further analysis on naming conventions and versioning.

3.1.2 Naming conventions and versioning

Following a consistent set of naming conventions in the development of the project's data files can greatly enhance their searchability. With that in mind, AGIMUS creates consistent data file names that provide clues to their content, status and versioning, while also increasing their discoverability.

In doing so, project partners as well as interested stakeholders can easily identify a file as well as classify and sort them.

According to the UK Data Archive (UK Data Service, 2017b), a best practice in naming convention is to create brief yet meaningful names for data files, that facilitate classification. The naming convention should avoid the utilisation of spaces, dots and special characters (such as & or !), whereas the use of underscores is endorsed, to separate elements in the data file name and make them understandable. At the same time, versioning should be a part of a naming convention to clearly identify the changes and edits in a file.

With that in mind and to facilitate the reference of the datasets that will be produced during its implementation, AGIMUS employs a **standard naming convention that integrates versioning and takes into account the possibility of creating multiple datasets** during an activity that entails data collection/generation. Indeed, AGIMUS's naming convention considers this issue and addresses it by employing a unique element that captures the number of datasets that are produced under the same activity.

In particular, the **naming convention employed by the project is described below**.

[Name of project] _ [Name of Study] _ [Number of dataset] _ [Issue Date] _ [Version number]

- **Name of project:** AGIMUS
- **Name of Study:** A short version of the name of the activity for which the dataset is created.
- **Number of dataset:** An indication of the number assigned to the dataset.
- **Issue Date:** The date on which the latest version of the dataset was modified (YYYYMMDD).
- **Version number:** The versioning number of a dataset.

With the above in mind, some **indicative examples** to showcase the naming structure applied in the context of AGIMUS are provided below:

- **AGIMUS_NeedChallengesOpportunities_Dataset1_20230331_v1 –** The first dataset generated from the semi-structured interviews conducted to identify their perceived needs, barriers, and opportunities in versatile autonomous robotics for agile production. This is the first version of the dataset that was last modified on the 31st of March 2023 (31/03/2023).
- **AGIMUS_VisualControlExperiment_Dataset2_20241231_v2 –** The second dataset created in the process of the Visual Control Experiment under T5.3, presenting data and results from the experimental procedure in the testing zones. The last modification of this dataset, which in this case produced the second version of the dataset, was on the 31st of December 2024 (31/12/2024).

Versioning of information makes a revision of datasets uniquely identifiable and can be used to determine whether and how data changed over time and to define specifically which version the creators/editors are working with. Moreover, effective data versioning enables understanding if a newer version of a dataset is available and which are the changes between the different versions allowing for comparisons and preventing confusion. In this context, **a clear version number indicator**

is used in the naming convention of every data file produced during AGIMUS in order to facilitate the identification of different versions.

3.1.3 Metadata allowing discovery

In addition to consistent naming conventions and versioning, the project also follows a metadata-driven approach so as to allow discovery and further increase the searchability of the data, while also facilitating its understanding and re-use. Metadata is defined as “data about data” or “information about information”[8]. It is usually structured textual information that describes the creation, content, or context of a digital resource – be it a single file, part of a single file, or a collection of many files. Metadata is the glue which links information and data across the world wide web. It is the tool that helps people to discover, manage, describe, preserve and build relationships with and between digital resources [9].

In particular, three distinct types of metadata exist[10], as presented below:

- **Descriptive metadata**, used to identify and describe collections and related information resources. Descriptive metadata at the local level helps with searching and retrieving. In an online environment, descriptive metadata helps to discover resources. Most of the times includes information such as the title, author, date, description, identifier, etc.
- **Administrative metadata** is used to facilitate the management of information resources. It is helpful for both short-term and long-term management and processing of data. This is information that will not usually be relevant to the public but will be essential for staff to manage collections internally. Such metadata may be location information, acquisition information, etc.
- **Structural metadata** enables navigation and presentation of electronic resources. It documents how the components of an item are organized. Examples of structural metadata could be the way in which pages are ordered to form chapters of a book, a photograph that is included in a manuscript or a scrapbook or the JPEG and TIF files that were created from the original photograph negative, linked together.

With that in mind, **data produced/used during AGIMUS is discoverable with metadata** suitable to its content and format. The project employs **metadata standards** to produce rich and consistent metadata with a view to supporting the long-term discovery, use and integrity of its data More details on the metadata standards adopted by AGIMUS are provided on the following subsection.

[8] Huxley, L., & Jacobs, N. (2004). Online information services in the Social Sciences. Oxford: Chandos.
[9] Foulonneau, M., & Riley, J. (2008). Metadata for digital resources: Implementation, systems design and interoperability. Oxford: Chandos.
[10] Caplan, P. (2003). Metadata fundamentals for all librarians. Chicago: American Library Association.

3.1.4 Standards for metadata creation

AGIMUS employs standards for creating metadata for data collected/generated by the project, with a view to describing it with **rich metadata** and thus improving their discoverability and searchability. In result, effective searching, improved digital curation and easy sharing will be realized. In addition, the metadata standards applied enable the integration of metadata from a variety of sources into other technical systems.

With that in mind, **for AGIMUS' openly available data the metadata standards provided by Zenodo will be used**. Zenodo creates metadata to accompany the datasets that are uploaded to its repository, extending their reach to a wider audience of interested stakeholders. This metadata can be exported in several standard formats, including open and machine-readable ones (such as MARCXML, Dublin Core, and DataCite Metadata Schema), following the guidelines of OpenAIRE and are stored by Zenodo in JSON-format according to a defined JSON schema[11].

Project **data not available for re-use, will also be annotated with open and machine-readable metadata** following the **Dublin Core Metadata standard**. The Dublin Core Metadata element set (certified with the ISO Standard 15836) is a standard which can be easily understood and implemented and as such, is one of the best-known metadata standards. It was originally developed as a core set of elements for describing the content of web pages and enabling their search and retrieval. Among the reasons for selecting this standard is also the fact that **Zenodo is compatible with Dublin Core metadata formats** and thus any initially closed data, that may become open at a later stage (e.g., due to a change in the consortium's policy), will not lose its metadata. With that said, the Dublin Core metadata standard is a simple yet effective set for creating rich metadata that will describe a wide range of resources. The fifteen element "Dublin Core" described in this standard is part of a larger set of metadata vocabularies and technical specifications maintained by the Dublin Core Metadata Initiative (DCMI). The full set of vocabularies also includes sets of resource classes, vocabulary encoding schemes, and syntax encoding schemes. **An online metadata generator will be used** to produce the different metadata elements required (dublincoregenerator.com).

3.1.5 Search keywords included in the metadata

The project's data will be provided with search keywords with a view to optimizing its findability as well as its ultimate re-use by interested stakeholders during its entire lifetime. With that in mind, the metadata standards employed by AGIMUS provide opportunities for tagging the data collected/generated and its content with keywords. In general, keywords are a subset of metadata and include words and phrases used to name data. In the context of AGIMUS, keywords are used to add valuable information to the data collected/generated as well as to facilitate the description and interpretation of its content and value.

[11] For more information on the JSON format and the JSON schema visit the following website: http://json-schema.org/

Along these lines, the project's strategy on keywords is underpinned by the following principles:

- The who, the what, the when, the where, and the why should be covered.
- Consistency among the different keyword tags needs to be ensured.
- Relevant, understandable and clear keywording ought to be sought.

In general, the keywords will comprise terms related to agile production, versatile and autonomous robotics, AI-powered robotics, perception, learning, and task-and-motion. The keywords will accurately reflect the content of the datasets and avoid words used only once or twice within them.

3.1.6 Offering metadata that can be harvested and indexed

We know that the wild diversity of the metadata accompanying open data across the plethora of online repositories (e.g., disciplinary archives, institutional repositories, open access journals) can serve as barriers for their findability and sharing amongst different research communities. This is why in the context of AGIMUS we have aligned our metadata creating approach with the **Open Archives Initiative (OAI)**, which promotes the use of a standard protocol for metadata harvesting, designed for better sharing and retrieval of data residing in distributed repositories. This protocol, namely the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)[12] promotes interoperability standards that facilitate efficient dissemination of data amongst diverse communities[13].

All structured **metadata linked to the project's open data will be offered in a way that can be exported and harvested** via the OAI-PMH thanks to the standards we adopt for metadata creation (see section 3.1.4). The same standards will also help us produce **metadata that facilitate indexing**. For instance, the use of the Dublin Core Metadata Standard (as further elaborated in section 3.3) provides a vocabulary of concepts with definitions in open-machine readable formats that enable easier indexing of metadata. Along these lines, there are several tools[14] which implement the Archives Initiative Protocol for Metadata Harvesting, such as **Arc source**, **EnhancedOAIServer** and **eprints.org**, and can be used for harvesting our data by different repositories.

AGIMUS' openly available data will be uploaded in Zenodo, which is in line with FAIR principles, including "To be Findable" principle. Metadata of each record uploaded in Zenodo is indexed and searchable directly in Zenodo search engine immediately after publishing. Metadata of each record is sent to DataCite servers during DOI registration and indexed there.

[12] Retrieved from: https://www.openarchives.org/pmh/

[13] Corrado, E.M. (2005) 'The importance of open access, open source, and open standards for libraries', Issues in Science and Technology Librarianship.

[14] For more information about the tools implementing the OAI-PMH: https://www.openarchives.org/pmh/tools/

3.2 Making data accessible

3.2.1 Repository

The **data produced by AGIMUS and deemed open for sharing and re-use, will be deposited to and securely stored by Zenodo** (www.zenodo.org), **which constitutes an open data repository** and has been specifically selected to enable access to the project's open data free of charge. In fact, Zenodo builds and operates a simple service that enables researchers, scientists, EU projects and institutions, among others, to share and showcase research results (including data and publications) that are not part of the existing institutional or subject-based repositories of the research communities. It accepts any file format, promotes peer-reviewed openly accessible research, allows the creation of own collections and it is available free of charge both for AGIMUS to upload and share data as well as for other stakeholders to explore, download and re-use this data.

Figure 1: Typical DOI created by Zenodo

Moreover, as a digital repository, Zenodo registers **Digital Object Identifiers (DOIs)** for all submitted data through DataCite[15], which is the leading global non-profit organisation that provides PIDs (and specifically DOIs) for research data and preserves these submissions using the safe and trusted foundation of CERN's data centre, alongside the biggest scientific dataset in the world, the LHC's 100PB Big Data store[16]. This means that the data preserved in Zenodo will be accessible for years to come, and the DOIs will function as perpetual links to the resources. DOIs remain valuable since they are future proofed against Uniform Resource Locator (URL) or even protocol changes, through resolvers (such as DOI[17]). With that in mind, an example of a DOI retrieved from this open repository follows the structure illustrated by Figure 1.

Accordingly, open-source solutions and related data will be deposited and securely stored by GitHub/GitLab hosted by CNRS (https://gitlab.laas.fr/gepetto). The AGIMUS repository has already been created[18].

3.2.2 Data

Openly available and closed data

AGIMUS follows the guidelines of the Horizon Europe Data Management Plan Template, and is in line with FAIR principles and the rule "as open as possible, as closed as necessary", aiming to "*make*

[15] For more information on DataCite: https://www.datacite.org/
[16] Retrieved from: https://www.software.ac.uk/tags/zenodo
[17] Retrieved from: http://dx.doi.org/
[18] https://gitlab.laas.fr/gepetto/agimus-doc

the data collected/generated openly available with as few restrictions as possible, while at the same time protecting sensitive data from inappropriate access"[19]. This calls for project partners to disseminate the project's data that have the potential to offer long-term value to external stakeholders and do not harm the confidentiality and privacy of the stakeholders that contributed to the collection/generation of this data, maximising the beneficial impact of AGIMUS.

Only anonymised and aggregated data will be made open to ensure that data subjects cannot be identified in any reports, publications and/or datasets resulting from the project, unless explicit consent has been collected following the respective GDPR provisions. The relevant project partner in each case will undertake all the necessary anonymisation procedures to anonymise the data in such a way that the data subject is no longer identifiable (more details on data management responsibilities are provided in Section 5.2).

To this end, it is important to keep in mind that during the process of data anonymisation, data identifiers need to be removed, generalised, aggregated or distorted. Moreover, **anonymisation is different than pseudonymisation**, which falls under a distinct category in the GDPR - anonymisation theoretically destroys any way of identifying the data subject, while pseudonymisation allows for the data subject to be re-identified with additional information. Along these lines, the table which follows provides a **list of good practices** for the anonymisation of quantitative and qualitative data derived from the tour guide on data management of the Consortium of European Social Science Data Archives (CESSDA).

Table 5: Good practices for data anonymisation

Type of data	Good practices
Quantitative data	• Remove or aggregate variables or reduce the precision or detailed textual meaning of a variable. • Aggregate or reduce the precision of a variable such as age or place of residence. As a general rule, report the lowest level of geo-referencing that will not potentially breach respondent confidentiality. • Generalise the meaning of a detailed text variable by replacing potentially disclosive free-text responses with more general text. • Restrict the upper or lower ranges of a continuous variable to hide outliers if the values for certain individuals are unusual or atypical within the wider group researched.
Qualitative data	• Use pseudonyms or generic descriptors to edit identifying information, rather than blanking-out that information.

[19] Koulocheri, E. (2017). What is the Open Research Data Pilot? Retrieved from https://www.openaire.eu/what-is-the-open-research-data-pilot

Type of data	Good practices
	• Plan anonymisation at the time of transcription or initial write-up, (longitudinal studies may be an exception if relationships between waves of interviews need special attention for harmonised editing). • Use pseudonyms or replacements that are consistent within the research team and throughout the project. For example, using the same pseudonyms in publications and follow-up research. • Use 'search and replace' techniques carefully so that unintended changes are not made, and misspelt words are not missed. • Identify replacements in text clearly, for example with [brackets] or using XML tags such as <seg>word to be anonymised</seg>. • Create an anonymisation log (also known as a de-anonymisation key) of all replacements, aggregations or removals made and store such a log securely and separately from the anonymised data files.

Source: Tour guide on data management of the CESSDA[20]

With that in mind, the following table presents the data collected/generated during the course of the project that will be made openly available. In case certain data cannot be shared (or need to be shared under restrictions), a justification for that choice is provided.

Table 6: Data Availability

No	Data	Availability	Notes
1	Identification of needs, challenges, opportunities, data, and platforms for agile production	Open	Any personal information data will be aggregated/ anonymized before becoming openly available.
2	Identification of ethical, trustworthy, and secure design guidelines, for next-generation robotics for agile production	Open	Any personal information data will be aggregated/ anonymized before becoming openly available.
3	Technical and non-Technical specifications and opinions for the AGIMUS system architecture	Open & Closed	It is envisaged that data underpinning the co-design activities may be made available in an aggregated format through deliverable D1.2. Information related to the robot shall remain

[20] Retrieved from: https://www.cessda.eu/Research-Infrastructure/Training/Expert-Tour-Guide-on-Data-Management/5.-Protect/Anonymisation

No	Data	Availability	Notes
			closed as it reflects proprietary knowledge.
4	Motion Solver Toolkit Benchmark Dataset	Open	-
5	Sensor readings recorded during experiments (real or simulated)	Open	Any personal information data will be aggregated/ anonymized before becoming openly available.
6	Motion dataset	Open	-
7	Videos of motions performed by human actors	Open	Any personal information data will be aggregated/ anonymized before becoming openly available.
8	Dataset for 6D object pose estimation training and evaluation	Open	-
9	Collection of data for learning sensor-to-state predictive model	Open	-
10	Robot performance benchmarking and stress-testing	Closed	This contains internal definition of a new proprietary platform
11	Visual Control Experiment Dataset	Open	Any personal information data will be aggregated/ anonymized before becoming openly available.
12	Physical Interaction Experiment Dataset	Open	Any personal information data will be aggregated/ anonymized before becoming openly available.
13	Multi-step Task Planning Experiment Dataset	Open	Any personal information data will be aggregated/ anonymized before becoming openly available.
14	AIRBUS Use Cases Dataset	Closed	Due to strict data privacy regulations in the AIRBUS industrial sites, any data related to the AIRBUS industrial site will remain closed. However, the possibility of a “replica” of the use cases at CNRS will be investigated, in which the data

No	Data	Availability	Notes
			will be re-evaluated if they can be shared.
15	KLEEMANN Use Cases Dataset	Closed	Industrial processes are considered confidential and only selected sub-sets may be made publicly available
16	THIMM Use Cases Dataset	Closed	Industrial processes are considered confidential and only selected sub-sets may be made publicly available
17	Cross-pilot Comparative Assessment Dataset	Open	Any personal information data will be aggregated/ anonymized before becoming openly available.
18	Business planning exercise	Closed	Data closely related to business planning may include sensitive financial information that cannot be made publicly available.
19	Website and social media analytics	Open & Closed	Website analytics will be available only to AGIMUS consortium and the EU Commission. In cases where statistics were shared, data were aggregated and anonymized before being made openly available (e.g., reported in the publicly available Dissemination and Communication Plan of the project), while personal data were treated as expected by the GDPR.
20	Data collected from project events	Open & Closed	Personal data of participants will remain closed, as well as confidential information of SMEs. Aggregated statistics may be published for promotion and reporting purposes.
21	Data collected from dissemination and communication actions	Open & Closed	Any personal information data will be aggregated/ anonymized before becoming openly available.

No	Data	Availability	Notes
22	Newsletter subscriptions	Closed	Data from newsletter subscriptions will remain closed as it contains personal information and is useful only for internal reporting purposes.
23	Setup and Operation of AGIMUS' Advisory Boards	Open & Closed	Contact details and other key personal information from the advisory board members will not be publicly shared. However the constitution of the Advisory Board Members will be open.

It is important to note that **all personal data collected / generated will be considered as closed data prior to their anonymisation and aggregation** to safeguard the confidentiality of the data subjects.

Data accessibility and availability

Public access to the open data will be made available through Zenodo, which will automatically link to OpenAIRE, and GitHub. The data will be fully accessible thanks to the included metadata and the search facility available on both online repositories. At the same time, closed data are intended to be stored and shared amongst authorised members of the consortium through cloud storage and file sharing providers which constitute structures that maintain and manage data and make these data accessible over a network, usually the internet (i.e. CNRS/LAAS ownCloud or CIIRC cloud data repository). Before starting using these cloud services from providers situated both inside and outside the European Economic Area (EEA) we have ensured that they comply with the relevant GDPR requirements.

The following table presents where data will be made accessible in the context of AGIMUS.

Table 7: Data Accessibility

No	Data	Accessibility
1	Identification of needs, challenges, opportunities, data, and platforms for agile production	AGIMUS Website / Zenodo
2	Identification of ethical, trustworthy, and secure design guidelines, for next-generation robotics for agile production	AGIMUS Website, Zenodo

No	Data	Accessibility
3	Technical and non-Technical specifications and opinions for the AGIMUS system architecture	AGIMUS Website, Zenodo, GitHub/GitLab hosted by CNRS
4	Motion Solver Toolkit Benchmark Dataset	GitHub/GitLab hosted by CNRS
5	Sensor readings recorded during experiments (real or simulated)	GitHub/GitLab hosted by CNRS
6	Motion dataset	GitHub/GitLab hosted by CNRS
7	Videos of motions performed by human actors	AGIMUS Website / Big Data Repository hosted by CTU
8	Dataset for 6D object pose estimation training and evaluation	AGIMUS Website / Big Data Repository hosted by CTU
9	Collection of data for learning sensor-to-state predictive model	AGIMUS Website / Big Data Repository hosted by CTU
10	Robot performance benchmarking and stress-testing	Closed data, sensible data not publicly available
11	Visual Control Experiment Dataset	GitHub/GitLab hosted by CNRS / Big Data Repository hosted by CTU
12	Physical Interaction Experiment Dataset	GitHub/GitLab hosted by CNRS / Big Data Repository hosted by CTU
13	Multi-step Task Planning Experiment Dataset	GitHub/GitLab hosted by CNRS / Big Data Repository hosted by CTU
14	AIRBUS Use Cases Dataset	Closed data, sensible data not publicly available
15	KLEEMANN Use Cases Dataset	AGIMUS Website, Zenodo, GitHub/GitLab hosted by CNRS / Big Data Repository hosted by CTU
16	THIMM Use Cases Dataset	AGIMUS Website, Zenodo, GitHub/GitLab hosted by CNRS / Big Data Repository hosted by CTU
17	Cross-pilot Comparative Assessment Dataset	AGIMUS Website, Zenodo
18	Business planning exercise	Closed data
19	Website and social media analytics	Closed data

No	Data	Accessibility
20	Data collected from project events	AGIMUS Website, Zenodo
21	Data collected from dissemination and communication actions	AGIMUS Website, Zenodo
22	Newsletter subscriptions	Closed data
23	Setup and Operation of AGIMUS' Advisory Boards	AGIMUS Website

Restrictions on use

By utilising Zenodo for sharing the project's openly available data, AGIMUS can apply **different levels of accessibility** for this data taking into account any relevant issues (such as ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related, etc.).

More specifically, **Zenodo offers the following levels of data accessibility**:

- **Open access**: Data remains available for re-use. Nevertheless, the level in which this data can be re-used is determined also by their accompanied licence for re-use (see subsection 3.4.3).
- **Embargoed status**: Access to the data will be restricted until the end of the embargo period, at which time, the content will automatically become publicly available.
- **Restricted access**: The data will not be made publicly available, and sharing will be made possible only by the approval of the project partner that have the responsibility of the data.
- **Closed access**: The data is protected against unauthorized access at all levels and only members of the consortium have the right to access it.

The same principle applies also for GitHub/GitLab hosted by CNRS, within which **private and public repositories** can be created with differing sharing and access options.

Accordingly, the repository for larger datasets hosted by CTU can allow both publicly available access (for publicly shared data) or private use only to project partners through password protected access.

Project partners will mainly use the open access level to disseminate the project's data amongst the interested stakeholders. Data that will not be available for re-use will be accessible only by authorised partners of AGIMUS' consortium and /or authorised personnel from the funding authority of the project.

Moreover, **AGIMUS will ensure open access to all peer-reviewed scientific publications** that may be produced in the framework of the project. In particular, according to the Grant Agreement, AGIMUS will:

- At the latest at the time of publication, deposit a machine-readable electronic copy of the published version or final peer-reviewed manuscript accepted for publication in a trusted repository for scientific publications.
- Ensure immediate open access to the deposited publication — via the repository — under the latest available version of the Creative Commons Attribution International Public License (CC BY) or a license with equivalent rights. Moreover, for monographs and other long-text formats, the license may exclude commercial uses and derivative works (e.g., CC BY-NC, CC-BY-ND).
- Ensure information is given – via the repository – about any research output or any other tools and instruments needed to validate the conclusions of the scientific publication.

Beneficiaries (or authors) must retain sufficient intellectual property rights to comply with the open access requirements.

Along these lines, this section has provided the methodology applied in the frame of AGIMUS to ensure that its data are as openly accessible as possible by any stakeholder that may find it interesting for re-use. In this context, AGIMUS also focuses on providing metadata standards and appropriate metadata vocabularies to increase data interoperability. The following section provides further details in this respect.

It should be noted here that CNRS and INRIA scientific publications are automatically available on the open Library HAL[21]. In addition, CTU, PAL, and TOWARD, upload scientific publications to arXiv[22].

Identity ascertainment and data access committee

The identity of stakeholders who want to access the open data on Zenodo, GitHub, or CTU's database, is not necessary to be ascertained, as the uploaded data are publicly open and no authorization is needed. On the other hand, closed for the public data will be available only to authorized consortium partners through dedicated mechanisms provided by the cloud storage service employed by the respective partners in order to deposit the data. As further elaborated in Section 6 of this DMP, technical access controls are built into the AGIMUS repositories, in order to ascertain the identity and access rights of those who want to access the data.

The need for a data access committee to evaluate or approve access requests to personal data, is not foreseen because only authorized partners will have access to the project's closed data, accessible only by using their credentials (username/password), and no third-party will re-use them for their benefit.

[21] https://hal.archives-ouvertes.fr
[22] https://arxiv.org/

3.2.3 Metadata

Availability and licences

Metadata of deposited publications generated in the context of AGIMUS will be **open under a Creative Common Public Domain Dedication (CC 0)** or equivalent, in line with the FAIR principles for data management adopted by the project (in particular machine-actionable). Such **metadata will provide information, at least, about the following**:

- The publication at hand (author(s), title, date of publication, publication venue);
- Reference to the Horizon Europe funding;
- The name of the project, including its acronym and Grant Agreement number;
- Any particular licensing terms which may apply (depending on the chose license);
- Persistent identifiers that have been attributed to the publication;
- Authors involved in the action, their organisations and the project itself.

Where applicable, the metadata will also include persistent identifiers for any research output, or any other tools and instruments needed to validate the conclusions of the publication. The metadata will be available through Zenodo and GitHub. It is quite unlikely that Zenodo or GitHub will terminate its operation and stop providing its services, but in such a case all data, metadata, code and documentation uploaded will be transferred and hosted to other suitable repositories without undue delay. In this respect, it is important to note that, since all of AGIMUS' openly available data will make use of PIDs (i.e., DOIs), which is further elaborated in subsection 3.1.1, the links to the data will not be affected. In parallel, the project's data that will not be openly available for sharing will be deposited, together with their accompanying metadata, code and documentation (if necessary), to the cloud web storage service employed by the project.

Methods, software tools and documentation to access the data

AGIMUS emphasises the accessibility of the data collected / generated during the project. With that in mind, no specialised method, software tool and / or documentation is expected to be needed at the moment, in order to access the data. Stakeholders will have the ability to access the data by simply using their web browser (e.g., Mozilla, Google Chrome, Internet Explorer, Safari, etc.) through their computers (either desktop or laptop), smart phones and/or tablets.

More specifically, through the AGIMUS website clear links to the selected repositories will allow direct navigation to the publicly available datasets in all three online repositories (i.e., Zenodo, GitHub, CTU's repository). For Zenodo and GitHub, stakeholders will also be able to utilise the search engine of the repositories to search for interesting data. By typing the name of the project (or any other relevant keyword connected to the AGIMUS data) the search engine will direct the user to the project's data, ready to be explored and re-used. Moreover, since the data will be available in open formats, we will be ensuring that they can appropriately be read by a range of different software that are widely and freely accessible to all potential users of the data.

Closed data will only be accessed by authorised project partners through usage of a cloud storage service. Again, no specialised method, software tool and / or documentation is needed to this end.

3.3 Making data interoperable

Data interoperability refers to the ability of systems and services that create, exchange, and use data to have clear, shared expectations for the contents, context and meaning of that data[23]. With that in mind, AGIMUS has adopted in its data management methodology the use of metadata vocabularies, standards and methods that will increase the interoperability of the data collected/generated through its activities.

More specifically, **the interoperability of the data that will not be publicly shared will be facilitated by the use of the Dublin Core Metadata standard.** This standard is a small "metadata element set" which accounts for issues that must be resolved in order to ensure that data meet traditional standards for quality and consistency, while still remaining broadly interoperable with other data sources in the linked data environment. The fifteen elements of the standard provide a vocabulary of concepts with natural-language definitions (e.g., title, creator, author, etc.) that are instantly converted into open machine-readable formats (such as XML, HTML, etc.), enabling machine-processability. Each element is optional and may be repeated, while the standard itself offer ways exist for refining them, encouraging the use of encoding and vocabulary schemes. The vocabulary of the Dublin Core Metadata standard is presented by the following table[24].

Table 8: Dublin Core Metadata standard vocabulary

No	Element	Element definition
1	Title	A name given to the resource.
2	Creator	An entity primarily responsible for making the content of the resource.
3	Subject	The topic of the content of the resource.
4	Description	An account of the content of the resource.
5	Publisher	An entity responsible for making the resource available.
6	Contributor	An entity responsible for making contributions to the content of the resource.

[23] L. Steele & T. Orrell (2017). The frontiers of data interoperability for sustainable development. Publish What You Fund and Development Initiatives

[24] Sugimoto, S., Baker, T., & Weibel, S. L. (2002). Dublin Core: Process and Principles. Lecture Notes in Computer Science Digital Libraries: People, Knowledge, and Technology, 25-35.

No	Element	Element definition
7	Date	A date associated with an event in the life cycle of the resource
8	Type	The nature or genre of the content of the resource.
9	Format	The physical or digital manifestation of the resource.
10	Identifier	An unambiguous reference to the resource within a given context.
11	Source	A reference to a resource from which the present resource is derived.
12	Language	A language of the intellectual content of the resource.
13	Relation	A reference to a related resource.
14	Coverage	The extent or scope of the content of the resource.
15	Rights	Information about rights held in and over the resource.

Along similar lines, **the interoperability of openly available data will be facilitated through Zenodo** since its metadata will be stored internally in JSON format according to a defined JSON schema. This encloses HTML microdata that allows machine-readable data to be embedded in HTML documents in the form of nested groups of name-value pairs. Moreover, the JSON schema provides a collection of shared vocabularies in microdata format that can be used to mark-up pages in ways that can be understood by the major search engines. Moreover, all metadata linked to the open data is exported via the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) and can be harvested. The OAI-PMH develops and promotes interoperability standards that facilitate the efficient dissemination of data amongst diverse communities[25].

In a similar manner, GitHub/GitLab can facilitate storing metadata either through git notes[26] or by associated description to a branch. The metadata stored can follow the same JSON format.

[25] Corrado, E.M. (2005) 'The importance of open access, open source, and open standards for libraries', Issues in Science and Technology Librarianship.

[26] https://git-scm.com/docs/git-notes

3.4 Increase data re-use

3.4.1 Documentation for validating data analysis and facilitating data re-use

By utilising Zenodo for sharing the project's openly available data, AGIMUS ensures the facilitation of data access, validation and re-use, according in compliance to the general policies of Zenodo regarding content, access and reuse. More specifically, the following principles are followed by Zenodo to make data re-useable according to the FAIR principles:

- R1: (meta)data are richly described with a plurality of accurate and relevant attributes

Each record contains a minimum of DataCite's mandatory terms, with optionally additional DataCite recommended terms and Zenodo's enrichments.

- R1.1: (meta)data are released with a clear and accessible data usage license

License is one of the mandatory terms in Zenodo's metadata, and is referring to an Open Definition license. Data downloaded by the users is subject to the license specified in the metadata by the uploader.

- R1.2: (meta)data are associated with detailed provenance

All data and metadata uploaded is traceable to a registered Zenodo user. Metadata can optionally describe the original authors of the published work.

- R1.3: (meta)data meet domain-relevant community standards

Zenodo is not a domain-specific repository, yet through compliance with DataCite's Metadata Schema, metadata meets one of the broadest cross-domain standards available.

On the other hand, by utilising GitHub for sharing the project's openly available data, and specifically the open-source solutions developed along the datasets used for the various experiments, AGIMUS offers access, and thus increased re-use, to the scientific community that use it as their main knowledge sharing online repository.

3.4.2 License schemes to permit the widest use possible

The application of a licence to AGIMUS open data is a simple way to ensure that any interested third-party can re-use it. In this context, licences are the instrument which permit a third-party to copy, distribute, display and/or modify the project's data only for the purposes that are set by the licence. Licences typically grant permissions on condition that certain terms are met. While the precise details vary, three conditions are commonly found in licences which are the attribution, non-derivative, and non-commerciality.

Along these lines, AGIMUS publishes openly available data under the **Creative Commons licencing scheme** to foster their re-use and build an equitable and accessible environment for them. Zenodo

provides AGIMUS the **opportunity to publish its open data under five Creative Common licences** as follows:

- **Creative commons Attribution-Share Alike 4.0** (CC BY-SA 4.0) according to which any third party can freely copy, distribute, display and modify the datasets for any purpose. Remix, transform, or built upon data, must be distributed under the same license as the original. Third parties must give appropriate credit, provide a link to the license, and indicate if changes were made.

Figure 2: CC BY-SA 4.0

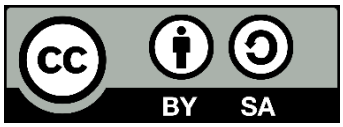
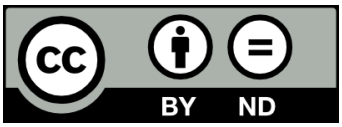


Figure 3: CC BY 4.0

- **Creative Commons Attribution 4.0 International** (CC BY 4.0) according to which any third party can freely copy, distribute, display and modify the datasets for any purpose. Third parties must give appropriate credit, provide a link to the license, and indicate if changes were made.
- **Creative Commons Attribution-No Derivatives 4.0 International** (CC BY-ND 4.0) during which any third party can freely copy, distribute, display and modify the datasets for any purpose. Remix, transform, or built upon data, however must not be distributed. Third parties must give appropriate credit, provide a link to the license, and indicate if changes were made.

Figure 4: CC BY-ND 4.0



- **Creative Commons Attribution-NonCommercial 4.0 International** (CC BY-NC 4.0) based on which third parties can copy, distribute, display and modify the datasets for any purpose other than commercial unless they get a permission by project partners first. Third parties must give appropriate credit, provide a link to the license, and indicate if changes were made.

Figure 5: CC BY-NC 4.0

- **Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International** (CC BY-NC-ND 4.0) according to which third parties can copy, distribute, display and modify the datasets for any purpose other than commercial unless they get a permission by project partners first. Remix, transform, or built upon data, however, must not be distributed. Third parties must give appropriate credit, provide a link to the license, and indicate if changes were made.

Figure 6: CC BY-NC-ND 4.0

Beyond Zenodo licence schemes, Github offers a wider superset of licenses[27] that extends to other types of licenses besides Creative Commons, including but not limited to Academic Free License v3.0, Berkeley Software Distribution (BSD) 2 and 3 licences, European Union Public License 1.1, GNU General Public License v2.0/v3.0, Open Software License 3.0, Educational Community License v2.0, and more.

Different licensing schemes may be selected to better fit the need of AGIMUS' open data ensuring not only their long-term preservation and re-use but also the interests of the consortium along with the rights of individuals for whom the data is about. In such a case, this subsection of the DMP will be updated accordingly.

3.4.3 Availability for re-use

The re-use of data is a key component of AGIMUS' methodology for making data FAIR. In fact, making data available for re-use ensures interested stakeholders, other than project partners, can benefit from this data, contributing towards maximising the impact of the project. Rich metadata created based on metadata standards that enable proper discovery as well as appropriate licensing schemes facilitate the re-use of AGIMUS' open data, allowing them to find valuable utility.

In principle, it is expected that data will become available for re-use no later than 120 days after the end of its processing in the framework of the project (i.e., collection, anonymisation, aggregation, etc.) to ensure that any additional data management activities required to this end do not compete with the timely delivery of the project's planned outputs.

With that in mind, the expected time that AGIMUS' data will be made openly accessible and uploaded to Zenodo is indicatively provided in the following table:

Table 9: Expected time that data will be made open

No	Data	Expected time for making data open	Notes
1	Identification of needs, challenges, opportunities, data, and platforms for agile production	31/10/2023	After approval of D1.1
2	Identification of ethical, trustworthy, and secure design guidelines, for next-generation robotics for agile production	31/10/2023	After approval of D1.1

[27] https://docs.github.com/en/repositories/managing-your-repositorys-settings-and-features/customizing-your-repository/licensing-a-repository

No	Data	Expected time for making data open	Notes
3	Technical and non-Technical specifications and opinions for the AGIMUS system architecture	31/08/2026	After approval of D1.2, D1.3 and at the end of the project
4	Motion Solver Toolkit Benchmark Dataset	31/03/2025	Upon WP2 completion
5	Sensor readings recorded during experiments (real or simulated)	31/08/2025	Upon WP3 completion
6	Motion dataset	31/08/2025	Upon WP3 completion
7	Videos of motions performed by human actors	31/08/2026	Upon WP3 completion
8	Dataset for 6D object pose estimation training and evaluation	31/10/2025	Upon WP4 completion
9	Collection of data for learning sensor-to-state predictive model	31/10/2025	Upon WP4 completion
10	Robot performance benchmarking and stress-testing	- (closed data)	The robotic platform is proprietary, so the data coming from the benchmarking will be sensible and private. No public sharing will be available
11	Visual Control Experiment Dataset	31/08/2024	Upon T5.3 completion
12	Physical Interaction Experiment Dataset	31/03/2025	Upon T5.4 completion
13	Multi-step Task Planning Experiment Dataset	31/08/2025	Upon T5.5 completion
14	AIRBUS Use Cases Dataset	- (closed data)	No public sharing will be available for data related to industrial processes
15	KLEEMANN Use Cases Dataset	- (closed data)	No public sharing will be available for data related to industrial processes
16	THIMM Use Cases Dataset	- (closed data)	No public sharing will be available for data related to industrial processes
17	Cross-pilot Comparative Assessment Dataset	31/08/2026	Upon T6.3 completion
18	Business planning exercise	- (closed data)	No public sharing will be available

No	Data	Expected time for making data open	Notes
19	Website and social media analytics	- (closed data)	No public sharing will be available
20	Data collected from project events	31/08/2026	Upon T7.1 completion
21	Data collected from dissemination and communication actions	31/08/2026	Upon T7.1 completion
22	Newsletter subscriptions	- (closed data)	No public sharing will be available
23	Setup and Operation of AGIMUS' Advisory Boards	30/01/2023	Information about the AGIMUS Advisory Boards will be published as soon as the initial members have been onboarded.

3.4.4 Data provenance

Data provenance is the documentation of where a piece of data comes from and the processes and methodology by which it was produced. Put simply, provenance answers the questions of why and how the data was produced, as well as where, when and by whom[28]. Accurately recording data provenance is a cornerstone of good data management. AGIMUS will use specific elements of the **Dublin Core Metadata Standards**[29] and the **W3C Provenance Data Model**[30], to generate specific text files (e.g., README) that will accurately capture the history of each data entity throughout its versions (e.g., based on the DOI versioning Zenodo provides)[31].

3.4.5 Data quality assurance processes

Quality Assurance (QA) and **Quality Control** (QC) activities[32] are an integral part of AGIMUS' data management methodology and are implemented prior to the publication of any data to online repositories (e.g., Zenodo, GitHub, etc.), safeguarding the transparency, consistency, comparability, completeness, and accuracy of the data.

QA is a planned system of review procedures conducted outside the framework of developing a dataset, by personnel not directly involved in the dataset development process. In the context of

[28] https://ardc.edu.au/resource/data-provenance/
[29] https://www.dublincore.org/resources/userguide/creating_metadata/#Provenance
[30] https://www.w3.org/TR/prov-dm/
[31] https://help.zenodo.org/
[32] 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Vol. 1 General Guidance and Reporting, CHAPTER 6 Quality Assurance / Quality Control and Verification.

AGIMUS, it takes the form of **peer-reviews of methods and/or data summaries** to assess the quality of the dataset and identify any need for improvement, ensuring that the dataset correctly incorporates the scientific knowledge and data generated.

QC is defined as a system of checks to assess and maintain the quality of the dataset being compiled. The relevant procedures of AGIMUS are designed to provide routine technical checks as they measure and control data consistency, integrity, correctness, and completeness as well as identify and address errors and omissions. In this context, QC checks cover everything from data acquisition and handling, application of approved procedures and methods, and documentation. Some of the general quality checks undertaken in the framework of the project include checking (i) for transcription errors in data input; (ii) that scale measures are within the range of acceptable values; and (iii) whether proper naming conventions are used.

4. Other research outputs

The implementation of AGIMUS entails activities that will generate or re-use other research outputs. This includes digital (e.g., software, workflows, protocols, models, etc.) or physical (e.g., robotic experimental platform, mechatronic or robotic components, etc.) research outputs, besides the data described in the second chapter of this DMP. The management of these other research outputs in the context of AGIMUS will be in line with FAIR data principles, as concisely outlined in the table that follows.

Table 10: Other digital or physical research outputs

WP / Task	Research Output	Brief Description	Type	Expected Size	Interested stakeholders and benefits	Availability	Accessibility (repository)	Expected time for making open
WP2	Pinocchio	Software for modelling complex robotic systems	Software	100k lines of code	Same as Section 2.5	Open	GitHub	09/06/2022
WP2	Simple	Software for differentiable simulation	Software	100k lines of code	Same as Section 2.5	Open	GitHub	01/06/2023
WP3	Aligator	Software for optimal control and model predictive control	Software	100k lines of code	Same as Section 2.5	Open	GitHub	01/06/2023

5. Allocation of resources

5.1 Estimated costs for making data FAIR

The costs required for making the data collected/generated during AGIMUS' activities FAIR are integrated in the budget of the project. With that in mind, the table which follows provides an overview of the estimated costs of making data FAIR as well as their budget source within the framework of AGIMUS.

Table 11: Estimated costs for making data FAIR

No	Data Processing / Management Activity	Budget source	Total estimated effort in Person Months[33]	Total estimated cost in Euro[34]
1	Collection	Budget allocated to the WP under which the respective data are processed	51.56	293319.1
2	Documentation	Budget allocated to the WP under which the respective data are processed	12.89	73329.78
3	Storage	Budget allocated to the WP under which the respective data are processed	6.45	36664.89
4	Access and security	Budget allocated to the WP under which the respective data are processed	6.45	36664.89
5	Preservation	Budget allocated to the WP under which the respective data are processed	3.22	18332.44
6	Availability and re-use	Budget allocated to the WP under which the respective data are processed	19.34	109994.7
7	Overall data management	WP8	2.62	14904.89

[33] The total estimated effort for each data processing / management activity reflects the cumulative effort for the implementation of the respective activity for all data collected / generated across the different WPs of AGIMUS.
[34] The total cost of each data processing / management activity is calculated by multiplying the effort estimated for the respective activity with the weighted average cost of a person month in the framework of AGIMUS.

No	Data Processing / Management Activity	Budget source	Total estimated effort in Person Months[33]	Total estimated cost in Euro[34]
		Total	**102.52**	**583210.7**

In order to produce the estimations of the costs for making data FAIR in the context of AGIMUS, a series of **assumptions** were made, taking into account the respective **guidelines** provided by the Research Data Management Support, a multidisciplinary network of data experts within Utrecht University[35], as well as of the UK Data Service and its data management costing tool[36]. With that in mind, the estimated costs for making AGIMUS' data FAIR cover **data-related activities and resources across the data lifecycle**, spanning from collection and documentation through storage and preservation over to sharing and re-use.

In particular, costs for **data collection** cover activities necessary for acquiring external datasets (if required), gathering/generating new data, transcribing (if applicable), formatting and organising this data as well as acquiring informed consent from data subjects. This data processing activity reflects the majority of the costs required for making data FAIR as the majority of AGIMUS' data constitutes new data collected/generated over the course of the project. At the same time, **data documentation** costs address the effort required for describing data (e.g., marking data with variable and value labels, code descriptions, etc.) as well as creating well-defined metadata along with a meaningful description of the context and methodology of how data was collected/generated and processed (where necessary).

Costs for **data storage** include the resources required for ensuring adequate storage space for the data as well as the effort necessary for conducting data back-ups, while **data access and security** costs encompass costs related to ensuring access to the data as well as for protecting it from unauthorised access or use or from disclosure. Given that the storage of AGIMUS's data will not require the procurement of additional space (other than what is already available to project partners) as well as that no special measures or software are required to access and secure the data (other than what is inherently built in to the repositories of AGIMUS' data), such costs are kept to a minimum.

Data preservation costs, on the other hand, are estimated relatively higher than data storage, access and security costs, as additional effort will be required in several cases in order to convert the collected/generated data from their original form (e.g. physical interview transcripts) to an open and/or machine readable format suitable for long-term preservation (e.g. to an .xlsx format.). Adequate effort for **data availability and re-use** costs is also foreseen to safeguard the appropriate digitisation and anonymisation of the data as well as cover any resources required for data sharing and cleaning. Along the same lines, appropriate effort is foreseen for **overall data management** as well, in order

[35] Research Data Management Support. Guides: Costs of data management. Utrecht University. Retrieved from: https://www.uu.nl/en/research/research-data-management/guides/costs-of-data-management

[36] UK Data Service. Costing Data Management. Retrieved from: https://www.ukdataservice.ac.uk/manage-data/plan/costing

to cover the effort related with the operationalisation of data management in the framework of AGIMUS.

Finally, costs for **long-term preservation** in the framework of AGIMUS are assumed to be negligible since the open data of the project will be hosted in the repository of Zenodo free of charge.

5.2 Data management responsibilities

For the effective, proper and secure handling of the data collected/generated in the frame of AGIMUS, specific data management roles have been established within the data management methodology and procedures of the project. These responsibilities are outlined in this section of the DMP and are as follows.

Project Coordinator (PC) and Project Management Office (PMO): responsible for overall data management in the framework of AGIMUS, including the elaboration of the DMP and its updates (when necessary, along with support of all partners). At the same time, the PC is responsible for the elaboration of proper templates for the Informed Consent Form and the Data Subject Request Form to be appropriately adjusted and utilised by project partners during the relevant activities of the project as well as for drafting the project's Privacy Policy that has been uploaded on the project's website. The PC in collaboration with the relevant project partners (e.g., Task Leaders) will examine if additional specific privacy policies are required for certain project's tasks and will coordinate the elaboration of such privacy policies. Finally, the PC coordinates with Work Package Leaders and Co-leaders, Task Leaders, and Responsible Partners to determine whether and how the data collected / generated by the project are shared and become available for re-use, contributes to its quality assurance and uploads the project's openly available data to Zenodo.

Work Package Leaders (WPL): The WPL is responsible for coordinating the implementation of the data processing activities performed under the WPs they are leading. Moreover, they align with the PC and the respective Work Task Leader on whether and how the data gathered/produced under the tasks that fall within the WP they are leading will be shared and/or re-used. This includes the definition of access procedures as well as potential embargo periods along with any necessary software and/or other tools which may be required for data sharing and re-use. Finally, the WPL are the main responsible for assuring the quality of the data stemming from the activities of the WP they are leading, including assessing their quality and indicating any need for improvement to the respective Work Task Leaders.

Work Task Leaders (WTL): WTLs are responsible for the data collected / generated in the frame of the tasks that fall under their leadership as well as for safeguarding their appropriate and timely processing. Moreover, they are responsible for properly adjusting the Informed Consent Form and Data Subject Request Form templates, to the needs and specificities of the activities carried out in the task they are leading. WTLs are responsible for identifying the need for a specific privacy policy regarding the task they are leading and collaborate with the PC for drafting and releasing it to the public. Finally, they undertake any necessary actions to prepare the data collected / generated through the tasks they are leading for sharing either within the consortium or openly (including the

use of proper naming conventions, application of suitable anonymisation techniques, creation of appropriate metadata and documentation, etc.).

Partners: All project partners are tasked to collect, digitise, anonymise, store, destroy and / or otherwise process data for the specific purpose of the activity in which it has been collected / generated within the project. They are responsible for appropriately collecting the necessary consent for processing data as well as for ensuring that the Informed Consent Form and the Data Subject Request Form used to this end are properly adjusted to the needs of the activity they are participating (including references to the project's Privacy Policy and any other applicable specific privacy policies) and, in any particularities, applicable to their organisation while ensuring adherence to provisions of relevant national data protection legislation in their respective country. Moreover, they are responsible for managing the consents they have collected with a view to demonstrating their compliance with the relevant applicable EU and national regulation(s). Finally, they perform quality checks to assess and maintain the quality of the dataset(s) held within their records.

Data repositories: Data repositories are tasked with the storage and long-term preservation of the project's data. In this respect, Zenodo and GitHub will maintain and preserve the openly available data of AGIMUS, enabling its sharing and re-use. To this end, Zenodo assigns metadata and DOIs to the data, while also taking all necessary measures to securely back-up the data and restore it, safeguarding its long-term preservation. Accordingly, metadata will be included in the GitHub publicly available data, with state-of-the -art security ensured[37]. Large datasets will be stored at CTU 900TB disk array and will become either publicly available or available to project partners through password protected access. The public datasets will be linked from AGIMUS website.

In this context, the following table illustrates the allocation of data management responsibilities amongst the members of the AGIMUS consortium per data collected/generated under each WP.

Table 12: Data management responsibilities of AGIMUS' partners per data collected/generated under each WP

WP	WPL	Data	Tasks	WTL	Responsible Partners
1	TOWARD	Identification of needs, challenges, opportunities, data, and platforms for agile production	T1.1	Q-PLAN	ALL
		Identification of ethical, trustworthy, and secure design guidelines, for next-generation robotics for agile production	T1.2	INRIA	ALL
		Technical and non-Technical specifications during the co-design (technical) workshops	T1.3	TOWARD	ALL

[37] https://github.com/security

WP	WPL	Data	Tasks	WTL	Responsible Partners
2	INRIA	Motion Solver Toolkit Benchmark Dataset	T2.1-T2.3	INRIA / CNRS	INRIA, CNRS, CTU
3	CNRS	Sensor readings recorded during experiments (real or simulated)	T3.1	CTU	INRIA, CNRS, CTU
		Motion Dataset	T3.2	CNRS	INRIA, CNRS, CTU
		Videos of motions performed by human actors	T3.3	CNRS	INRIA, CNRS, CTU
4	CTU	WP4 6D object pose estimation training and evaluation	T4.1	CTU	CTU, CNRS, INRIA
		WP4 Collection of data for learning sensor-to-state predictive model	T4.3	CNRS	CNRS, CTU, INRIA
5	PAL	Robot performance benchmarking and stress-testing	T5.1, T5.2	PAL	CNRS, CTU, INRIA, PAL, TOWARD
		Visual Control Experiment Dataset	T5.3	INRIA	CNRS, CTU, INRIA, PAL, TOWARD
		Physical Interaction Experiment Dataset	T5.4	CTU	CNRS, CTU, INRIA, PAL, TOWARD
		Multi-step Task Planning Experiment Dataset	T5.5	CNRS	CNRS, CTU, INRIA, PAL, TOWARD
6	KLEEMANN	AIRBUS Use Cases Dataset	T6.1, T6.2	KLEEMANN	ALL
		KLEEMANN Use Cases Dataset	T6.1, T6.2	KLEEMANN	ALL
		THIMM Use Cases Dataset	T6.1, T6.2	KLEEMANN	ALL
		Cross-pilot Comparative Assessment Dataset	T6.3	AIRBUS	ALL
7	Q-PLAN	Website and social media analytics	T7.1	Q-PLAN	ALL
		Data collected from project events	T7.1, T7.4	Q-PLAN	ALL
		Data collected from dissemination and communication actions	T7.1	Q-PLAN	ALL

WP	WPL	Data	Tasks	WTL	Responsible Partners
		Newsletter subscriptions	T7.1	Q-PLAN	ALL
		Photos, number of event participants, presentations, joint activities	T7.1	Q-PLAN	ALL
		Business planning exercise	T7.3	Q-PLAN	PAL, TOWARD, AIRBUS, KLEEMANN, THIMM
8	CNRS	Setup and Operation of AGIMUS' Advisory Boards	T8.2	INRIA	ALL

6. Data security

AGIMUS will securely handle any collected / generated data throughout its entire lifecycle as it is essential to safeguard these data against accidental loss and / or unauthorised access. To achieve this the project will apply appropriate technical and organisational measures based on a risk assessment of the relevant data that takes into account the impact and the likelihood of a potential data breach. With that in mind, the project's data security strategy aims at minimizing the probability that a data breach will occur during the course and after the completion of AGIMUS, either from human error or hardware failure, as well as inhibit any unauthorised access. Particularly, in case of personal data collection / generation it is crucial that these **data can only be accessible by those authorised to do so**.

All project partners are responsible for processing[38] data using appropriate means, such as private servers or cloud service providers that adhere to the relevant legal data protection requirements (e.g., GDPR) and will ensure that these **data are protected**, and any **necessary data security controls have been implemented**, to minimize the risk of information leak and destruction. This case refers to the data that will be closed and therefore will not be shared and / or re-used within the framework of the project. In this case, to minimize the consequences of potential data losses, the data will be **backed up at regular time intervals based on change frequency and criticality. The backed-up files will be stored in appropriate storage media including external hard drives, flash drives, NAS devices and reputable cloud services**, so as to safeguard their preservation, while also enabling their recovery at any time. Moreover, **integrity checks**[39] will be carried out regularly ensuring that the stored data have not been changed or corrupted.

Access to closed data will only be permitted to authorised project partners. In case there is a personal data breach, the responsible project partner will notify, without undue delay and, where feasible, not later than 72 hours after having become aware of it, its competent national supervisory authority (e.g., data protection authority) as well as the data subject(s) that may be affected by the breach. Moreover, the responsible partner will document any personal data breaches, including information such as the facts relevant to the breach, its effects and the remedial action(s) taken.

Identification and authentication access controls play an important role in the context of the project, as they help partners to protect the data collected / generated during AGIMUS and especially personal data. To this end, each project partner is responsible for and committed to ensuring the application of appropriate access controls to the data they are processing. Finally, in order to safeguard the privacy of the users of the AGIMUS website, a dedicated **privacy policy** will define the

[38] Processing, according to Regulation (EU) 2016/679 of the European Parliament (General Data Protection Regulation), means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.

[39] An integrity check is the process of comparing the current state of stored data and/or programs to a previously recorded state in order to detect any changes.

way in which this online space collects, processes and uses personal data, the security procedures followed, the users' rights as well as the cookies policy employed.

On another note, openly available data will be stored safely for long-term preservation on Zenodo, in the same cloud infrastructure as research data from CERN's Large Hadron Collider, using CERN's battle-tested repository software INVENIO, which is used by some of the world's largest repositories (such as INSPIRE HEP and the CERN Document Server). Along these lines, data is stored and backed-up in CERN's EOS service in an 18 petabytes disk cluster. Both data files and metadata are kept in multiple online replicas and independent replicas ensuring their long-term preservation as well as their recovery when necessary. Moreover, for each file two independent MD5 checksums are stored. One checksum is stored by INVENIO, used to detect changes to files made from outside of it whereas the other checksum is stored by EOS, and used for automatic detection and recovery of file corruption on disks. In this context, access control is applied by the different level of openness that Zenodo allows (i.e., open, embargoed, restricted and closed).

Accordingly, data stored at GitHub can also be considered to be stored safely for long-term preservation. In general, GitHub is a GDPR compliant platform that ensures data privacy and security. GitHub is a Trusted Cloud Provider (™) with the Cloud Security Alliance (CSA)[40], whereas its Information Security Management System (ISMS) has been certified against the ISO/IEC 27001:2013 standard.

[40] https://cloudsecurityalliance.org/

7. Ethics and other issues

This Chapter addresses the ethical aspects of AGIMUS' Data Management Plan and the ethical compliance of the underlying data foreseen to be collected / generated under the project's activities. The project will process data that is not included in any special category of personal data (i.e., non-sensitive data) according to the relevant data protection legislation (e.g., GDPR). In accordance with the **Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (GDPR)**, all personal data processed for project's activities shall be:

- processed lawfully, fairly and in a transparent manner in relation to the data subject;
- collected for specified, explicit and legitimate purposes relative to project's objectives and not further processed in a manner that is incompatible with those purposes;
- adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed;
- accurate and, where necessary, kept up to date;
- kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed;
- processed in a manner that ensures appropriate security of the personal data (see section 6).

For all personal data processing activities within the framework of the project at least one lawful basis as of Art. 6 GDPR applies. Where informed consent is chosen as the lawful basis for processing, all relevant provisions of the data protection legislation (e.g., Art.7 GDPR) are observed. The project's Privacy Policy and the templates of the Informed Consent Form and the Data Subject Request Form, used in the implementation of the project's activities, are compliant with the General Data Protection Regulation and annexed to this DMP (see Annex). Last but not least, **no transfer of personal data outside the EU is foreseen as part of the project's implementation.** In case of data storage providers situated both inside and outside the EEA, partners are committed to ensure their compliance with the relevant GDPR requirements before start using their services.

It is important to highlight that each partner is responsible for ensuring that the templates for the Informed Consent Form and Subject Data Request Form (including references to the project's Privacy Policy and any other applicable specific privacy policies) are appropriately adjusted according to (i) the needs of the activity for which they are being used by them as well as to (ii) the relevant data protection laws and regulations applicable to their respective countries and / or organisation. All partners should keep records to demonstrate that data subjects have consented to the processing of their personal data and use consent management mechanisms that make it easy for individuals to withdraw their consent.

Finally, **no other national/funder/sectoral/departmental procedures for data management were used in the framework of AGIMUS**.

8. Conclusions and way forward

This initial version of the AGIMUS DMP aims at safeguarding the sound management of the data collected, processed and/or generated during the project's activities across their entire lifecycle, while also making them FAIR. It describes all the underlying processes of the AGIMUS data management, collection, process and generation, in accordance with the GDPR guidelines, and sheds light on (i) the data being collected, processed and/or generated under the project activities, (ii) the specific objectives under which each dataset is collected, processed and/or generated, (iii) the allocation of resources and data management responsibilities and (iv) the data security and ethical aspects of the data.

In the framework of AGIMUS, the DMP is a living document and is updated throughout the course of the project, considering its latest developments and available results. It is expected to be further developed and updated at least twice by the end the project (i.e., as D8.3 by M24 and as D8.4 by M48). If necessary, additional ad hoc updates may be realised in order to include new data, better detail and/or reflect modifications in the methodologies applied or other aspects relevant to data management (such as costs for making data FAIR, size of data, etc.), changes in consortium policies and plans or other potential external factors.

Annexes

Annex I – Privacy Policy

Within this Annex, the reader can find the draft project's overall Privacy Policy, which is uploaded on the project's website. Task leaders are responsible for developing any additional privacy policy needed in their tasks and activities in close collaboration with the PC. Moreover, they are responsible for appropriately adjusting and translating the templates to fit the needs and specificities of their task's activities.

PRIVACY POLICY

LAST UPDATE: < date of the privacy policy's last update in the form of DD/MM/YYYY >

1. Who we are:

AGIMUS is a project funded by the European Union's Horizon Europe Research and Innovation funding programme. AGIMUS aims to deliver an open-source breakthrough innovation in AI-powered agile production, introducing solutions that push the limits of perception, planning, and control in robotics, enabling general-purpose robots to be quick to setup, autonomous and to easily adapt to changes in the manufacturing process. To achieve such agile production, AGIMUS leverages on cutting-edge technologies and goes beyond the state-of-the-art to equip current mobile manipulators with a combination of (i) an advanced task and motion planner that can learn from online available video demonstrations; (ii) optimal control policies obtained from advances in reinforcement learning based on efficient differentiable physics simulations of the manufacturing process; as well as (iii) advanced perception algorithms able to handle objects and situations unseen during initial training. Along the way, optimization of energy efficiency and the use of 5G technology will support further pushing the limits of autonomy. The AGIMUS solutions and their impact will be demonstrated and thoroughly stress-tested in 3 testing zones, as well as 3 industrial pilots in Europe, under numerous diverse real-world case studies and scenarios (different tools, environments, processes, etc.). In every step, and from the very beginning, AGIMUS will go beyond current norms and involve a wide range of stakeholders, starting from the production line itself, to identify the essential ethical-by-design principles and guidelines that can maximise acceptance and impact.

AGIMUS entails several activities that involve collecting, producing, and processing data to generate meaningful insights that will feed into the project and fuel the co-creation and delivery of demand-driven and evidence-based results. Thus, the AGIMUS consortium partners, listed below, process certain types of personal data for the purposes of the project. Each partner is responsible for the personal data they collect and process during their activities under the framework of the project:

- CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE [CNRS] (Coordinator), https://www.laas.fr

- CESKE VYSOKE UCENI TECHNICKE V PRAZE [CTU], https://www.cvut.cz/en
- INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET AUTOMATIQUE [INRIA], https://www.inria.fr/en
- PAL ROBOTICS SL [PAL], https://pal-robotics.com/
- TOWARD SAS [TOWARD], https://toward.fr/en/
- Q-PLAN INTERNATIONAL ADVISORS PC [Q-PLAN], https://qplan-intl.gr
- AIRBUS [AIRBUS], https://www.airbus.com/en
- KLEEMANN HELLAS SA [KLEEMANN], https://kleemannlifts.com/
- THIMM OBALY, K.S. [THIMM], https://www.thimm.cz/

For further information, we can be contacted at: info@agimus-project.eu

2. How we collect your personal data

We collect your personal data both directly and indirectly:

Directly. We obtain personal data directly from individuals in a variety of ways, including but not limited to the following cases:

- an individual subscribes to our newsletter/s;
- an individual registers to attend in meetings and events we host and during attendance at such events;
- we establish cooperative relationships with an individual;
- we provide professional services pursuant to our contract with the European Commission;
- an individual participates in an interview or survey organized by us.

Indirectly. We obtain personal data indirectly about individuals from a variety of sources, including:

- our research partners;
- our networks and contacts;
- public and open data sources such as public registers, news articles and internet searches;
- social and professional networking sites (e.g. LinkedIn).

3. What types of data we collect?

We only collect the data that are necessary for the smooth implementation of our project. These data fall into the following categories:

- **contact details** (name/ surname, e-mail address, street address, mobile phone number, land line phone number);
- **professional information** (job title, organization, field of expertise);
- **demographics** (e.g. age, gender, nationality);
- **information about what a person knows or believes.**
- **videos and photos** (e.g. from people that attend our events).

4. Bases of lawful processing

We process personal data on the following legal bases:

Legal obligations - for processing activities required for compliance both with applicable national and European legislation as well as with the specific legal and regulatory framework of the European Union's Horizon Europe Research and Innovation funding programme.

Consent – for processing activities such as organization of surveys and interviews, completing of questionnaires and dissemination of project's results.

Contractual obligations - for processing activities such as reporting to the European Commission and complying with project's publicity obligations.

5. What we do with your personal data

We process your personal data with the purpose of:

- Conducting research (e.g., interviews, surveys);
- Disseminating our project's results to different types of stakeholders;
- Sending invitations and providing access to guests attending our events and webinars;
- Administering, maintaining, and ensuring the security of our information systems, applications, and websites;
- Processing online requests or queries, including responding to communications from individuals;
- Complying with contractual, legal, and regulatory obligations.

6. How we secure your personal data when we process it

We continuously apply a personal data risk assessment process to identify, analyse, and evaluate the security risks that may threat your personal data. Based on the results of this risk assessment, we define and apply a set of both technical and organizational measures to mitigate the above security risks, including but not limited to:

- Data Protection Policies to guide our personnel when processing your data;
- Written contracts with organizations that process personal data on our behalf;
- Non-Disclosure Agreements with our personnel;
- Back up process, antimalware protection, access control mechanisms, etc.
- Some of our partners have appointed a Data Protection Officer.

7. Do we share personal data with third parties?

We may occasionally share personal data with trusted third parties to help us deliver efficient and quality services. When we do so, we ensure that recipients are contractually bound to safeguard the data we entrust to them before we share the data. We may engage with several or all the following categories of recipients:

- Parties that support us as we provide our services (e.g., cloud-based software services such as Dropbox, Microsoft Sharepoint, Google, ownCloud);
- Our professional advisers, including lawyers, auditors, and insurers;
- Dissemination services providers (e.g., MailChimp);
- Law enforcement or other government and regulatory agencies or other third parties as required by, and in accordance with applicable law or regulation;
- The European Commission according to our relevant contractual obligations.

8. Do we transfer your personal data outside the European Economic Area?

We do not own file servers located outside the European Economic Area (EEA). However, some partners may use cloud and / or marketing services from reputable providers such as SharePoint, DropBox, MailChimp, Google, etc., situated both inside and outside the EEA. We always check that such providers comply with the relevant GDPR requirements before start using their services.

9. Do we use cookies?

Our website uses cookies. Where cookies are used, a statement will be sent to your browser explaining the use of cookies. Cookies are small text files which are saved on your computer, mobile phone or tablet. They allow the website to remember your actions and preferences (such as login, language, font size and other display preferences) so you don't have to keep re-entering them whenever you come back to the site. You can control and/ or delete cookies as you wish. If you do this, however, you may need to manually adjust your preferences every time you visit a site. For more information on how to manage cookies, please visit: http://www.aboutcookies.org/

We use tools like Google Analytics to better understand how visitors interact with our website. This provides us with important information to enable the site to work better. The information collected is

not linked to your personal data. For more information on the cookies set by Google Analytics, please visit: http://code.google.com/apis/analytics/docs/concepts/gaConceptsCookies.html

The following cookies are used by Google Analytics:

Name	Typical content	Cookie expires after
_ga	Used to distinguish users	2 years
_gat	Used to throttle request rate	1 minute
_gid	Used to distinguish users	24 hours

We use cookies that identify you when you're logged in, in order to give you the possibility to comment using your profile and also to edit your newsletter preferences. We also use anonymous cookies that tell us where our users come from and what pages they visit. These cookies sometimes collect anonymous statistics about the user (such as gender, age, interests) and the data is stored by the analytics services that we use: Google Analytics, Content Insights and Facebook Analytics. These cookies also gather data about a person's behaviour on our website: what pages they visit, how long they stay on the page, what videos they watch or files they download. We also use tracking cookies from social networks such as Facebook, Twitter, Youtube and LinkedIn in order to be able to serve better advertising to our users on these platforms and to assess the performance of our ads on these platforms. The data that these platforms collect is anonymized. That means that we cannot see your social media profiles, but you will sometimes see ads to AGIMUS pages advertised to you on social media.

10. Your rights

You have the following rights regarding our processing of your personal data:

- **Right to withdraw consent** – You can withdraw consent that you have previously given to one or more specified purposes to process your personal data. This will not affect the lawfulness of any processing carried out before you withdraw your consent.
- **Right of access** – You can ask us to verify whether we are processing personal data about you and, if so, to have access to a copy of such data.
- **Right to rectification and erasure** – You can ask us to correct our records if you believe they contain incorrect or incomplete information about you or ask us to erase your personal data after you withdraw your consent to processing or when we no longer need it for the purpose it was originally collected.
- **Right to restriction of processing** – You can ask us to temporarily restrict our processing of your personal data if you contest the accuracy of your personal data, prefer to restrict its use rather than having us erase it, or need us to preserve it for you to establish, exercise or defend a legal claim. A temporary restriction may apply while verifying whether we have overriding

legitimate grounds to process it. You can ask us to inform you before we lift that temporary processing restriction.

- **Right to data portability** – In some circumstances, where you have provided personal data to us, you can ask us to transmit that personal data (in a structured, commonly used, and machine-readable format) directly to another entity.
- **Right to object –** You can object to our use of your personal data for direct marketing purposes, including profiling or where processing has taken the form of automated decision-making. However, we may need to keep some minimal information (e.g., e-mail address) to comply with your request to cease marketing to you.
- **Right to make a complaint to your local Data Protection Authority** (DPA) (see https://ec.europa.eu/justice/article-29/structure/data-protection-authorities/index_en.htm) regarding any concerns you may have about our data handling practices.

To ask us to do anything of the above, you can contact us by email: info@agimus-project.eu. We will promptly examine your request against the relevant requirements of the laws and regulations governing privacy and personal data protection and we will answer the latest within 30 days after receiving your request. We will ask from you some kind of identification (e.g., photocopy of your identity card or passport) to avoid non-authorized reveal of your personal data. If, for reasons of complexity of the request or a multitude of requests, we are unable to respond promptly, we will notify you within 30 days of any delay, which in no case may exceed two months from the expiration of the 30-day deadline.

11. How long do we retain personal data?

We retain personal data to provide our services, stay in contact with you and to comply with applicable laws, regulations, and contractual obligations to which we are subject. Please note that we have an obligation to retain data concerning projects funded by the European Union's Horizon Europe Research and Innovation funding programme for up to five years after the end of the project (unless further retention is requested by auditors). After the expiry of the retention period, and unless further legitimate grounds for retention arise, we will dispose of personal data in a secure manner.

12. Disclaimer of liability for third party websites

Although our site may contain links to third-party sites, including the sites of the consortium partners, we are not responsible for the privacy practices or content of these sites, and we expressly disclaim any liability for any loss or damage that may be caused by the use of these links. We do not monitor the privacy practices or the content of these sites. If you have any questions about the privacy practices of another site, you should contact the site's responsible personnel. We suggest you read the privacy policy of each website you interact with, before allowing the collection and use of your personal data.

We may also provide social media features that allow you to share information on your social networks and interact with our project on various social media sites. The use of these social media features may result in the collection or sharing of information about you. We recommend that you check the privacy policies and regulations of the social networking sites you interact with, so that you can be sure that you understand what information may be collected, used and disclosed by these sites.

13. Children

We do not knowingly collect, use, or disclose information from children under the age of 16. If we learn that we have collected the personal information of a child under 16 we will take steps to delete the information as soon as possible. Please immediately contact us if you become aware that a child under 16 has provided us with personal information.

14. Revisions of this Privacy Policy

This Privacy Policy is valid from < date of the policy's last update in the form of DD/MM/YYYY > and replaces any other previous notifications that we had issued in the past regarding our personal data management practices. We reserve the right to revise this Policy at any time. The current version will always be uploaded to our website indicating the date of entry into force, so you know when the most recent revision took place. If there are critical changes in this Policy or our personal data practices change significantly in the future, we will notify you by posting the changes on our website.

Annex II – Informed Consent Form

Text in red colour contains guidelines for adjusting this template and should be deleted.

Text included in < > and/or highlighted with yellow should be replaced with content that is suitable to the context of each activity & project as well as to the organisation seeking to obtain the consent.

Before using this template take the time to carefully read and adjust it to the needs of the activity at hand as well as to any relevant regulations and particularities applicable to your country and organisation. Specifically, this template has been developed for a data processing activity involving interviews, however you can easily adapt it to other common data processing activities such as surveys, events organizing, etc.

INFORMED CONSENT FORM

Who we are:

We are < Insert Partner Name > and we are contacting you in the framework of AGIMUS, a project funded by the European Union under the Horizon Europe Research and Innovation funding programme. A detailed description on how AGIMUS handles personal data is presented in the project's Privacy Policy that accompanies this Consent Form.

Project: AGIMUS - Next generation of AI-powered robotics for agile production (GA Number 101070165).

Partner:

Organisation name: < Insert Partner Name >

Address: < Insert Partner Address >.

Phone: < Insert Partner Phone >.

E-mail: <Insert Partner Generic E-mail Address >

Responsible persons:

You may delete the line referring to the Data Protection Officer if your organisation does not have one.

#	Role	Name	E-mail
1	AGIMUS Project Manager	<Insert name of project manager>	<Insert email of project manager>
2	Interviewer	<Insert name of interviewer>	<Insert e-mail of interviewer>
3	Data Protection Officer	<Insert name of DPO >	<Insert e-mail of DPO >

What do we need from you?

Please explain in a brief paragraph (4-5 lines) the activity and its purpose under the frame of the project. You can find below an indicative example.

Example: We need you to participate in an interview that will be carried out by AGIMUS with a view to monitor and assess AGIMUS' outcomes, impacts and perceptions change towards the new agile production paradigm.

The interview is expected to last for no more than < Insert number of minutes > minutes. We will take written notes and we will be making a sound recording of the interview.

Please adapt the following text to accurately depict the type of personal data to be collected.

To effectively conduct this interview, we need to process some of your personal data:

- Your contact details (full name, email, phone number);
- Some basic demographics (age, gender);
- Your professional info (organization, job position, field of expertise);
- Your opinions on the subject matter.

Why do we need your data & what will we do with them?

We need your data to contact you in order to plan and carry out the aforementioned interview and to resolve any ambiguities, questions and other issues that may arise after and as a result of the interview. We also need to record your data to keep track of the interview process. The project's deliverables that will be derived by the interview will not include your personal data or any other information that could identify you. Your personal data will remain on our written notes (interview's transcript) and/or the sound recording we will make during the interview.

We will share your data with a few other AGIMUS project partners that are also involved in this task and will participate in the drafting of the relevant deliverables. We are also obliged to grant access to your data to:

- EU officials such as our Project Officer for purposes related to project's evaluation;
- EU agencies and other authorities for project's auditing purposes.

We would also be very happy if you gave us your consent to contact you in the future to ask you to participate in other project's activities (e.g., surveys, interviews, project events, AGIMUS Industrial Advisory Board, etc.) and also to inform you about the project's progress (e.g., by sending you a newsletter or similar messages).

How can you withdraw your consent?

You should know that you can withdraw your consent at any time by communicating either on the phone or by email with the responsible persons listed in the previous page. With regards to the informational messages and newsletters you can always opt out by simply clicking the link "Unsubscribe" or something similar included at the end of all the relevant messages.

I hereby give my consent to the processing of my personal data needed for:

*(Please, tick the boxes below to confirm that you give us your consent for the respective subject. Any boxes left unticked mean that **you do not consent to the relevant subject**.)*

#	Consent Subject	Tick box
1	My participation in an interview that will be carried out by AGIMUS to < insert key objective of the interview >	
2	My participation in future activities of AGIMUS	
3	Receiving newsletters and messages regarding AGIMUS activities	

Name of participant _____ Date _____ Signature _____

Annex III – Data Subject Request Form

Text in red colour contains guidelines for adjusting this template and should be deleted.

Text included in < > and/or highlighted with yellow should be replaced with content that is suitable to the context of each activity & project as well as to the organisation seeking to obtain the consent.

AGIMUS

Data Subject Request form

You may delete the data referring to the Data Protection Officer if your organisation does not have one.

Contact

<Insert name of responsible Project Manager>	<Insert name of DPO > **(Data Protection Officer)**
<Insert email of responsible Project Manager>	<Insert e-mail of DPO >

Data Subject Request Form

This form should be used to submit a data subject request under the provisions of the European Union General Data Protection Regulation (GDPR).

Submitter Details

Title:	
Name:	
Address:	

Type of Request

Please select the type of request you are making:

☐ *Consent Withdrawal*

☐ *Access request*

☐ *Rectification of personal data*

☐ *Erasure of personal data*

☐ *Restriction of processing of personal data*

☐ *Personal data portability request*

☐ *Objection to processing of personal data*

☐ *Request regarding automated decision making and profiling*

Personal data involved

Request details

Request reason/justification

Name: .....

Signature: .....

Date: .....

***

Once completed, this form should be submitted via e-mail to < Insert contact e-mail of Partner > or posted to:

< Insert Partner Name >

< Insert Partner Address >