This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº 101017331

D1.3 Data Management Plan v2

Deliverable No.	D1.2	Due Date	31.01.2023
Description	The Data Management Plan provides the main principles, as well as ethical and data protection strategies adopted within the project regarding the data management, knowledge and IPR issues. It also includes the data management strategies of each consortium partner and pilots. This is a living document that will be constantly updated during the lifetime of the project according to the changes that might happen at a project level and at the pilots' level.		
Type	Report	Dissemination Level	PU
Work Package No.	WP1	Work Package Title	Project Management and Coordination
Version	1.0	Status	Final

Authors

Name and surname	Partner name	e-mail
Adrian Quesada Rodriguez	UDGA	aquesada@udgalliance.org
Vasiliki Tsiompanidou	UDGA	vtsiompanidou@udgalliance.org

History

Date	Version	Change
09/02/2023	0.1	Creation of initial draft
28/02/2023	0.2	Finalization of structure and general content
15/03/2023	0.3	Inclusion of pilots initial data flows
31/03/2023	0.4	Introduction of questionnaire inputs
07/04/2023	0.5	Expansion on the questionnaire findings
28/04/2023	0.6	Finalization of the questionnaire findings
08/05/2023	0.7	Introduction of the pilots data flows
15/05/2023	0.8	Finalization of the deliverable and submission to peer review
18/05/2023	0.9	Peer review
23/05/2023	1.0	Final Deliverable

Key data

Keywords	Data; Data Protection; Data Management; FAIR Data; IPR; Ethics; Privacy
Lead Editor	Adrian Quesada Rodriguez; Vasiliki Tsiompanidou
Internal Reviewer(s)	Daphne Plati (FORTH); Ilias Kalamaras (CERTH)

Abstract

This document describes the Data Management Plan (DMP) and serves as a guide for the partners of the ODIN project. This deliverable is the second version of the DMP, which outlines the data processing activities of the partners, the security measures implemented, as well as their IPR and FAIR data strategies. Further, this document identifies the data which will be generated during ODIN's execution, and the already existing data used within the tasks, as well as the data flows where data is to be shared within the Consortium. The findings leverage on inputs provided by the consortium, as the partners and pilots' managers will continue to update this deliverable. A final version of the DMP will be presented at the very end of the project and will include conclusions on the data processing and ODIN consortium agreements.

Statement of originality

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Acronyms

Acronym	Definition
DMP	Data Management Plan
DoA	Description of Action
DPIA	Data Protection Impact Assessment
DPO	Data Protection Officer
EDPS	European Data Protection Supervisor
EEA	European Economic Area
EU	European Union
GDPR	General Data Protection Regulation
GPL	General Public Licence
FLOSS	Free/Libre and Open Source Software
IoT	Internet of Things
IPR	Intellectual Property Rights
LSPs	Large Scale Pilots
PS	Pilot Site
WP	Work Package

Table of contents

List of tables

List of figures

Executive Summary

The current document presents the second iteration of the Data Management Plan (DMP) designated to the partners of the ODIN project for their data processing-related activities, as well as for the data processing activates in the different hospital use cases (pilots). The plan specifies the Data Governance and handling of personal and sensitive data during the project activates; outlining what types of data are expected to be generated and used, if and how it will be shared and made accessible internally and, after the lifetime of the project, externally for verification and re-use. The plan explains how partners intend to store and protect data, considering, in particular, ethical, privacy, and security issues. The guidelines on collecting and characterizing datasets follow the ethics, privacy and legal framework delivered in D8.2. All findings are based on the answers provided by the partners to dedicated Data Management Questionnaires.

The DMP covers the entire research data life cycle and is consistent with exploitation and Intellectual Property Rights (IPR) requirements, while at the same time respecting FAIR (Findable, Accessible, Interoperable and Reusable) data principles. Sensitive and personal data of patients and participants will be kept strictly confidential and either anonymized or pseudonymized, to maintain compliance with General Data Protection Regulation (GDPR).

1 Introduction

1.1 The Odin Project: Overview

The ODIN project focuses on identified hospitals' critical challenges which will be faced by combining robotics, Internet of Things (IoT) and artificial intelligence (AI) to empower workers, medical locations, logistics and interaction with the hospital's territory. According to their expertise, the project's consortium has divided its management responsibilities into the areas showcased below in *Figure 1*. ODIN's aspiration to enhance healthcare for patients, leveraging on AI, robotics, emerging techniques, approaches and methods results in critical ethical and data protection issues, such as potential harms to autonomy, dignity, privacy, moral responsibility, equality, transparency, safety, accountability, and liability. To meaningfully address these challenges and develop a commonly followed strategy for risk mitigation and compliance, the project management and coordination work package (WP1) has dedicated a task (T1.4) to Data Management and Ethics.

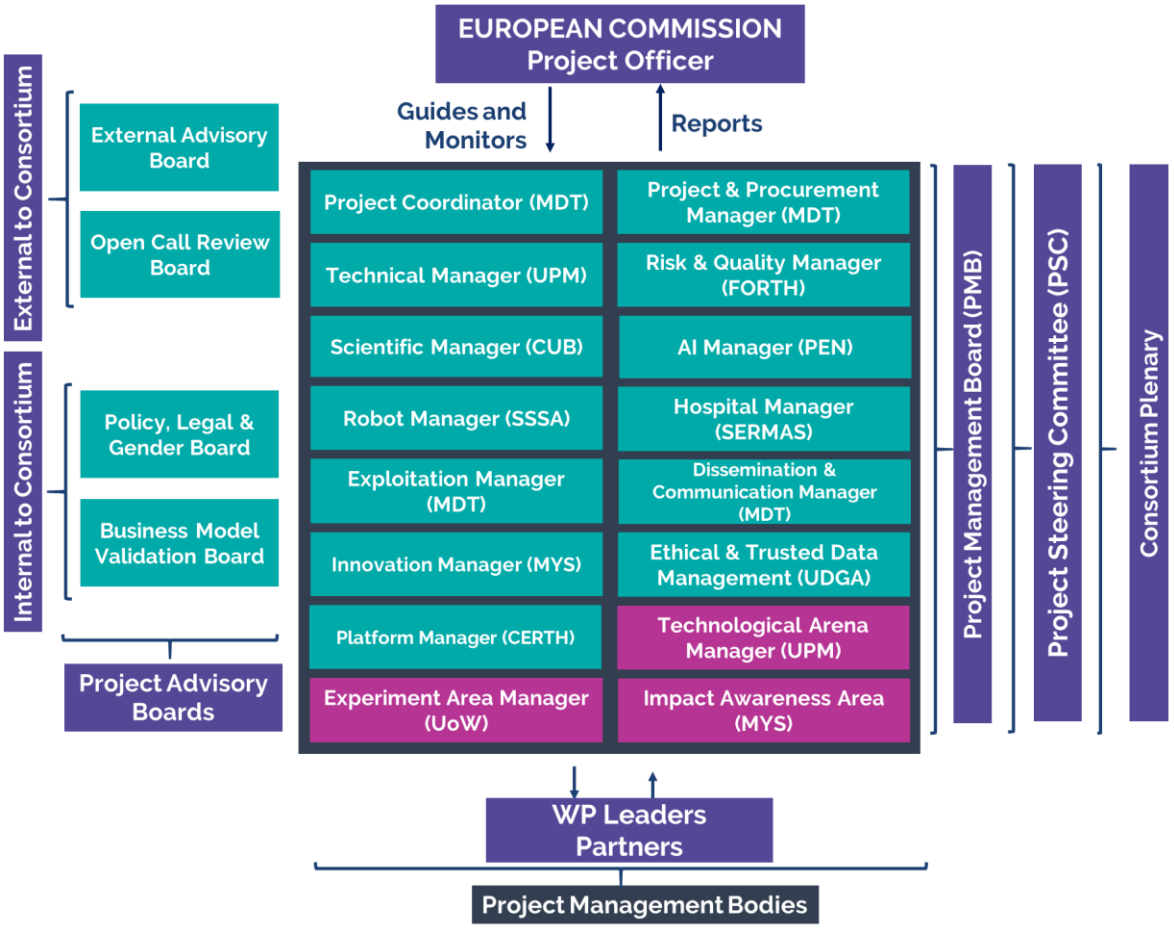

Figure 1: The position of D1.2 in ODIN Management

1.2 About this Deliverable

The European Commission defines[1] Data Management Plans (DMPs) as key elements of good data management. A DMP should describe the data management life cycle for the data to be collected, processed and/or generated. As part of making research data findable, accessible, interoperable and re-usable (FAIR), a DMP should include information on:

- The handling of research data during & after the end of the project;
- What data will be collected, processed and/or generated;
- Which methodology & standards will be applied;
- Whether data will be shared/made open access;
- How data will be curated & preserved (including after the end of the project).

Deliverable D1.2 is a plan for ethical and GDPR-compliant data management among the ODIN consortium. It is a product of task T1.4 "Data Management and Ethics" and aims at summarizing the data to be generated within the project and the envisioned data processing. The current deliverable is the first version of the plan and is produced at month 10. It is a recurrent live deliverable, which will be constantly updated according to partners' inputs and will reflect on any changes regarding data generation, usage, processing, storage, and ethical management. A second version in month 24 will update the data processing activities and include the initial plan for exploitation and preservation. The deliverable's final version, which will be produced in month 42, will include the final work done in terms of data processing in alignment with the ODIN consortium agreements.

1.3 Context of the Deliverable

Table 1. Deliverable context

PROJECT ITEM IN THE DOA	RELATIONSHIP
Project Objectives	In providing a management plan for compliant handling of data in the scope of the ODIN project, the DMP contributes to the realization of the objectives, tightly dependent on a compliance with legal, ethical and security frameworks. In particular, the DMP suffices the requirements in O1 to support the interoperable and effective implementation of the decentralized ODIN platform; O2 to guarantee the delivery and scale-up of innovative services in accordance with national and European legal frameworks; and O4 to help set up an exploitation strategy for data, specifically in terms of delivery to European Data Space.
Exploitable results	The deliverable presents a model for data management and will serve as a reference for the consortium. Exploitation strategy for open access data can be further explored outside the consortium.

[1] https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm.

Workplan	The deliverable will be constantly updated according to the DoA. Our partners will be encouraged to provide constant up-to-date inputs regarding their data processing activities. Pilots' progress in this regard will be monitored and documented.
Milestones	D1.2 is a key deliverable for the Preparation milestone.
Deliverables	D1.2 defines the Data Management plan of the project. It is also connected to other deliverables such as D8.2, D11.1, D11.2, and D11.3.
Risks	The constant update of information from the pilots will need to be monitored. Particular attention will need to be devoted to data sharing issues.

1.4 Methodology

The Data Management plan combines both data protection and ethical aspects of the compliant handling of data, lays down regulatory principles, and identifies best practises. For the purposes of the deliverable, a dedicated ethics and data protection questionnaire has been developed (Appendix A). The questionnaire consists of six parts: 1) Part A required partners to describe the datasets that are either generated and/or collected and/or processed within the context of ODIN (personal or non-personal), 2) Part B, focusing on the measures implemented to make data FAIR (Findable, Accessible, Interoperable and Re-usable), 3) Part C, analysing any cost implications that making data FAIR would entail, 4) Part D, requiring that the partners describe in detail the security measures implemented to ensure data remains safe, 5) Part E, expanding on any ethical or legal issues identified by your organisation that have impacted or might impact data sharing, and 6) Part F, requiring an analysis of the Intellectual Property Rights (IPR) brought to the project or generated through it, as well as the steps to exploit it. The representatives of each partner organization within the ODIN consortium received the questionnaire in a digitalized form and were asked to provide their inputs.

These interactive activities are complemented by research on the GDPR and other applicable EU legislation (as comprehensively mapped in the previous version of this deliverable and D8.2). The data management plan acknowledges and takes into account any complementary EU Member State legislation around the processing of special categories of data and demands declaration from partners that they abide by these rules and thus efficiently safeguard data subjects' rights.

The figure below (Figure 2) describes the creation of the Data Management Plan. As outlined above, this is a living document, which incorporates best practices and main principles in the field of data protection and ethics; it offers mitigation strategies for various issues, including IPR management. Being a living document, the DMP presents the current state of partner's data management activities and anticipated actions. Identified issues and the effectiveness of applied mitigation measures will be further evaluated in the final iteration of this deliverable.

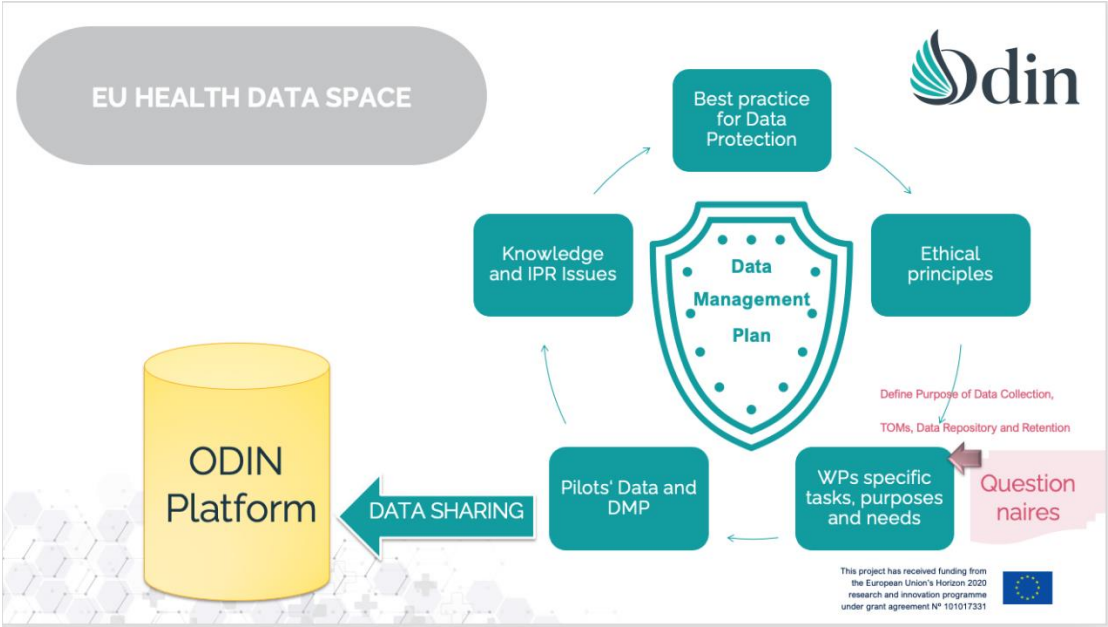

Figure 2: DMP Methodology

2 Data summary

2.1 Data generation and data flows in ODIN

2.1.1 Purpose of Data Generation and Relation to Objectives

In the context of the ODIN project, personal data are processed with the aim of working on solutions with technologies for the better quality of life and (health) care. The purpose of the treatment is to carry out the management of stakeholder participation in the project. Likewise, the data may be processed to develop ODIN's own dissemination activities or to send information about participation in the project-to-project users.

Personal data of participants will only be used for the development of the implementation in the region where the ODIN project is developed, being stored with all the possible guarantees of confidentiality and privacy. Figure 3 below demonstrates the information flow within the project.

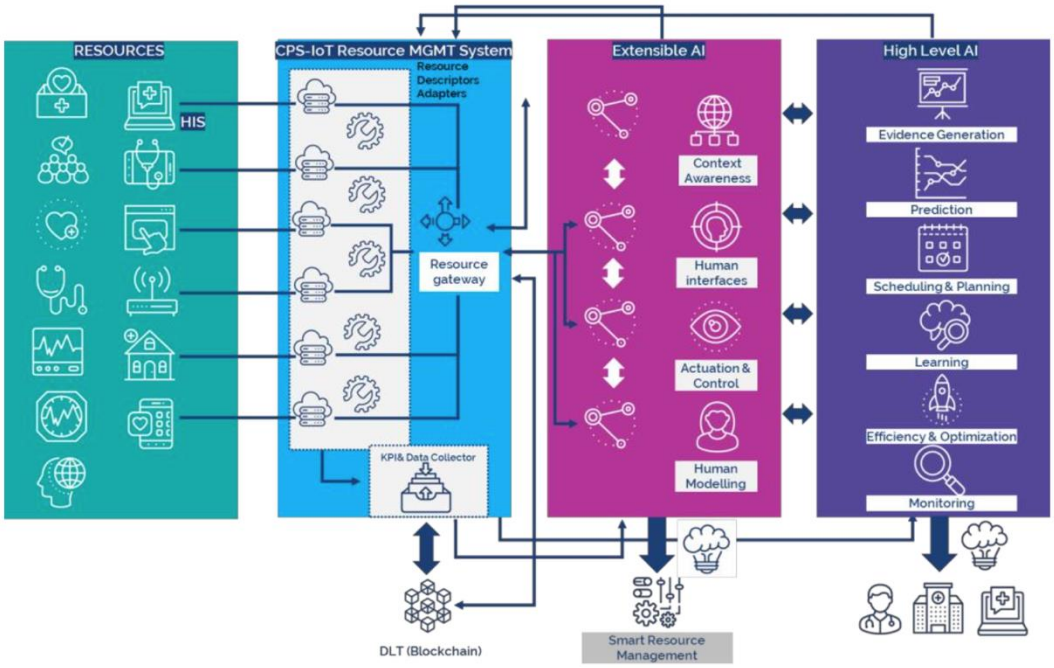

Figure 3: Information Flow in ODIN

2.1.2 Type and Format of Generated Datasets for ODIN

Data in the context of ODIN are collected at the project level by non-pilot owners and at the pilots' level by pilot-owners. In the context of ODIN, data is collected for production of deliverables, for training AI algorithms and, consequently, robots, for initial analysis of requirements and use-case catalogue production, as well as for model validation. The following table presents the datasets which each partner will generate, irrespective of their personal data protection relevance. The established interconnections would facilitate data mapping, which is particularly relevant in terms of data sharing within the consortium. Detailed information about the generated data by each partner is provided in the submitted Data Management Questionnaires.

Table 2: Data generated by the partners in different work packages for their deliverables

Partner	Work Package/ Task	Asset	Type & Format
MDT	WP2, T2.1	Interviews 1:1 with hospitals	.docx, .pptx
MDT	WP2, T2.1	Requirements questionnaire	.xlsx
MDT	WP2, T2.1	Stakeholder mapping workshop	Miro board (online); .docx, .pptx
MDT	WP2, T2.5	Pilot Sites Legislation on Public Procurement	.docx
MDT	WP2, T2.5	Form to capture needs, requirements, and problems towards Public Procurement Processes	.docx
MDT	WP2, T2.5	Public procurement form from suppliers	.docx
MDT	WP8, T8.2	A report summarising applicable standards to ODIN, and a standardisation and certification strategy, plus a sustainability plan, together with a general overview on the relevance of standardisation and certification to introduce the topic to Hospital partners	.docx, .ppt
MDT	WP9	Interviews 1:1 and workshops with hospitals and relevant partners to understand the needs for exploitation of project's results	Exploitation plan of each partner: .docx, .pdf, .xlsx
MDT	WP10, T10.	ODIN Community of Interest	ODIN Website, .pdf, .xlsx, .docx
MDT	WP10, T2.2	Focus Group on Public Procurement with hospitals	.docx, .pdf
MDT	WP10, T2.3	Focus Group on Public Procurement with suppliers	.docx, .pdf
MDT	WP10, T10.4	Open Call submission portal	ODIN Website, .pdf, .docx, .xlsx

CERTH	WP5, T5.2	Video format, RGB image format, Pointcloud format from depth sensors	either .mp4 or .avi, .png, and .pcd.
FORTH	WP6, T6.1, T6.2, T6.3	Pilot Data	JSON data types in either excel format, in SQL databases, in EHR systems, in xml format
UoW	WP3, WP6, WP7	Information referred to the pilots' representatives, their experiment definition and the procedure followed by each partner to obtain the ethical approval	.doc; .pdf; xcl
SSSA	WP5	Data coming from cameras and used for human awareness, robot navigation, human detection and tracking, social interaction models, monitoring and security, human action and behavioural recognition, human-robot interaction, etc.	Data coming from sensors or HMIs (e.g., images); digital data
SSSA	WP5	Data coming from sensors for localization of devices and robots that could be transported by people or wearables for cognitive performance monitoring and user's state estimation (stress, cognitive load, sleep quality, etc.)	Data coming from sensors or HMIs (e.g., images); digital data
SSSA	WP5	Sensitive data of patients and workers that are transmitted/processed through robotic modules that come from human-machine interfaces installed in the robots (for accessing to services or registrations) or coming from the Hospital's ICT infrastructure (other WP's)	Data coming from sensors or HMIs (e.g., images); digital data
ROBOTNIC		Data related to the movement of the robot and its commands	JSON format
MYS	WP4	Technical requirements that are needed to achieve Use	Free text from datasheet surveys.

		Case objectives from each Work Package. Opinions about the type of documentation and support service levels that partners can offer. Designs of the ODIN architecture.	Free text documentation and images of the designs.
THL	WP5, T5.3	Technical data related to robots' operation and status	Robotic data formats will be custom defined through the ROS messaging and service interface (.msg and .srv file format)
PEN	WP6	Analytics and AI in specific clinical use cases to process de-identified patient data and relevant process and administrative data. The output of the work will be models.	Not defined yet.
UPM	WP2, T2.4	Participant data collected through interviews and workshops	Text documents
UPM	WP3, T3.3	User data collected for identity management (i.e., credentials)	JSON or another interoperable format
UPM	WP4, T4.6	Data related to metrics such as logs, usage stats of the platform	JSON or another interoperable format
UPM	WP5	Data related to interaction of users with social robots	JSON or another interoperable format
UPM	WP7, T7.1	Data from pilots collected through questionnaires	Text documents
UPM	WP10, T10.3	Data from open calls submissions (participant forms, project specification, etc.)	Digital format, not yet defined
UCBM	WP7	Robot data: robot data recorded during testing, debugging and verification of the developed software modules (e.g., positions, velocities, forces, torques, RGB-D camera data).	Possible format will include .csv or .txt.

UCBM	WP7	Physiological data: physiological data recorded during testing, debugging and verification of the developed software modules (e.g., electromyography, galvanic skin response, heart rate, respiration rate).	Possible format will include .csv or .txt.
UCBM	WP5, WP7	Patient ID and HIS data: patient data (e.g., ID, pathologies, allergies, intolerances, diet) extracted from the HIS for testing, debugging and verification of the developed software modules.	Possible format will include:.xslx or .json or .csv.
UMCU	WP7	UC3: Generation of patient data from patients using the Luscii app for monitoring	.cvs
SERMAS	WP7	UC1: Information concerning material and equipment consumptions and purchases for a yet to be defined medical procedure	CSV file
SERMAS	WP7	UC1: Data from patients who undergo the yet to be defined medical procedure	CSV file
SERMAS	WP7	UC2: Internal data from a robot used to transport materials from a storage room to an operation room	To be defined.
SERMAS	WP7	UC7: Video image of a hospital area (either the emergency service or a surgical area), geographical position of equipment and personnel/ patients (RFID)	Video files.
CUB	WP7, WP8, WP9	Questionnaires distributed to stakeholders, students, and pilot participants to collect medical data and economic data.	Text data on paper; .pdf, textual data from medical records; EDF format data from sleep recording equipment
MUL	WP7	Architectural data from the hospital administration	To be defined

MUL	WP7	IoT data from tagging devices	To be defined
MUL	WP7	Data related to clinical staff and patients, i.e., the final users of equipment and consumables	CSV Comma Separated Values, XLS Excel Spreadsheets
MUL	WP7	EHR data, originating from the P1 EHR system, made available under nationwide P1 universal eHealth system, currently under implementation	SNOMED, ICD 10, ICD 9
M&S	WP9, T9.1 WP10, T10.1	Contact list	.xls/.xlsx, .csv
M&S	WP10, T10.1	Information on similar projects	.xls/.xlsx
M&S	WP10, T10.1	Data on supply and demand of ODIN related products	.xls/.xlsx
M&S	WP10, T10.1	Questionnaires for Trust building and Ecosystem enlargement	.xls/.xlsx
UDGA	WP1, T1.4	Questionnaires for information on partner data management activities	.docx; .pdf
UDGA	WP8, T8.3	Partner inputs on certification demand	.docx; .pdf
UDGA	WP8, T8.4	Partner inputs on data ethics for hospital procurement	.docx; .pdf
MEDEA	WP7, T7.2, T7.5, T7.7	Data referring to technological components provided by the partners; • the viewpoints of top-managers, lead users (doctors, nurses, technical staff) and end-users (patients and relatives), including user experience, user acceptance, usability, ergonomics, safety and ethics aspects • the impact on hospital management and cost effectiveness of the	.xlxs or .docx

		solutions according to specific identified KPIs	
MEDEA	WP9, T9.2	Data for the PESTLE analysis to analyse events and trends in areas that commonly affect business operations and performance	.xlxs or .docs

2.1.3 Pilots' Data Flows

In order to create a more comprehensive image of the data flows within the pilots, UML drawings were designed, as below:

A. CUB

CUB will be collecting and sharing data with partners in four instances, as can be demonstrated below. Data related to sleep disorder management is collected, also including sensitive personal data, and shared with Philips in a complete anonymized format for the purpose of training the algorithm. Data will be anonymized manually by qualified personnel within CUB, removing all personal data and identifiers that can lead to the re-identification of the data subjects.

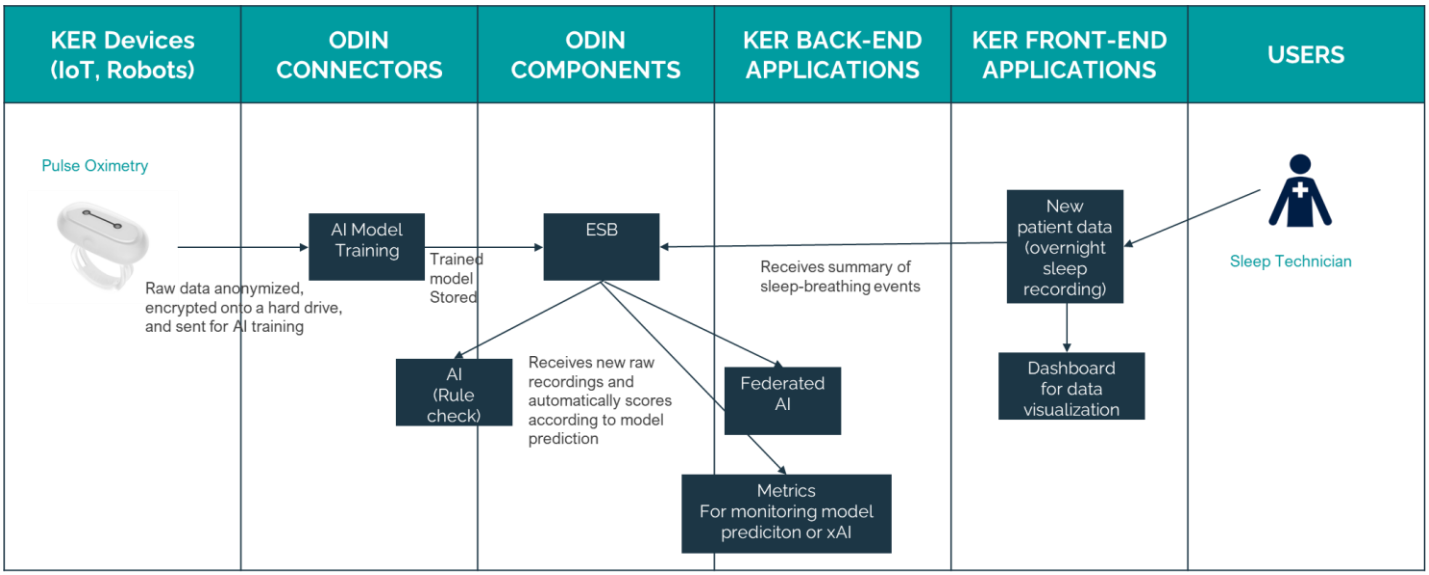

Figure 4: RUC A: Sleep disorder management through Pulse Oximetry Data Flow

Similarly, in the scenario of the sleep disorder management through Polysomnography, CUB will be sharing data with Philips for the training of the algorithm and with CERTH for the training of its analytics and visualization tools. All data will be strictly anonymized through a manual procedure performed by qualified personnel within CUB, removing all personal data and identifiers that can lead to the re-identification of the data subjects.

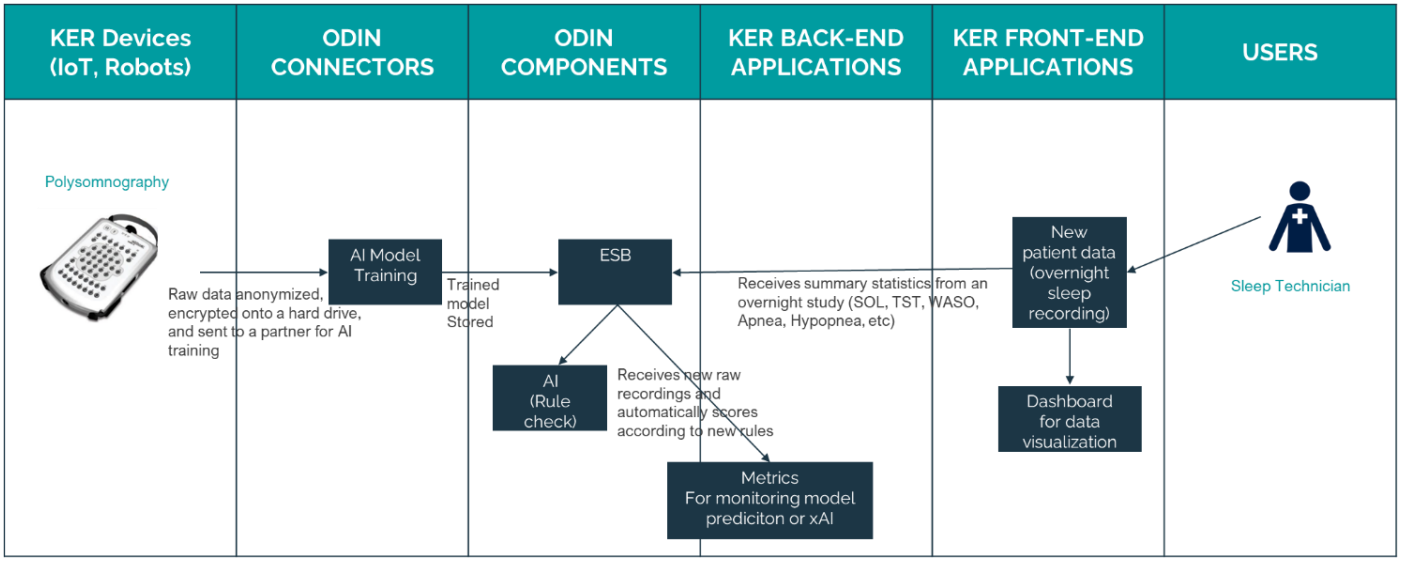

Figure 5: RUC A: Sleep disorder management through Polysomnography Data Flow

Additionally, data referring only to the number of the patient ward will be shared with the ODIN platform and the CERTHbot. Patients, if they desire so, can be provided with a QR code containing their ward number that can only be scanned by the CERTHbot, so that it can guide them to their room.

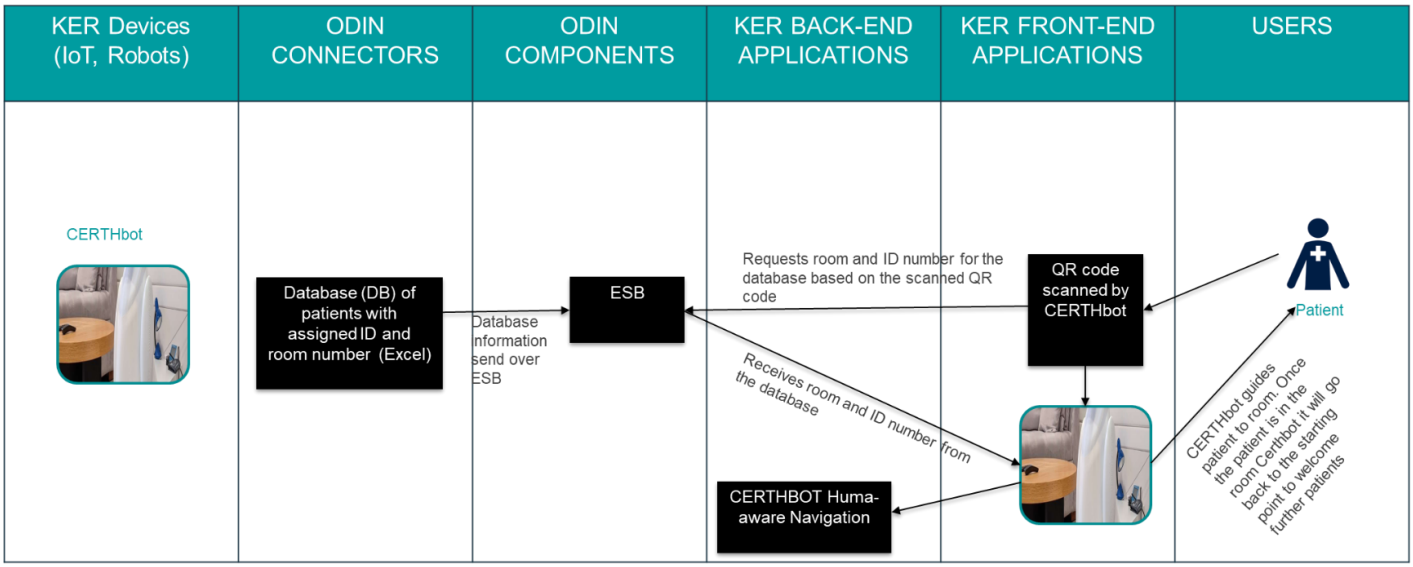

Figure 6- RUC A: UC 4 CERTH Bot Receptionist Data Flow

Finally, CUB will be using the CERTHbot in order to identify and inform the CUB personnel of abnormal behaviors during sleep. The bot will be creating an alert that an abnormal behavior was noted, providing only the room number in question, transmitting the data through the ODIN platform. From there, sleep technicians will be notified with a pop-up message containing only the room number so that they can check in person the status.

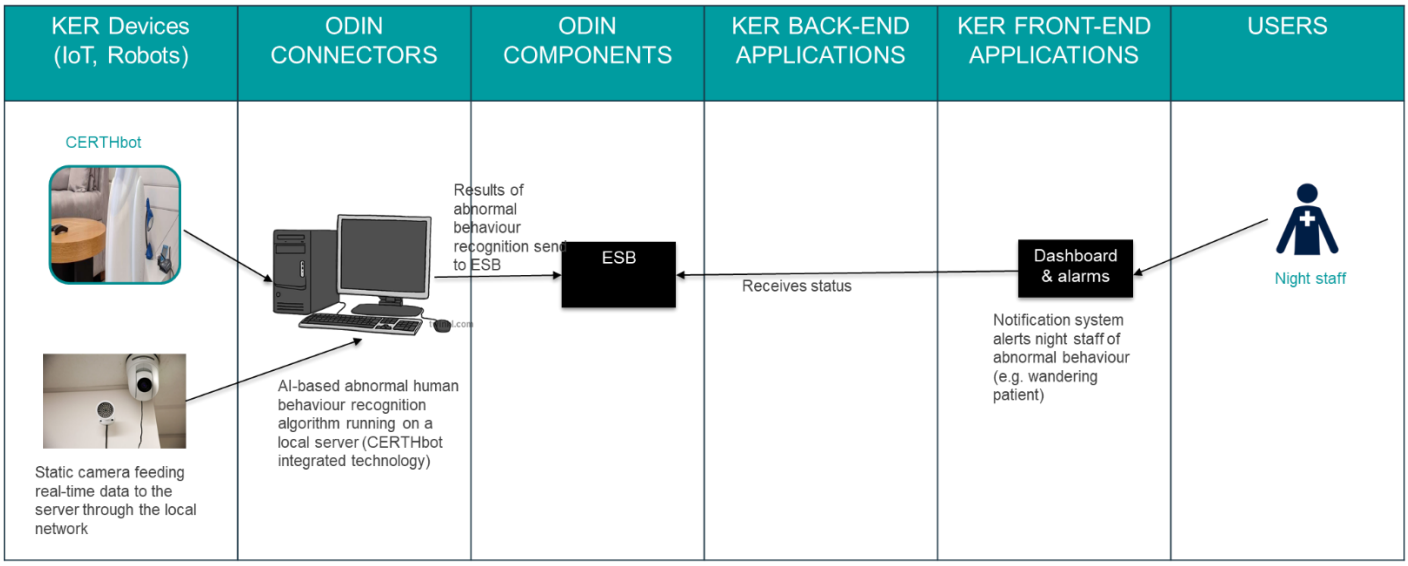

Figure 7: RUC C: UC 7 Patient monitoring/ Evacuation Data Flow

B. MUL

MUL will be sharing data with the Consortium on two instances, regarding blood sample transport and asset/medical device finding. As such, it does not envision any personal data sharing at this moment. All data sharing activities will be protected by robust encryption mechanisms in order to ensure that the data remains secure.

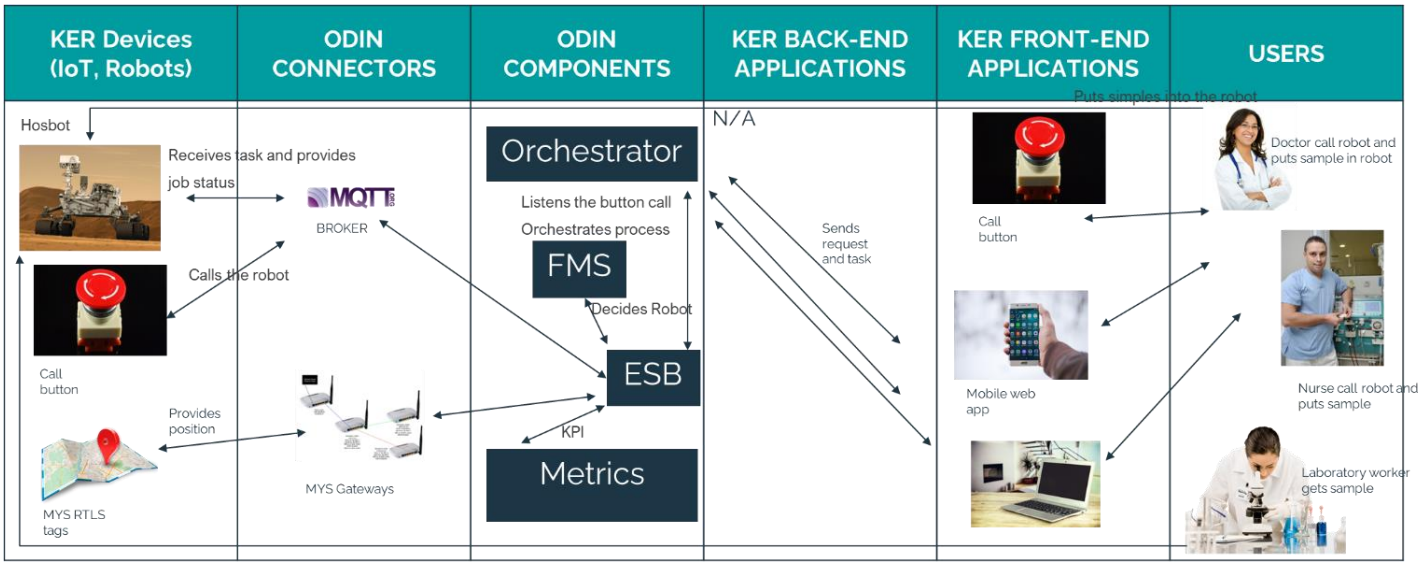

Figure 8: RUC A2-UC4 – Blood sample transport Data Flow

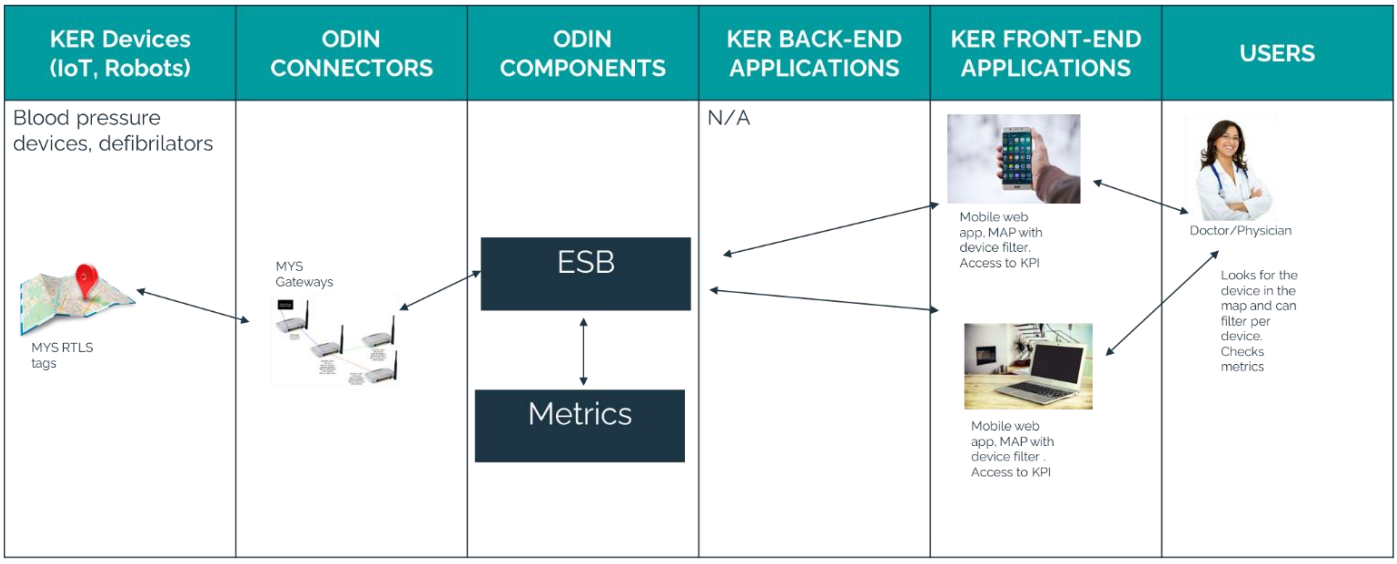

Figure 9: RUC B2-UC2 - Asset/medical device finding Data Flow

C. SERMAS

SERMAS will be sharing data with partners on three occasions, related to logistics support, clinical engineering, MD locations and real-time management, as well as to disaster preparedness, as demonstrated below. During the use case regarding the logistics support, SERMAS will be sharing data with FORTH that are connected to information regarding stents' purchase and consumption, patient hospitalization, emergency and outpatient episode information, as well as variables related to the stents, patients and patient episodes. Personal data, notably sensitive data referring to the patients' health may be shared in order to develop an algorithm for predicting the need to purchase hospital consumables. As such, in order to ensure the data's security and patients' privacy and confidentiality, SERMAS shall only share anonymized data, employing a K-anonymity analysis. The anonymization will be performed within SERMAS by adequately trained personnel.

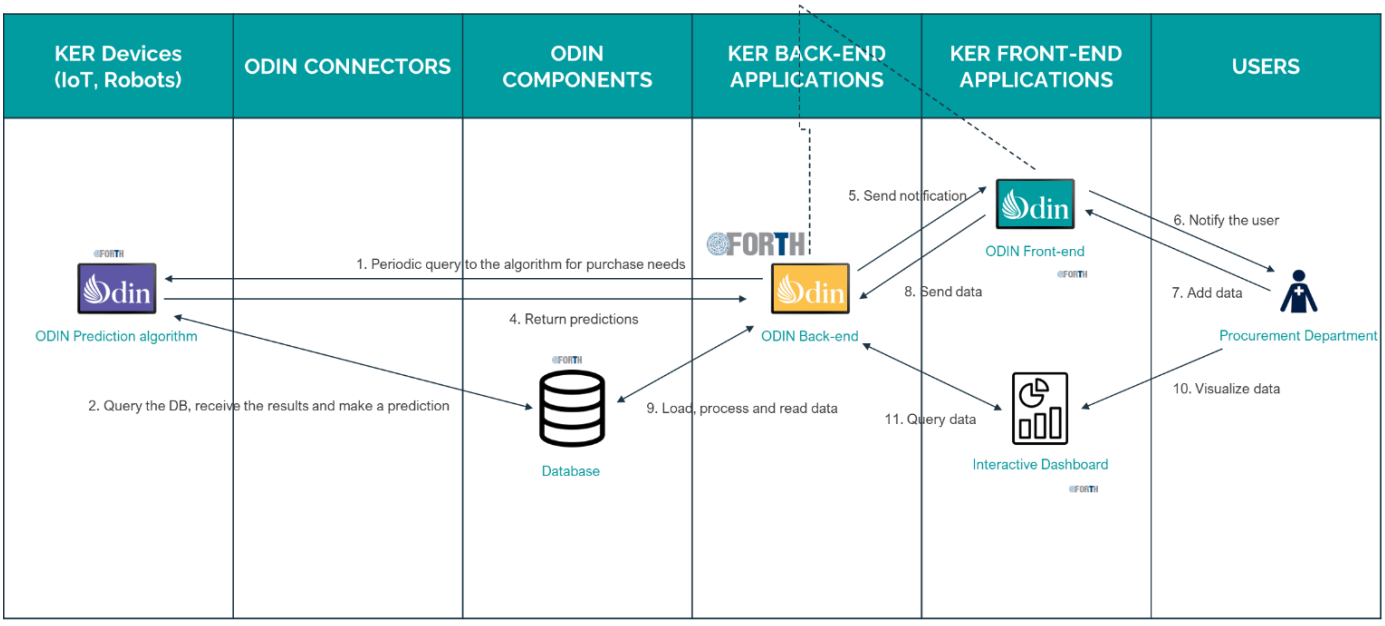

Figure 10: RUC B1 UC1: Aided logistics support Data Flow

For the use case demonstrated below, SERMAS will be sharing data, in particular the plans of the hospital area used by the HOSBOT developed by Robotnik, with MYS in order for the latter to design and install the RTLS sensors that will monitor the robot's movement. Similarly, the plans will also be shared with SSSA that will use them to analyze the requirements of the robot's potential trajectories. SSSA will also receive pictures of the stents and further medical equipment to analyze the dimension requirements that the robot must have. No personal data shall be shared.

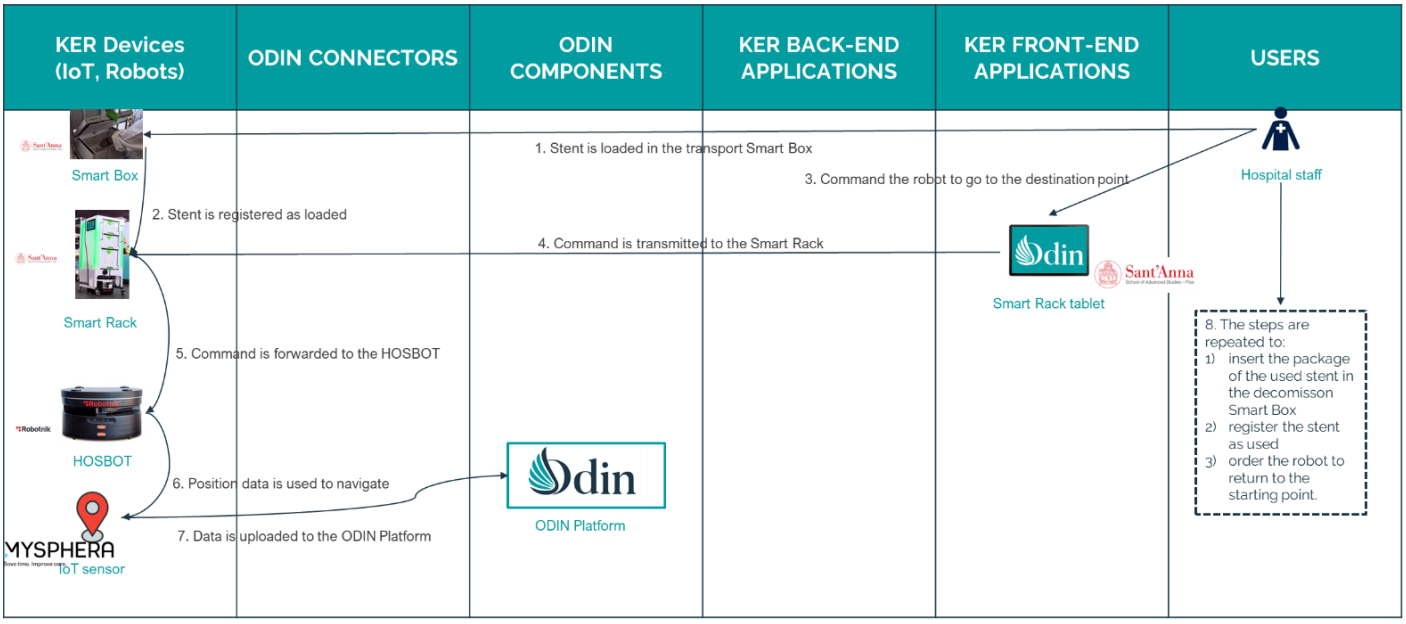

Figure 11: RUC B2 UC2: Clinical engineering, MD locations, real-time management Data Flow

With regards to the disaster preparedness use case, SERMAS does not intend to share any personal data. The video surveillance system is part of the hospital's already existing equipment and will not be transmitting any personal data. Data will be shared for the purpose of enhancing the hospital's disaster preparedness plan, while INETUM's algorithm will be utilized in the premises of SERMAS to transform the images into coordinate vectors. As such, neither INETUM nor any other partner will have access to the data in question.

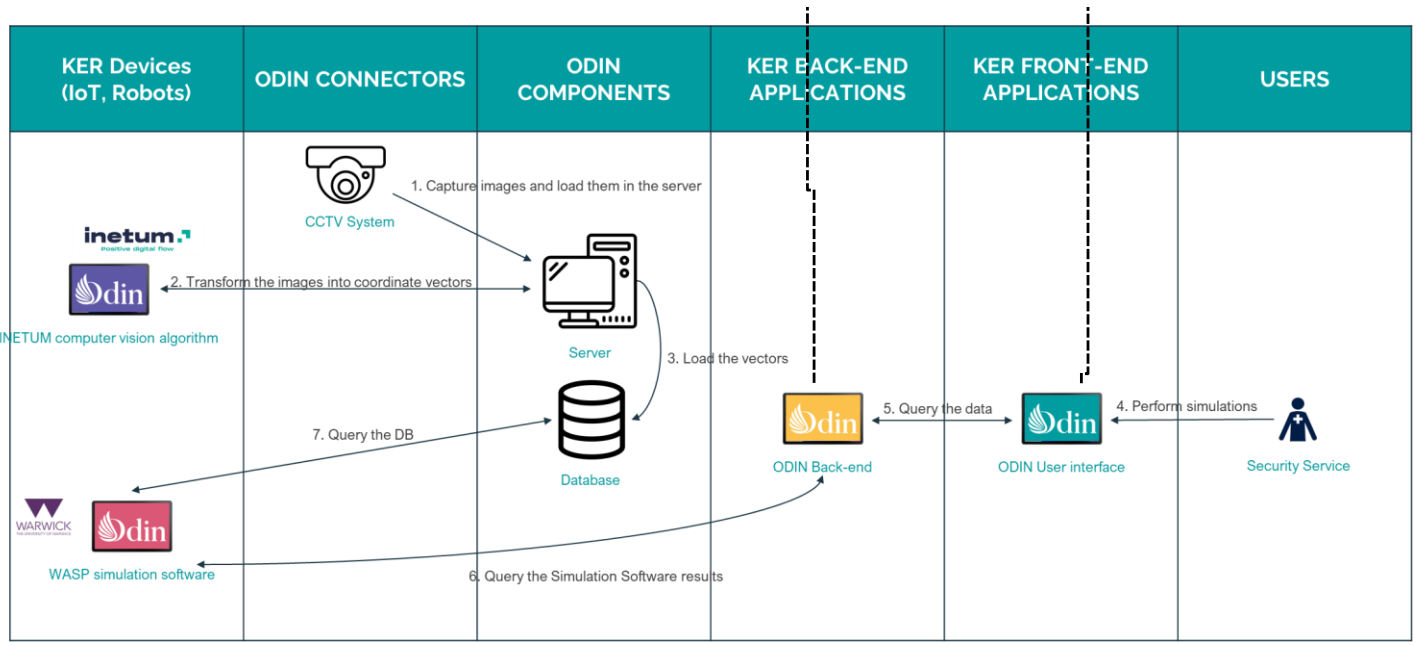

Figure 12: RUC C UC1: Disaster Preparedness Data Flow

D. UCBM

UCBM will be sharing a number of data with FORTH in order for the latter to develop the corresponding algorithms. In particular, in the use case below, UCBM will be sharing food pictures with FORTH in order to develop and train an algorithm meant to estimate food consumption, as well as calories and macronutrients assumption. No personal data will be transmitted.

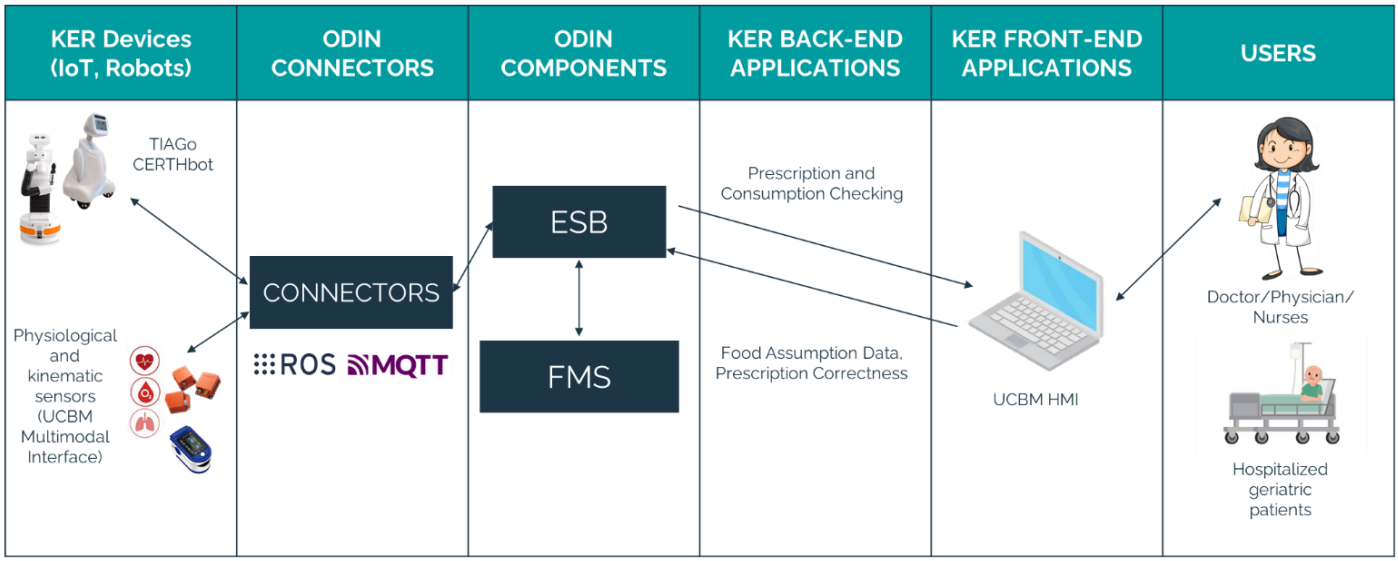

Figure 13: RUC A2.1-UC4 - Monitoring of food consumption to prevent undernutrition Data Flow

Similarly, for the use case related to rehabilitation to prevent loss of mobility, UCBM will be sharing motion data and video during rehabilitation exercises with FORTH in order to develop and train the algorithm that will be used for the estimation of motion performance during rehabilitation exercises. In order to generate the data, joint angle data are produced by wearable sensors placed on UCBM healthy researchers, who have provided their prior, written, freely given and informed consent. Video was simultaneously recorded. The researchers personal data were not transmitted.

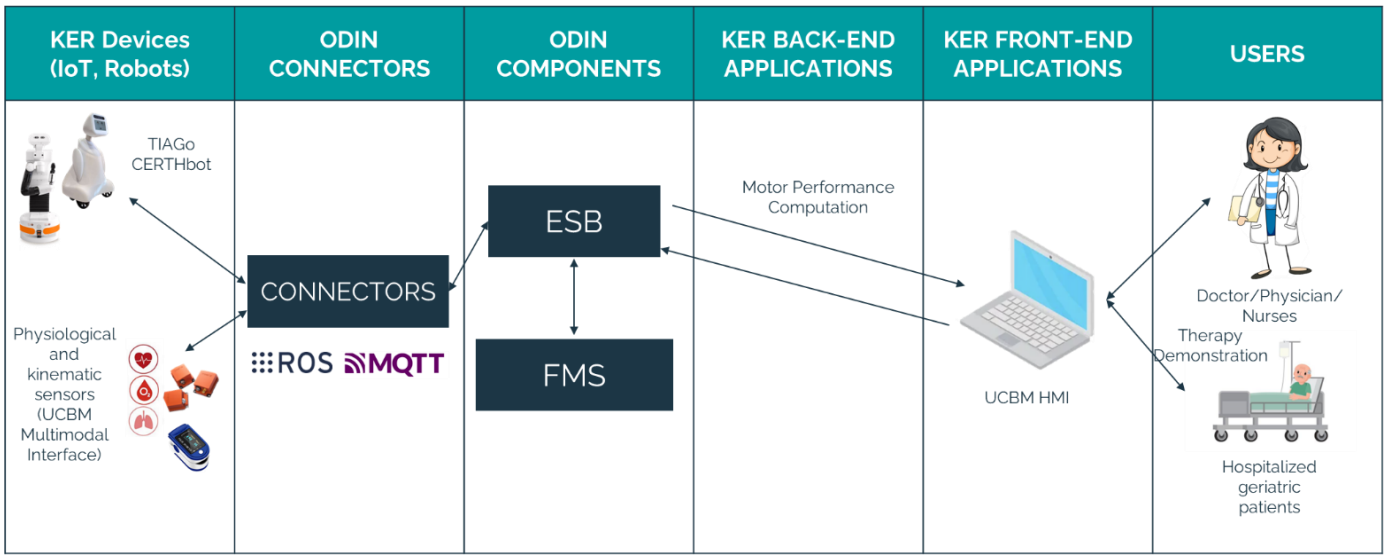

Figure 14: RUC A2.2-UC4 - Rehabilitation to prevent loss of mobility Data Flow

Finally, oxygen mask pictures are shared with FORTH in order to develop and train the algorithm that will be used for the estimation of motion performance during rehabilitation exercises. The pictures refer to healthy subjects, in particular UCBM researchers, wearing sample oxygen masks. The researchers had provided their prior, written, freely given and informed consent. Their personal data were not transmitted.

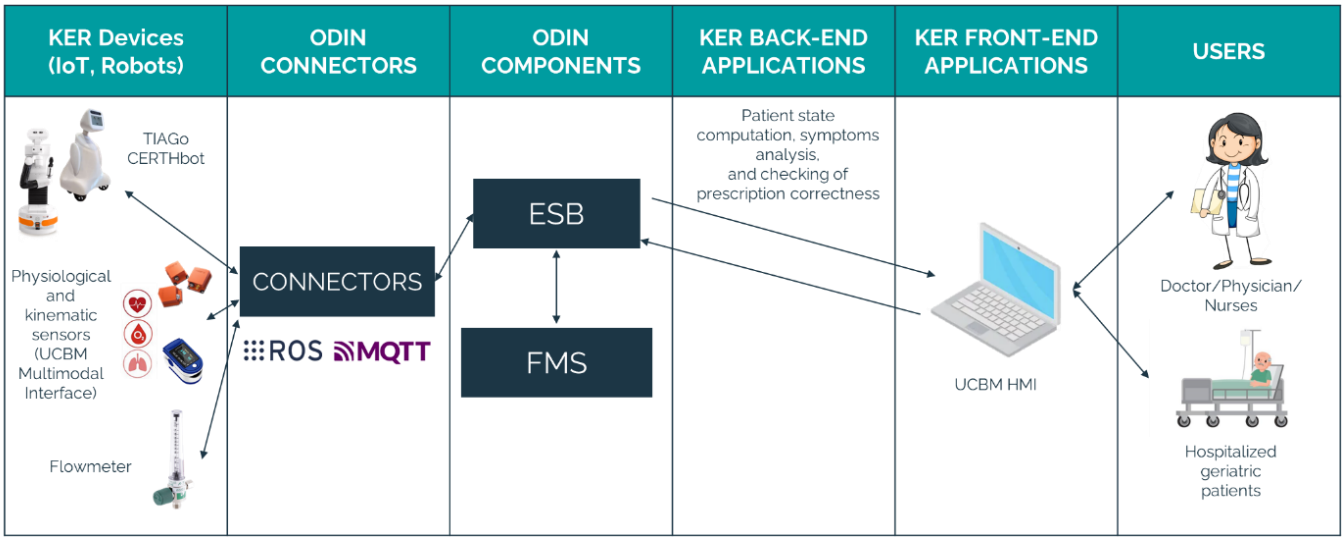

Figure 15: RUC A3 - Monitoring of oxygen therapy to prevent hypoxia Data Flow

E. UMCU

UMCU intends to process data that involve patients' health data and is related to EHR data. Data will be collected using questionnaires, the Utrecht Patient Oriented Database (UPOD) and the Luscii app. This data generation, collection and processing is needed for the execution of the use cases in their pilots with the purpose of performing a variety of health services related tasks, including optimization and personalization of patients' diagnostic trajectory, monitoring vascular surgery patients at home post-surgery and locating pathogen carriers inside the UMCU. In order to achieve this while ensuring protection of personal data, the data will be pseudonymized and any analysis will be performed within the hospital, in secured servers subject to authorization and access restrictions. In view of this, the following diagram is applicable for all relevant use cases.

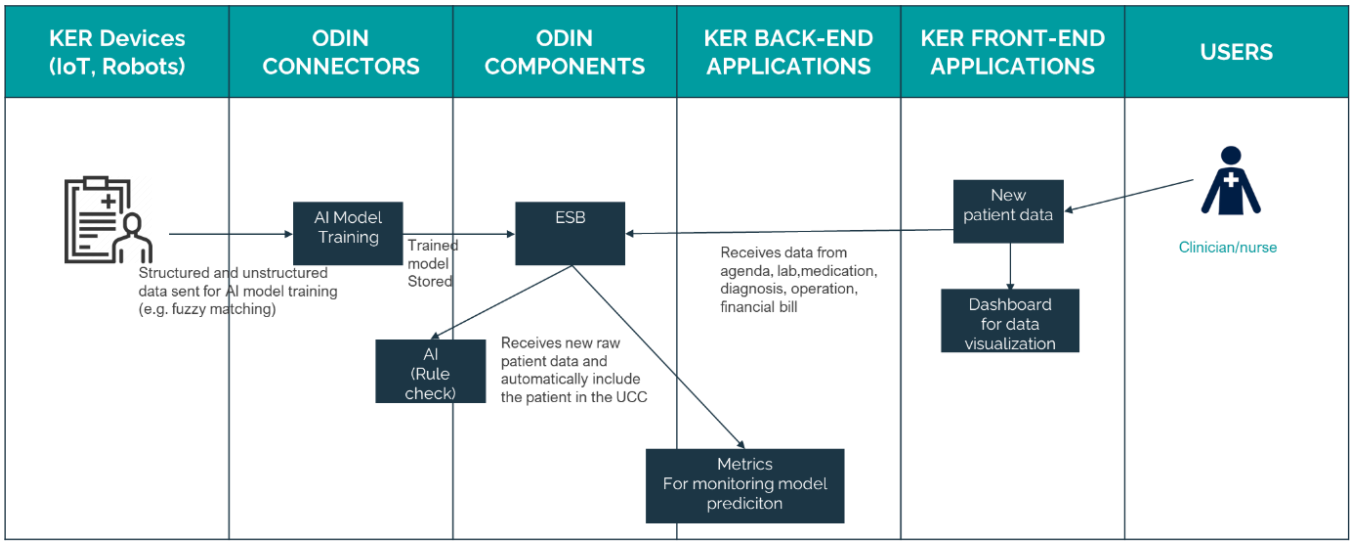

Figure 16: Indicatively, RUC A: Automated patient inclusion system in the UCC Data Flow

2.1.4 Findings

Data are a central component of the research project. At the project level WPs will mainly manage data related to the production of the different deliverables. Although there are several work packages and tasks, where the concrete data format and type is still to be identified, a positive commitment to ensuring anonymization, pseudonymization, purpose limitation and data minimization is visible from the provided responses to the questionnaires.

2.2 Processing of Existing Data

In order to determine, if and what previously available data will be processed by the consortium members, the following table has been provided to be filled out by all partners.

Table 3: Processing of Existing Data

	Please provide your answers in this column:
Dataset(s) name	*What is the name of the used dataset(s)?*
Dataset(s) description	*Please provide a short description of the dataset(s).*
Personal Data	*Does the dataset include personal data? If yes, please specify the type of personal data.*
Purpose	*What is the purpose for which you use/ process the dataset(s)?*
Data format	*What format(s) are your dataset(s)?*
Data Storage	*Where will you store the dataset(s)?*
Main Data Source	*What is the main source of the dataset(s)?*
Data Ownership	*Who owns the dataset(s)?*
Country of Origin	*Where does the dataset come from?*
Restrictions on the use	*Are there any restrictions for the use of the datasets?*
Access	*Who has access to the datasets? Please include other work packages which will also access the datasets.*
Retention Period	*How long will you keep the datasets?*
Licence	*Under which licence did you obtain access to the datasets?*
WP and task	*For which work package and which task do you need to use the datasets?*
Additional Comments	*Please add here any additional comments.*

In principle, as outlined in the previous sections of the current document, the processing of an already existing dataset for scientific research, is permitted under the GDPR. The processing activities and the already existing datasets are nevertheless subject to conditions such as appropriate safeguards, technical and organizational measures, pseudonymization, etc. A general good practice is, whenever possible, to share only anonymized or pseudonymized data when reusing already existing data sets, and, in case of pseudonymized data, the researcher should does not receive the link file, which will possibly allow re-identification.

For the reuse of special categories of personal data (Art. 9(2)(j) GDPR) for the purposes of academic research, necessary and proportionate safeguarding measures should be undertaken in order to be compliant.

As per Art. 14 (5)(b) GDPR, the data subjects have the right to be informed immediately about the processing if personal data have not been received from the data subject unless this requires a disproportionate effort. In the latter case, the data subject must be informed either publicly or by means of a privacy statement.

The current analysis is carried-out on a per-partner basis. Nine partners from ODIN's consortium, CUB, MUL, SERMAS, UCBM, UMCU, CERTH, FORTH, INETUM and Philips, will process already existing datasets.

Table 4: Processing of Existing Data: CUB

Berlin ESADA data	Sleep studies recorded in Berlin from 2008 – 2021 as part of the ESADA project
Data identification	
Data set description	This dataset contains roughly 200-230 sleep apnea patients that underwent PG and/or PSG recordings between 2007 and 2021. Generally, multiple recordings (nights) using the same device. This dataset is the contribution of the Charité towards the ESADA project. Ages Eligible for Study: 18+ (Adult) Sexes Eligible for Study: All Accepts Healthy Volunteers: No Criteria Inclusion Criteria: •Adult subjects (18+) with suspicion of sleep apnea (obstructive/central). Exclusion Criteria: •History of any sleep disorder, or any Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) axis I disorder other than sleep apnea.
Source (i.e. which device?)	PG devices: Embletta, Nox, Miniscreen, Somnotouch PSG devices: Embla, Alice V, Alice Light Edition, Somnoscreen PG 6 channel electrode PSG with a minimum of 12 electrodes
Partners responsibilities	
Partner in charge of data collection	Charité
Partner in charge of data analysis	Philips

Partner in charge of data storage	Philips
Standards and metadata	
Info about metadata (Production and storage dates, places) and documentation?	Still being defined: All documentation, dates, descriptive statistics, inferential statistics, device information, and so on are available. They just need to be defined before sharing the data.
Standards, Format, Estimated volume of data	The sleep recordings will be provided either in the format of EDF (EU data format) or as their default format. Further descriptive data is available via PDF, personal details are anonymized.
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	An AI model will be trained on the retrospective data. From here, the model will have enough information to successfully classify sleep disorders from sleep recordings of the same type. For instance, if a new sleep study measured by PSG were entered into the model after training, the model would be expected to predict what type of sleep disorder is present.
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	Available upon request: only for members of the consortium.
Data sharing, re-use and distribution (How?)	If all the data is fully anonymized, the data sharing agreement is extremely straight forward. It is not possible to trace back to the original patients since all information which could identify them would be removed.
Embargo periods (if any)	
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	The Charité owns the data and it is only available to use in this consortium. There are no limitations on what can be conducted with this dataset. As the data will be fully anonymized it will be available should adhere to all partner policies.
Insomnia Dataset	**Sleep studies recorded in Berlin from a randomized control trial on insomnia patients, only baseline nights are used**
Data identification	
Data set description	This dataset contains 64 insomnia patients that underwent PSG recording between 2008 and 2010. The patient data collected from 2008 – 2009 were used in the clinical trial “A polysomnography study to evaluate the effect, safety and tolerability of oral administration of almorexant (ACT 078573) in

	adult subjects with primary insomnia". The remaining patients were recorded under the same criteria. Ages Eligible for Study: 18 Years to 64 Years (Adult) Sexes Eligible for Study: All Accepts Healthy Volunteers: No Criteria Inclusion Criteria: •Adult subjects (18-64 years) with a diagnosis of primary insomnia. Exclusion Criteria: •History of any sleep disorder, or any Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) axis I disorder other than primary insomnia. •Sleep apnea, or restless legs syndrome. •Daytime napping of more than 1 hour per day. •Important caffeine consumption, heavy tobacco use, alcohol or drug abuse within 2 years prior to the screening visit. •Unwillingness to refrain from drugs, over-the-counter or herbal medication having an effect on sleep or behavior.
Source (i.e. which device?)	PSG with a minimum of 12 electrodes
Partners responsibilities	
Partner in charge of data collection	Charité
Partner in charge of data analysis	Philips
Partner in charge of data storage	Philips
Standards and metadata	
Info about metadata (Production and storage dates, places) and documentation?	Still being defined: All documentation, dates, descriptive statistics, inferential statistics, device information, and so on are available. They just need to be defined before sharing the data.

Standards, Format, Estimated volume of data	The sleep recordings will be provided either in the format of EDF (EU data format) or as their default format from the original recording device. Further descriptive data is available via PDF, personal details are anonymized
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	An AI model will be trained on the retrospective data. From here, the model will have enough information to successfully classify sleep disorders from sleep recordings of the same type. For instance, if a new sleep study measured by PSG were entered into the model after training, the model would be expected to predict what type of sleep disorder is present.
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	Available upon request: only for members of the consortium.
Data sharing, re-use and distribution (How?)	If all the data is fully anonymized, the data sharing agreement is extremely straight forward. It is not possible to trace back to the original patients since all information which could identify them would be removed.
Embargo periods (if any)	
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	The Charité owns the data and it is only available to use in this consortium. There are no limitations on what can be conducted with this dataset. As the data will be fully anonymized it will be available should adhere to all partner policies.
Finger ring Dataset	**Sleep studies recorded in Berlin from two validation studies regarding pulse oximetry wearable finger rings**
Data identification	
Data set description	This dataset contains approximately 60 patients in each study that are suspected of having obstructive sleep apnea. patients underwent PSG recording whilst wearing pulse oximetry finger rings in 2022.
Source (i.e. which device?)	PSG with a minimum of 12 electrodes Pulse oximetry finger rings: Oura, SleepOn, Circul, and SleepImage
Partners responsibilities	
Partner in charge of data collection	Charité
Partner in charge of data analysis	Philips, CERTH

Partner in charge of data storage	Philips, CERTH
Standards and metadata	
Info about metadata (Production and storage dates, places) and documentation?	Still being defined: All documentation, dates, descriptive statistics, inferential statistics, device information, and so on are available. They just need to be defined before sharing the data.
Standards, Format, Estimated volume of data	The sleep recordings will be provided in their default format from the original recording device. Further descriptive data is available via PDF, personal details are anonymized
Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	An AI model will be trained on the data. From here, the model will have enough information to successfully classify sleep breathing disorders from sleep recordings of the same type. For instance, if a new sleep study measured by PSG were entered into the model after training, the model would be expected to predict what type of sleep disorder is present. The PSG recordings here can be anonymized and used as test data for the ESADA AI model to see if the model can accurately classify the disorder and the severity.
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	Available upon request: only for members of the consortium.
Data sharing, re-use and distribution (How?)	If all the data is fully anonymized, the data sharing agreement is extremely straight forward. It is not possible to trace back to the original patients since all information which could identify them would be removed.
Embargo periods (if any)	
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	The Charité owns the data and it is only available to use in this consortium. There are no limitations on what can be conducted with this dataset. As the data will be fully anonymized it will be available should adhere to all partner policies.

Table 5: Processing of Existing Data: SERMAS

	Please provide your answers in this column:
Dataset(s) name	*PACIENTE, CMBD_H, CMBD_U, CMBD_A, stent catalogue, stent purchase-consumption logs, stent use pdfs*
Dataset(s) description	*PACIENTE has the socio-demographic data of the hospital patients;* *CMBD_H, CMBD_U and CMBD_A have the hospitalization, emergency and outpatient information;* *The stent catalogue is the catalogue of the stents purchased by the hospital and it contains a technical description;* *The stent purchase-consumption logs have purchase and consumption information such as dates, price or vendors;* *The stent use pdfs are pdf files that contain the stent id, the name and id of the patient who received the stent and the date of the procedure.*
Personal Data	*PACIENTE, CMBD_H, CMBD_U, CMBD_A and the stent use pdfs contain socio-demographic and clinical variables.*
Purpose	*We will use existing datasets to produce the datasets for RUC B1 UC1.*
Data format	*Excel/CSV.*
Data Storage	*In our local computers.*
Main Data Source	*The hospital information system.*
Data Ownership	*Hospital Clínico San Carlos.*
Country of Origin	*Spain.*
Restrictions on the use	*The data can only be used for the objectives of RUC B1 UC1.*
Access	*No partner has access to them. Currently, only FORTH has access to the datasets produced after processing these ones.*
Retention Period	*The data will be stored as long as needed to comply with both ODIN and SERMAS policies.*
Licence	*Does not apply.*
WP and task	*For all WP7 tasks.*
Additional Comments	*None.*

Table 6: Processing of Existing Data: UCBM

ODIN UCBM	Data collected during UCBM UCs
Data identification	
Data set description	The data collected during the experiments on the UCBM use cases will include the following data categories: i. Clinical data: patients' status collected by using clinical scales ii. Physiological patients' data: heart rate, respiration rate, EMG (if available), M-IMU data (if available) iii. Environmental data: temperature, humidity, brightness, noise iv. Robot data: joint angles, end effector pose, RGB-D camera data, navigation data (position, velocity, etc.)
Source (i.e. which device?)	Data coming from - Clinical scales - Wearable sensors (UCBM multimodal interface) - Environmental sensors (camera, Transparent Robot, etc.) - TIAGo, CERTHbot A combination of these sources will depend on the specific UC/database under analysis.
Partners responsibilities	
Partner in charge of data collection	UCBM, CERTH, FORTH, SSSA, THL, UPM.
Partner in charge of data analysis	All
Partner in charge of data storage	UCBM
Standards and metadata	
Info about metadata (Production and storage dates, places) and documentation?	All acquired data will always be accompanied by metadata, containing at least: author, date created, date modified and file size. In the final version, acquired data on patients should contain the following classes of metadata: v. Technical Metadata: for decoding and rendering files vi. Preservation Metadata: for the long-term management and clinical archiving vii. Rights Metadata: for intellectual property and license (if needed).
Standards, Format, Estimated volume of data	Data will be collected in standard .CSV or .TXT or .JSON formats. The estimated volume of data is about 10 GB, including integration and clinical validations of the UCBM UCs on patients.

Data exploitation and sharing	
Data exploitation (purpose/use of the data analysis)	Scientific publication, conferences, invited talks.
Data access policy / Dissemination level (Confidential, only for members of the Consortium and the Commission Services) / Public	Confidential, only for members of the Consortium and the Commission Services. Public data access policies will only be evaluated after an embargo period has ended.
Data sharing, re-use and distribution (How?)	The sharing, re-use and sharing of data, in accordance with the ODIN DMP and compatibly with the UCBM policy, will be allowed to foster the reusability and to accelerate research through the typical channels of scientific dissemination (e.g. journals foreseeing publications of datasets). Moreover, the principle of 'as open as possible, as closed as necessary' will be applied, according to the EU H2020 Program Guidelines on FAIR Data.
Embargo periods (if any)	The collected data will undergo a period of embargo necessary for the elaboration and publication of the clinical results obtained in the ODIN project framework.
Archiving and preservation (including storage and backup)	
Data storage (including backup): where? For how long?	The acquired data must be kept in physical archives at UCBM, according to what has been recommended by the Ethics Committee and by indications of the Italian Ministry of Health.

UCBM will re-use their own previously generated datasets, which will be fully pseudonymized (see table below) in order to provide sufficient measures for safeguarding the rights and freedoms of data subjects. Further technical and organizational measures, such as access restriction, and data minimization will be accordingly implemented.

Table 7: Processing of Existing Data: UMCU

	Please provide your answers in this column:
Dataset(s) name	Utrecht Patient Oriented Database
Dataset(s) description	Routine care database consisting of all patients that ever visited the UMC Utrecht
Personal Data	Yes, although fully pseudonymized, these data are still considered personal data
Purpose	Research
Data format	Sasbdat files
Data Storage	According to the UMC Utrecht data management policy: within UPOD structures protected by authorization and outside UPOD

	structures in protected research folder structures protected by authorization
Main Data Source	The electronic health record system
Data Ownership	UMC Utrecht
Country of Origin	The Netherlands
Restrictions on the use	Yes, it cannot leave the hospital and users need to comply with Dutch data management and research guidelines before use
Access	Only UMC Utrecht
Retention Period	The full database is maintained indefinitely, the research subsets are kept for 15 years according to Dutch research guidelines
Licence	N/A
WP and task	For WP7 UC1/2/4
Additional Comments	It is UMCU own dataset so there is easy access.

Table 8: Processing of Existing Data: CERTH

	Please provide your answers in this column:
Dataset(s) name	Kinetics (400/600/700)
Dataset(s) description	A collection of large-scale, high-quality datasets of URL links of up to 650,000 video clips that cover 400/600/700 human action classes, depending on the dataset version. The videos include human-object interactions such as playing instruments, as well as human-human interactions such as shaking hands and hugging. Each action class has at least 400/600/700 video clips. Each clip is human annotated with a single action class and lasts around 10 seconds.
Personal Data	Yes. The dataset is depicting humans performing actions in various situations during their daily life.
Purpose	Train baseline deep neural network models to perform human action recognition in hospital environments
Data format	Video format, either .mp4 or .avi.
Data Storage	At specialized machines on our premises.
Main Data Source	YouTube
Data Ownership	DeepMind
Country of Origin	Not defined
Restrictions on the use	Restrictions as defined in dataset licence
Access	CERTH (is publicly available so practically anyone)
Retention Period	For the duration of the project
Licence	The kinetics dataset is licensed by Google Inc. under a Creative Commons Attribution 4.0 International License.
WP and task	Task T5.2 of WP5

In the case of CERTH (see table above), the existing datasets containing personal data will be collected from the source YouTube, which grants public access. As the datasets are licensed under a Creative Common Attribution, no specific permission is required. The datasets will be processed for a specifically defined purpose to "*train baseline deep neural network models to perform human action recognition in hospital environments*", in line with CERTH's obligations under WP5, Task 5.2. The datasets are stored at CERTH's specialized machines on their own premises, where technical and organizational measures for safeguarding of data subjects' rights and freedoms are implemented.

Table 9: Processing of Existing Data: FORTH

	Please provide your answers in this column:
Dataset(s) name	*MedGRFood Database*
Dataset(s) description	*Food images and nutritional analysis*
Personal Data	*No*
Purpose	*For AI model training*
Data format	*.jpg*
Data Storage	*In FORTH personnel private computer*
Main Data Source	*MedGr Project*
Data Ownership	*Dimitrios Fotiadis*
Country of Origin	*MedGr Project*
Restrictions on the use	*In case of publications we should refer to the MedGR Project*
Access	*Only authorized FORTH personnel*
Retention Period	*For the project*
Licence	*We got a licence from the project's coordinator (Dimitrios Fotiadis)*
WP and task	*WP6, T6.3*
Additional Comments	

Table 10: Processing of Existing Data: INETUM

	Please provide your answers in this column:
Dataset(s) name	*MS COCO (Common Objects in Context)*
Dataset(s) description	*It's a multipurpose dataset with 125 classes*
Personal Data	*No*
Purpose	*Training of AI algorithm*
Data format	*Images*
Data Storage	*No*

Main Data Source	https://cocodataset.org/#home
Data Ownership	https://cocodataset.org/#home
Country of Origin	*Unspecified*
Restrictions on the use	*No*
Access	*Open source.*
Retention Period	*Unspecified*
Licence	Creative Commons Attribution 4.0 License. Creative Commons — Attribution 4.0 International — CC BY 4.0
WP and task	*WP6-T6.2*
Additional Comments	-

Table 11: Processing of Existing Data: Philips

	Please provide your answers in this column:
Dataset(s) name	ESADA
Dataset(s) description	200 up to 230 sleep apnea patients that underwent PG and/or PSG recordings between 2007 and 2021. Generally, multiple recordings, usually performed during nights, use the same devices for acquiring PSG data (e.g., ECG, EMG or EOG devices). This dataset is the contribution of the Charité towards the ESADA project
Personal Data	DATA IS ANONYMIZED
Purpose	*Developing AI models for automatically detect sleep disorders and deploying a federated learning approach to leverage multiple hospitals without accessing the data*
Data format	*What format(s) are your dataset(s)?*
Data Storage	*CUB*
Main Data Source	*Polysomnography*
Data Ownership	*Question for CUB*
Country of Origin	*Berlin*
Restrictions on the use	*Question for CUB*
Access	*CUB*
Retention Period	*ODIN Project duration*
Licence	*Question for CUB*
WP and task	*WP6*
Additional Comments	*Please add here any additional comments.*

2.3 Data Storage Management & Retention Policy

As indicated in Section 3 regarding General Security Instructions, datasets that contain personal or confidential information need to be securely stored, and the data controller needs to ensure that the latest security updates are in place. Personal data obtained by the project's partners will be securely stored on local data servers. Additionally, best practices of data storage and handling (see Section 2, 3, and 4) are communicated to all project partners and team members.

At the beginning of the project (March 2021) a dedicated repository for project collaborative work was set up.

On ODIN's public website (https://www.odin-smarthospitals.eu) the following data and information will be publicly available:

- General information about the project, its mission, and objectives;
- The participating consortium;
- Project public deliverables;
- Publications.

Project datasets (presentations, reports, scientific publications etc.), which are not intended for public dissemination will be shared only within the consortium on its private repository (CBMLBox). The following deliverables are confidential:

- D1.1: Quality Management Plan;
- D1.5, D1.6, D1.7, D1.8: Annual Report v1, 2, 3, 4;
- D2.2: Hospital requirements report;
- D2.3: ODIN platform catalogue;
- D2.4: Acceptance, Trust and Change Management;
- D3.4, D3.5, D3.6: Privacy, Security and Trust report v1, 2, 3;
- D3.7, D3.8, D3.9: Technical Support Plan and Operations v1, 2, 3;
- All WP5 deliverables;
- D6.2, D6.3: Data results interpretation and data integration services v1, 2;
- D6.6, D6.7: Development of the High-Level AI-based models of planning, scheduling, and workflow modelling v1, 2;
- D7.1: Pilot Studies Use Case Definition and Key Performance;
- D7.8: New use case demonstrations conclusion (I to IX);
- D7.9: Pilot Studies Evaluation Results and sustainability;
- D8.3: Certification scheme strategy and sustainability plan;
- D8.4: Data ethics in public procurement for hospitals v1;
- All WP 9 deliverables;
- D10.3: Supply Open Innovation;
- D10.4-6: Open Calls Report v1, 2, 3;
- All WP11 deliverables.

Of the datasets intended to be openly accessible, particularly the datasets containing personal information (e.g., interviews, pilots, focus groups) require anonymisation prior to release.

Project data will also be stored in partners own facilities and servers. The storage time depends on the particular data but in general the rules are:

- During the lifetime of the project: the availability of and access to the data on the different servers will be maintained as long as they are needed.
- After the end of the project: the project public website and CBMLBox will be maintained until the end of the project. Afterwards, whenever possible, the data from the ODIN platform will be anonymized and made available in accordance with the FAIR data principles. Dedicated discussions around how this process will take place are ongoing and consider, particularly, the sensitive nature of the data.

The figure below offers a simplified scheme of the data processing, storage, and retention.

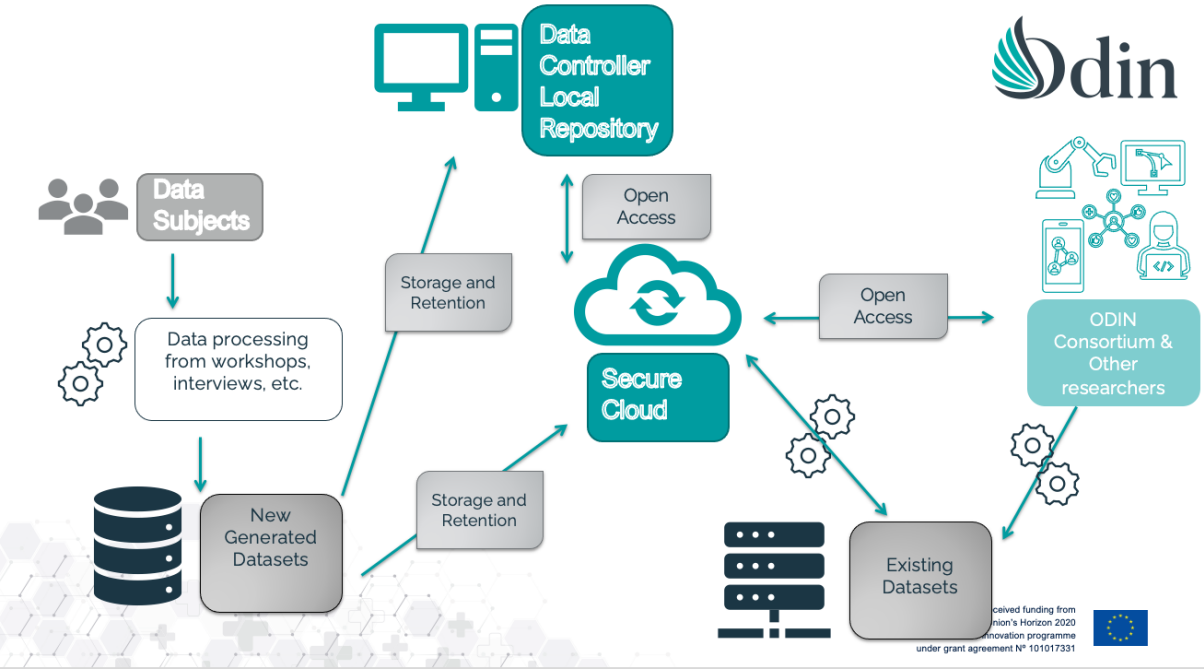

Figure 17: Storage and Flow of Data

All project data, except for the public website, is stored in password protected repositories and servers. The security strategy depends on the security policy of the partners in charge of these repositories and servers.

Beyond the provisions outlined, the table below presents the data storage management and retention policy on a per-partner basis.

Table 12: Data Storage Management & Retention

Partner	Data Storage Management & Retention Policy
Robotnik	Most of the data is stored inside the robot for internal algorithm optimization processes. There is data that is kept in only one work cycle of the robot and other data that will be kept during the pilot period.
SERMAS	The data will be stored by SERMAS, in a server independent from the Hospital Clínico San Carlos network. The data will be kept for the complete duration of the ODIN project.
MYSPHERA	MYSPHERA will store the data in the Cloud available from ODIN project. The data will be stored "as long as it is needed for project purposes".
M&S	Employees of MINDS & SPARKS GmbH will store the documents containing personal data locally on controlled secured, password protected personal computers, where only authorized access is allowed and to databases containing personal data, privacy by design techniques and encrypted file transfers. The data will be deleted three years after the end of the project.
MEDTRONIC	• Data capture by ODIN's website is stored in online servers • Miro board data in Miro tool's servers • Documents will be stored either in Medtronic's internal servers/OneDrive folders or ODIN project's repository (accessible to consortium). The data will be kept during ODIN project's lifetime and beyond, at least 5 years after the end of the project.
CERTH	Data will be stored locally on at CERTH premises by responsible researchers assigned this task. In compliance with GDPR core principles of proportionality and minimization, data will be processed and kept during the project's lifetime.
FORTH	The data received from pilots will be stored in their local repositories in order to ensure data privacy and protections. The current plan suggests that no data will be processed outside of the pilot site. Any data produced by FORTH will be stored in local repositories accessible only by FORTH personnel. Data will be kept for the project period.
University of Warwick (UoW)	UoW Team is collecting and they will be stored on the ODIN Project shared storage. No data will be stored locally.
SSSA	Data will be centrally stored (if needed) into the hospital's ICT infrastructure (mainly concerning WP3 and WP4). The aspect data retention does not concern WP5 and SSSA activities because data will not be stored by SSSA, but will be transmitted to the ODIN's ICT platform/infrastructure. Temporary dataset, used for local analysis, will be frequently replaced with the new generated data flow.

THL	The data will be stored on a server or a laptop on-premise (TBDL) until the end of the project.
Philips Electronics Netherland BV (PEN)	The data would be stored at the pilot sites. Only a limited amount of data (if allowed and needed) will be shared to develop base models. PEN will access data locally at pilot sites and nothing will be transferred to PEN.
UPM	T2.4, T7.1, T10.3 will store the data using the cloud repository of the project (NextCloud). T3.3, T4.6 and WP5, data will be stored at each pilot site of in a pilot cloud provider. Data will be kept for the time of the project.
INETUM	At this stage, there is no data collection, processing or sharing envisioned on behalf of INETUM.
UCBM	The data will be available to the UCBM team for the development and clinical validation of the ODIN platform. They will be strictly kept in closed archives and not connected to the network at UCBM. The data will be collected, analysed and managed by UCBM for the entire duration of the ODIN project and only for its purposes. Once project tasks have been completed, the above mentioned repository will be managed as a long term data locker for project files and deliverables (duration to be defined).
UMCU	UMC Utrecht will story the study data on its secure servers protected by authorization. Data will be kept, according to Dutch law, for the period of 15 years.
Charite (CUB)	The partner CHARITE will store the data on servers inside the hospital and inside the firewall protection of the hospital. The data collected will be kept for 10 years unless other regulations in Germany require a longer storage. Clinical study data need to be stored for 15 years according to good clinical practice regulations.
MUL	MUL stores the data in a dedicated Research Folder Structure for which authorization is secured using personal logins. We only store pseudonymised data there. The key is stored at the separate drive. They will not transfer sensitive data outside the hospital. The MUL IT centre will be responsible for security of data in the data centre. The data will be available for at least 10 years.
UDGA	Datasets will be stored in the ODIN's repository (CMBLBox) and kept for the duration of the project.
MEDICT MEDEA	MEDEA - Data will be only collected by reference MEDEA personnel and will be stored in the internal server of the company for the lifetime of the project.

2.4 Further Processing of Previously Collected Data

Recital 50 GDPR provides that "*the processing of personal data for purposes other than those for which the personal data were initially collected should be allowed only where the processing is compatible with the purposes for which the personal data were initially collected*". The purposes of processing must be specified prior to, and in any event, not later than, the time when the collection of personal data occurs.[2] As outlined in section 9.2.3 of the current Data Management Plan, in order to conduct a "compatibility assessment", the following aspects should be considered[3]:

- The **link** between the original purpose and the new/upcoming purpose;
- The **context** in which the data was collected (i.e., What is the relationship between your company/organisation and the individual?);
- The **type and nature** of the data (i.e., Is the data sensitive?);
- The possible **consequences** of the intended further processing (i.e., How will the further processing impact the individual?);
- The existence of appropriate **safeguards** (i.e., encryption or pseudonymisation).

The compatibility assessment is, however, not necessary if the data should be used for archiving purposes in the public interest, statistical or scientific research purposes.

The table below outlines if further data processing is being conducted within ODIN.

Table 13: Further Data Processing

Partner	Further Data Processing
Robotnik	No further data processing envisioned at the moment.
SERMAS	No further data processing envisioned at the moment.
MYSPHERA	No further data processing envisioned at the moment.
M&S	No further data processing envisioned at the moment.
MEDTRONIC	No further data processing envisioned at the moment.
CERTH	No further data processing envisioned at the moment.

[2] ARTICLE 29 DATA PROTECTION WORKING PARTY, Opinion 03/2013 on purpose limitation. Available here: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf.

[3] https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/principles-gdpr/purpose-data-processing/can-we-use-data-another-purpose_en.

FORTH	Further processing for scientific reasons.
University of Warwick (UoW)	No further data processing envisioned at the moment.
SSSA	No further data processing envisioned at the moment.
THL	No further data processing envisioned at the moment.
Philips Electronics Netherland BV (PEN)	No further data processing envisioned at the moment.
UPM	No further data processing envisioned at the moment.
INETUM	No further data processing envisioned at the moment.
UCBM	No further data processing envisioned at the moment.
UMCU	No further data processing envisioned at the moment.
Charite (CUB)	The patient data will be processed further for patient diagnostic purposes. The result of the diagnosis will be part of the electronic patient record as used in clinical work in the hospital.
AMIS	No information provided envisioned at the moment.
MUL	No further data processing envisioned at the moment.
UDGA	No further data processing envisioned at the moment.
MEDICT MEDEA	No further data processing envisioned at the moment.

3 FAIR DATA

3.1 FAIR Guidelines for Data Management

In its FAIR Data Management Horizon 2020 Guidelines[4], the European Commission notes that "*Good research data management is not a goal in itself, but rather the key conduit leading to knowledge discovery and innovation, and to subsequent data and knowledge integration and reuse*". Therefore, beneficiaries are explicitly encouraged to make their research data **f**indable, **a**ccessibly, **i**nteroperable and **r**eusable (FAIR). Specific discussions are ongoing in the ODIN project to identify avenues for FAIR data compliance in alignment with both privacy requirements and the expected exploitation of the project's developed solutions. A specific webinar on FAIR data compliance and IPR management will be organized in the upcoming months.

3.1.1 Findability of Data

According to this principle, metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services. For publicly available datasets, publications and reports (deliverables), the ODIN consortium is encouraged to attach or apply a DOI or any other unique identifier. Additionally, all communication and dissemination materials must include the following metadata:

- European Union's Horizon 2020 research and innovation programme letterhead
- Grant agreement No 101017331

For this, and as part of ODIN's communication and dissemination activities, a folder with templates has been uploaded to the common repository.

3.1.2 Accessibility of Data

Once research data have been found, they should be accessible to the user, possibly through mechanisms for access control, such as authentication and authorisation. As outlined in section 5.1.1.4 of the Grant Agreement, the ODIN project participates in the Open Research Data Pilot (ORDP) which aims to improve access to and re-use of research data generated by Horizon 2020 projects and applies primarily to the data needed to validate the results presented in scientific publications. In addition, the project is largely based on open data principles, as defined by the Open Knowledge Foundation[5], which define a set of policies and technical specifications for being compliant.

Project datasets (presentations, reports, scientific publications etc.), intended for public level of dissemination, will be openly accessible through:

- CBMLBox

[4] EC H2020 Programme Guidelines on FAIR Data Management in Horizon 2020, p.3. Available here: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.

[5] https://okfn.org/opendata/.

- Project's website (https://www.odin-smarthospitals.eu)
- Partners' own distribution channels

Most openly accessible data is accessible with a regular browser, with MS excel or a PDF reader.

In this respect, the data management in ODIN will be carried out in accordance with the Grant Agreement (article 29.3), which states that:

Regarding the digital research data generated in the action ('data'), the beneficiaries must:

> *(a) Deposit in a research data repository and take measures to make it possible for third parties to access, mine, exploit, reproduce and disseminate — free of charge for any user — the following:*
>
> > *(i) The data, including associated metadata, needed to validate the results presented in scientific publications, as soon as possible;*
> >
> > *(ii) Not applicable;*
> >
> > *(iii) Other data, including associated metadata, as specified and within the deadlines laid down in the 'data management plan' (see Annex 1);*
>
> *(b) Provide information — via the repository — about tools and instruments at the disposal of the beneficiaries and necessary for validating the results (and — where possible — provide the tools and instruments themselves).*

(…)

As an exception, the beneficiaries do not have to ensure open access to specific parts of their research data under Point (a)(i) and (iii), if the achievement of the action's main objective (…) would be jeopardised by making those specific parts of the research data openly accessible. In this case, the data management plan must contain the reasons for not giving access.

3.1.3 Interoperability of Data

As data usually need to be integrated with other data, they need to interoperate with applications or workflows for analysis, storage, and processing. ODIN's consortium is aware of this challenge, as the different WPs require interoperability of data in order to allow smooth dataflow between partners. Best practices for achieving interoperability are continuously sought by the partners.

Within this context, where appropriate, partners have established and shared data dictionaries regarding the variables they intend to share in order to facilitate interoperability. The project's ontology has been designed in accordance with standardized ontologies and will be further analyzed in Deliverable D3.3.

3.1.4 Reuse of Data

The ultimate goal of FAIR is to optimise the reuse of data.[6] This can be achieved, when metadata and data are well-described so that they can be replicated and/or combined in different settings.

[6] https://www.go-fair.org/fair-principles/.

Given the sensitive nature of the pilots' datasets and privacy requirements as per the GDPR, partners will only be making available data for reuse beyond the project's lifetime where national legislation permits so and only if all adequate measures have been adopted. Such measures may include the full anonymization of the data, an Ethics approval by the respective Ethics Committees and the signature of a data sharing agreement, specifying the purposes, condition, rights and obligations regarding the reuse of the data. When the European Health Data Space (EHDS) Regulation comes into force, the partners will further discuss the possibility and feasibility of making available their datasets for secondary use through the ODIN platform.

Further information will be reported in the final version of this deliverable, so as to better reflect the partners' final strategy with regards to FAIR data compliance.

4 IPR Management

Standard contracts will regulate the management of IPR (Intellectual Property Rights) in ODIN and protect the intellectual property of third parties and beneficiaries involved. The Consortium Agreement contains provisions regarding access rights.

Results from experiments are owned by the beneficiaries or third parties that generate them. Specifically, the product resulting from an experiment will be owned by third parties. Detailed IPR terms and conditions will be stated in the Consortium Agreement.

4.1 Intellectual Property Rights

IPR is a generic term that encompasses several different issues that are covered by different laws and practices. Generally, all issues related to copyright, patents, trademarks, trade secrets and sui generis database rights are collectively indicated as IPR.

IPR management is of fundamental importance in ODIN, because the main software artefacts that the project will release, will be distributed with the intent that Parties, either internal or external to ODIN, can use it freely. In order to grant Parties these rights, it should be carefully managed about the IPR distribution terms.

In the following paragraphs, we will discuss three basic issues about the knowledge the project are producing:

- Access rights: who will own the basic rights?
- Licences: under which conditions is the project going to exchange it?
- Use and dissemination: how will the project exploit it?

4.1.1 Copyright

Copyright is a corpus of laws that are harmonised in most nations in the world thanks to the Berne copyright convention. Copyright laws establish the rights that the authors have over their work. Copyright applies to most original and non-trivial works, be it writings, painting, music, most works of art and even software, both source and machine-readable code.

Copyright concerns the rights of copying, displaying, performing, printing, publishing, extending, modifying, translating a work. Application of copyright to software involves the rights to copy, modify or distribute the program. It does not involve the right to independently write a program performing the same actions as an original one. Generally speaking, the programmer who writes the program owns the rights. Where there is more than one programmer, the Directive (Directive 2009/24/EC) provides for co-ownership.

4.1.2 Patents

A patent is a set of exclusive rights granted by a national or international body to an inventor for a limited period of time in exchange for a public disclosure of an invention.

National legislations and international agreements significantly affect the procedures for granting patents, the requirements placed on patentees, and the extent of the exclusive rights leading to major variations between countries. A patent application must include one or more *claims* defining the invention which must be new, non-obvious, and useful or industrially applicable. The exclusive

right granted to a patentee in most countries is the right to prevent others from making, using, selling, or distributing the patented invention without permission.

Software patents are an important issue, because they can pose a real danger to Free/Libre and Open Source Software (FLOSS) software. When a software method or algorithm is covered by a patent, the patent office has recognised the inventor's claim that the software method is original (never invented before), and non-trivial (a knowledgeable person in the field would not be able to reproduce it from state of the art). The inventors have a 20-years monopoly on the exploitation of the software method, and no one can lawfully use it without their permission.

In Europe the law disallows patents on software per se. The legal status of the European software patents is unclear, though. Application of these recommendations depending a big extent on national laws, which are not harmonized.

FLOSS communities are particularly sensible to this risk, and in fact they avoid as much as possible to use software on which patent claims are known or suspected to exist. Modern FLOSS software licences often contain provisions against the most blatant abuses of software patents. The Apache licence, for example, contains some clauses that protect the software against the use of submarine patents: if a contributor's software is covered by a patent, and that contributor makes legal attacks against users of the software, that contributor loses all rights to using the software. The Mozilla, Eclipse and GPL licences all have some sort of protection against software patents.

4.1.3 Trademarks

A trademark is a distinctive sign, usually a word or a logo. Its usefulness is to give a brand to something and avoid that someone else takes credit for the product using the trademark or distributes a different version of it with the same name.

4.1.4 Database rights

Database rights are sui generis rights that protect the contents of a database, as the law recognizes that a substantial investment is required for the creation of the database, whether that is financial, material or related to human resources. As such, database rights are complementary to copyright protection that protects the structure of a database. Database rights may be of relevance to the ODIN project if datasets are to be shared through the ODIN platform, in order to ensure partners' work remains protected.

4.1.5 Trade Secrets

The most generic way of protecting IPR is to just not let slip the knowledge outside of the boundaries of your organization. For its very nature, this practice is utterly incompatible with FLOSS, which is based on openness. Keeping the development secret is a risky choice, because it can easily give the impression to outsiders that the openness of the developers is just a facade, rather than a real overall policy.

In ODIN, trade secrets would be kept to a minimum, and development should be organised around publicly available repositories as early as practically feasible.

4.2 IPR Management within ODIN

The management of IPR component in ODIN will be handled per the provisions of the DESCA 2020 Model Grant Agreement, as well as per the clauses defined within the ODIN Consortium Agreement.

4.2.1 Ownership of Background Knowledge

Ownership of background IP will not be affected by participation in this project, and it will remain the property of the corresponding participant during and after the project execution. The following principles apply to the use of background knowledge:

1) Access rights will be free of any administrative transfer costs.

2) Access rights are granted on a non-exclusive basis.

3) Background will be used only for the purposes for which access rights have been granted.

4) All requests for access rights will be made in writing.

5) Access rights to results and foreground needed for the performance of the own work of a party under the project will be granted on a royalty-free basis.

4.2.2 Open-Source Access

As part of its commitment to open science, ODIN promotes the development of an Open Access Software. Some of the project's innovations will be evaluated and made openly available for reuse and further development to ensure a level playing field in the market, targeting the production of new open-source software solutions. Hence, open data must be technologically neutral, licensed for reuse at low constraints and documented. Access rights may be granted for research purposes only, as this granting access rights would not include any rights to sublicense or to commercialize the information. The aim of the exploitation will be to ensure the optimal use of ODIN's outcomes and results after the project's completion, speed up the potential of their wider use within the ecosystem, and support in building and expanding the open-source community.

4.2.3 IPR Conflict Resolution

In order to handle IPR issues, which may occur during the project, the following procedures will be undertaken:

1. All project members themselves immediately notify the Project Management Board (PMB) as soon as they are aware of any issue that could be related to ODIN.

2. As soon as an issue of any kind is noticed, the coordinator will nominate one person to be responsible to solve the issue. This person may nominate a task force of persons within ODIN and, discretionally, outside the project, to help in the analysis of the problem and the finding an efficient solution in a timely manner. The nominated person will be responsible for notifying on the progress of the task force.

4.3 ODIN Software IPR Directory

The Software IPR Directory is the document where we store intellectual property information about software. From the exploitation and future business perspectives, it is considered of paramount importance that any piece of software used and produced in ODIN is registered in the Directory.

A respective entry should be created both with regards to background, and foreground work at the start or as soon as possible from the start date. The terms foreground knowledge and background knowledge were defined in the previous version of the present deliverable. Entries in the database contain critical information about copyright holders, patents pending, software licence to be used, distribution terms and willingness to contribute it to further possible initiatives that continues the exploitation of the ODIN assets. Consequently, the IPR directory contains important information from the partners. It is therefore critical that the data entered is reliable and non-refutable.

4.3.1 IPR Initial report

The following table showcases the initial background and foreground of the project, as reported by partners until the closing date of this deliverable.

Table 14: Partners Background and Foreground IPR

Organization	Background (Consortium Agreement)	Foreground: Aligned with Key Exploitable Results
CERTH	Human motion analysis toolkit Object detector module SLAM/mobile platform navigation toolkit Shared human-robot workspace toolkit IoT monitoring platform Data analytics and visualization platform Blockchain-based WoT certification CERTHbot endorsed with functionalities such as social navigation, human detection, tracking and action recognition. Data analytics platform for analyzing sleep monitoring data, and any machine learning methods developed within the project in this context.	Extension of: CERTHbot endorsed with functionalities such as social navigation, human detection, tracking and action recognition. - CERTH's data analytics platform for analyzing sleep monitoring data, and any machine learning methods developed within the project in this context.
CHARITE	Database of sleep recordings from healthy subjects and patients with sleep disorders	None
FORTH	None	None
INETUM	None	None
M&S	None	None
MDT	None	None
MEDEA	None	None
MUL	None	None
MYS	ORVital IoT – Real Time location System	Connector RTLS to Kafka to integrate in ODIN Resource choreographer

	ATLAS	connector and several workflow code and defiinition
PHILLIPS	None	None
ROBOTNIK	None	None
SERMAS	Clinical data lake of Hospital Clinico San Carlos (ECR, Clinical images, Drug Information)	None
SSSA	Sensitive system for increased proximity detection Two patents on the capacitive technology employed by SSSA. Another patent for HOSBOT robotic platform.	None
TWI	None	None
UCBM	Systems and approaches for low-level interaction control algorithms, high level interaction planners, approaches for human-robot interaction and manipulation, high-level software packages for interconnection of sensors, actuators and control units, mechanical tools and interfaces	None
UDGA	None	Certification-related criteria definition
UMCU	None	None
UOW	None	None
UPM	Know-how from R/D and previous research projects (See CA)	None

A specific webinar on FAIR data compliance and IPR management will be organized in the upcoming months in order to best guide partners during the design of their relevant strategy.

5 Data Security

5.1 Technical and Organizational Measures (TOMs) for Safeguarding the Rights and Freedoms of the Data Subjects

According to Art. 32(1) GDPR, the data controller and the data processor should implement appropriate technical and organisational measures (TOMs) to ensure a level of security appropriate to the risk, as well as to the state of the art, the costs of implementation and the nature, scope, context and purposes of processing. The Regulation deems, inter alia, the following TOMs appropriate[7]:

- Pseudonymisation and encryption of personal data;
- Ensuring the ongoing confidentiality, integrity, availability and resilience of processing systems and services;
- Restoring the availability and access to personal data in the event of a physical or technical incident;
- Regularly testing, assessing and evaluating the effectiveness of the TOMs for ensuring the security of the processing.

As a general guiding element, project partners are required to implement and document appropriate technical and organizational measures towards ensuring the security of any data collected, processed, transmitted, disclosed or deleted during the scope of the project. The following sections present an initial set of recommendations for partners to consider alongside those security practices implemented by their organizations. All partners are also invited to consider relevant standards on both data protection and security, including (but not limited to):

- ISO/IEC 15408-1:2009 Information technology — Security techniques — Evaluation criteria for IT security;
- ISO/IEC TS 19608:2018 Guidance for developing security and privacy functional requirements based on ISO/IEC 15408;
- ISO/IEC 27001:2022 information security management;
- ISO/IEC 27701:2019 Security techniques — Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management — Requirements and guidelines
- ISO/IEC 27002:2022 Information security, cybersecurity and privacy protection — Information security controls;
- ISO/IEC 27018:2019 Information technology — Security techniques — Code of practice for protection of personally identifiable information (PII) in public clouds acting as PII processors;
- ISO/IEC TS 27110:2021 Information technology, cybersecurity and privacy protection — Cybersecurity framework development guidelines;
- ISO/IEC 29100:2011 Information technology — Security techniques — Privacy framework;

[7] Art. 32 (1) GDPR.

- ISO/IEC 27559:2022 Information security, cybersecurity and privacy protection – Privacy enhancing data de-identification framework
- ISO/IEC 29101:2018 Information technology — Security techniques — Privacy architecture framework;
- ISO/IEC 29134:2017 Information technology — Security techniques — Guidelines for privacy impact assessment;
- ISO/IEC 29151:2017 Information technology — Security techniques — Code of practice for personally identifiable information protection;
- ISO/IEC 29190:2015 Information technology — Security techniques — Privacy capability assessment model

Furthermore, all Partners acting as data Controllers should ensure compliance with national and regional requirements for personal data processing, and should consider both EDPB and national authority guidance when developing and deploying their respective research actions in the context of the ODIN project. In particular, all partners must consider the content of the following EDPB guidance whenever relevant to their actions:

- Article 29 Data Protection Working Party - Guidelines on Data Protection Officers ('DPOs') (wp243rev.01) (10/2017)
- Article 29 Data Protection Working Party - Data Protection impact assessments High risk processing (10/2017)
- Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01)
- Guidelines 3/2019 on processing of personal data through video devices
- EDPB Guidelines 4/2019 on Article 25 Data Protection by Design and by Default
- EDPB Guidelines 03/2020 on the processing of data concerning health for the purpose of scientific research in the context of the COVID-19 outbreak
- Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak
- EDPB Guidelines 05/2020 on consent under Regulation 2016/679
- EDPB Guidelines 07/2020 on the concepts of controller and processor in the GDPR
- EDPB Guidelines 01/2022 on data subject rights - Right of access
- EDPB Guidelines 02/2021 on virtual voice assistants
- EDPB Guidelines 08/2022 on identifying a controller or processor's lead supervisory authority

Project partners have been granted free access to the Europrivacy Academy (https://academy.europrivacy.com/), where their experts and DPOs can receive training on the Europrivacy certification scheme, currently the only EU Data Protection Seal recognised under Art. 42 GDPR.

The table below represents the TOMs each partner of the consortium has undertaken to secure their data processing and to safeguard the rights and freedoms of the data subjects, whose data are being processed.

Table 15: Technical and Organizational Measures for Data Subjects' Rights

Partner	TOMs
Robotnik	The robot is protected by different security layers regarding network, code, and access.
SERMAS	• Data back-ups. • The server will be located in a secure area of the hospital with restricted physical access. • Different users and user rights will be defined for the project members accessing the data remotely.
MYSPHERA	Secure access through role base permission. Data is stored with version management and data loss protection.
M&S	MINDS & SPARKS GmbH ensures that both physical and technical measures will be taken for the protection of the personal data which are going to be processed during the lifetime of ODIN. Internal policies and confidentiality agreements will safeguard the data as well as secured storage where only authorized access is allowed with controlled password-protected access (see above).
MEDTRONIC	• Data stored in documents is regularly backed-up to Medtronic's OneDrive servers • Miro boards data, although available through Miro website, is backed up in Medtronic's internal servers • Website data
CERTH	In conformity with international (GDPR) and national law a number of technical and organizational measures (TOMs) will be initiated and implemented. For instance, among the technical measures that are defined to be implemented is to conduct regular back-ups in order to avoid unexpected data loss, enable physical and virtual secured storage and access to data. Particularly sensitive data will be stored in an anonymised form, explicitly stated in the consent form and in the context of the overall consent procedure. Access rights to this data will be clearly stated in relevant documentations.
FORTH	FORTH and the technology that is used is GDPR compliant for all data used. The private cloud facilities are ensuring data security and privacy.
University of Warwick (UoW)	UOW is not storing data on their premises but on the Cloud Services provided by FORTH.
SSSA	This aspect does not concern WP5 and SSSA activities because data will not be stored by SSSA, but will be transmitted to the ODIN's ICT platform/infrastructure.
THL	Equipment used in the project's data processing activities (including laptops etc) is password protected. Regular weekly back-ups are conducted.
Philips Electronics Netherland BV (PEN)	PEN will not store data in their infrastructure based on initial discussions with pilot sites. Data will be accessed at the clinical sites.

UPM	Security measures are going to be defined but in principle it is foreseen high availability of data services that implies backups, data redundancy, as well secure storage with encryption at rest and access control are expected for securely accessing the data only by authorized users.
INETUM	At this stage, there is no data collection, processing or sharing envisioned on behalf of INETUM.
UCBM	Thanks to its previous experience, UCBM will adopt several good practices to ensure data management in complete safety. In particular, the data will be password protected and access will be allowed only to a small group of the UCBM team. Anyone who works with confidential electronic data should identify themselves when they log on to the PC or laptop computer that gives them access to the data and the list of users will be kept up to date. Furthermore, only secure methods of data transfer will be used and systems for the secure destruction of the same will be adopted. Furthermore, all UCBM personnel are constantly trained and updated on the best practices to be adopted for data management. PCs and laptops will be used for short-term storage and data processing. In no case should these be relied upon for storing master copies, unless backed-up regularly. A private back-up area will be identified for research data collection, and it will be set and configured to act as the only backup area for any relevant project file. The file repository will be duly sized. Moreover, all data will be backed up and securely encrypted with a cadence to be defined.
UMCU	UMCU is ISO27001 certified (IT dept) and 9001 certified (lab).
Charite (CUB)	The hospital firewall is maintained by the IT department of the Charite university hospital. All computers in use are protected and administered by the university hospital IT department. One of our IT department members is member of the Charite group involved in ODIN.
MUL	Data access is available for researchers at the Department of Family Medicine, Medical University of Lodz based on the MUL data management policy and binding national regulations. Backups are made on the continuous basis by MUL IT department for all study-related data.
UDGA	Access to the data processed for UDGA's activity is secured with password, granted through an authentication mechanism only to participants in the consortium.
MEDICT MEDEA	MEDEA - Conduction of regular back-ups to avoid unexpected data loss; physical and virtual secured access to data.

6 Ethical and legal aspects

6.1 Task Management within the Project and the Pilots

Task management within the project is dependent on the partners' roles within the project. As such, each partner, including pilot owners, needed to clarify *who is in charge of what*, and more specifically, who are the data controller(s) and the data processor(s) and are there established joint controllerships. Pilot owners needed to clarify and explicitly define how their work is organized, to know and communicate what personal data are/will be collected and thus to allow a data flow mapping, as is reported in Section 2.1.3 of the present deliverable. In order to facilitate the identification of the roles of the partners, the previous Data Management Plan had already provided a guide to help partners determine what their role in the personal data processing is, and what obligations ensue.

For the definition of the data processing activities, partners have been provided with Data Management Questionnaires (Appendix A). As the work on local and general data management plans is understood as an ongoing work, the exercise in data mapping will continue over the course of the project. The figure below provides a high-level overview of the distribution of responsibilities and tasks among the different work packages.

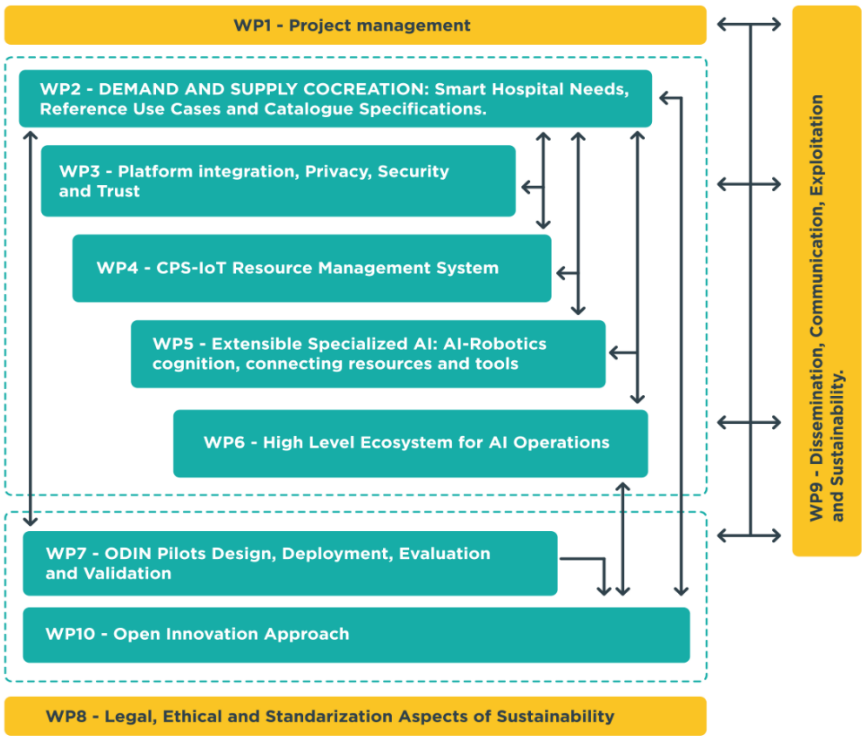

Figure 18: Work Package Governance

6.1.1 ODIN Ecosystem of Partners

The outcomes of T2.1 "Co-creation strategy, stakeholders' definition and mapping" of Work Package 2 will complement the identification process of the roles and obligations within the consortium. The following figure offers a comprehensive map of the different clusters in the consortium and partners can determine, according to their expertise, to which of the "circles" they belong, what are other stakeholders with similar profiles and explore the dimensions of the work related to their activities.

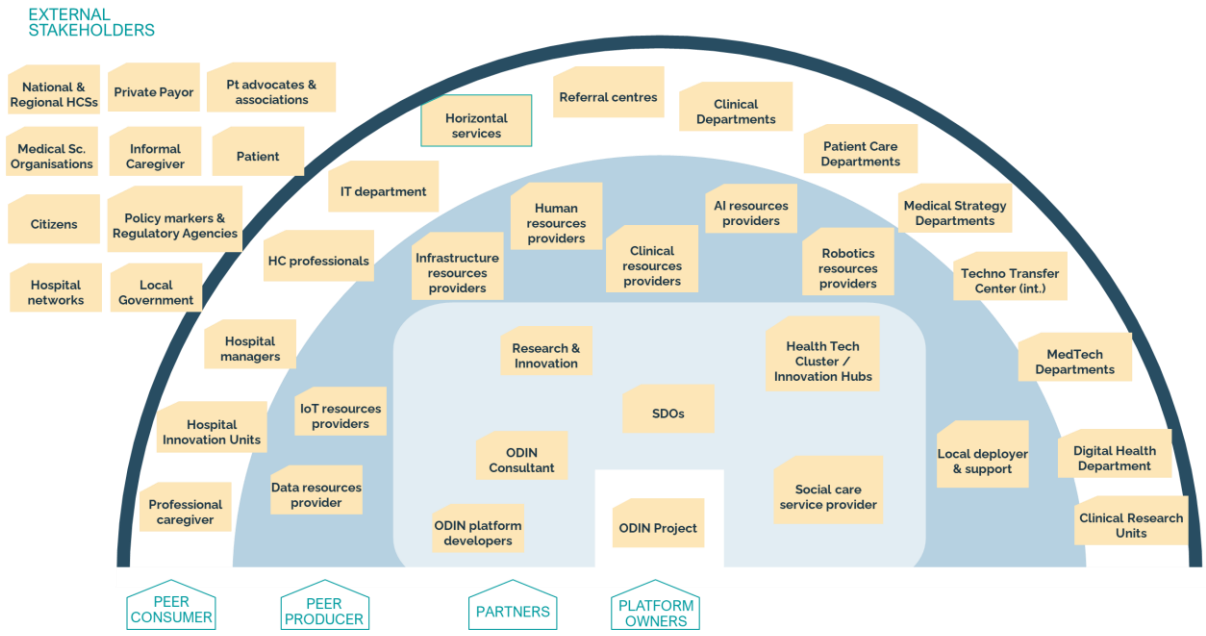

Figure 19: ODIN Stakeholder Analysis

Source: Medtronic, STAKEHOLDER MAPPING WORKSHOP (2nd Plenary Meeting)

6.1.2 Data Protection Officers (DPOs)

The GDPR defines the role and responsibility of a Data Protection Officer (DPO), Art. 37 GDPR. The DPO is in charge of monitoring the application of the GDPR within an organization and providing strategic advice to it on how to process personal data while respecting individuals' rights. Each Data Controller (pilot) should have a clearly identified DPO. ODIN also has appointed a DPO for the project, represented by the Ethical and Trusted Data Manager (ETDM) (UDGA) position currently held by MA. M.Sc. Adrian Quesada Rodriguez, in charge of:

- Establishing common rules and requirements for the consortium data protection policy;
- Coordinating the action and information among the various DPOs and organize, when needed, regular calls among the DPOs of the different data controllers;
- Serving as an entry point to answer questions and complaints from third parties when addressed to the project as a whole;
- Providing guidance on how to implement the privacy by design and by default principles.

Local DPOs should report and work in close coordination with the researchers responsible of the different pilots and the project's DPO. Thus, providing information about local DPOs in the data management questionnaires, and keeping it up to date, is of utmost importance.

6.2 Controller identification and initial instruction definition

The ODIN project's implementation is divided into four phases, showcased in Figure 4 below.

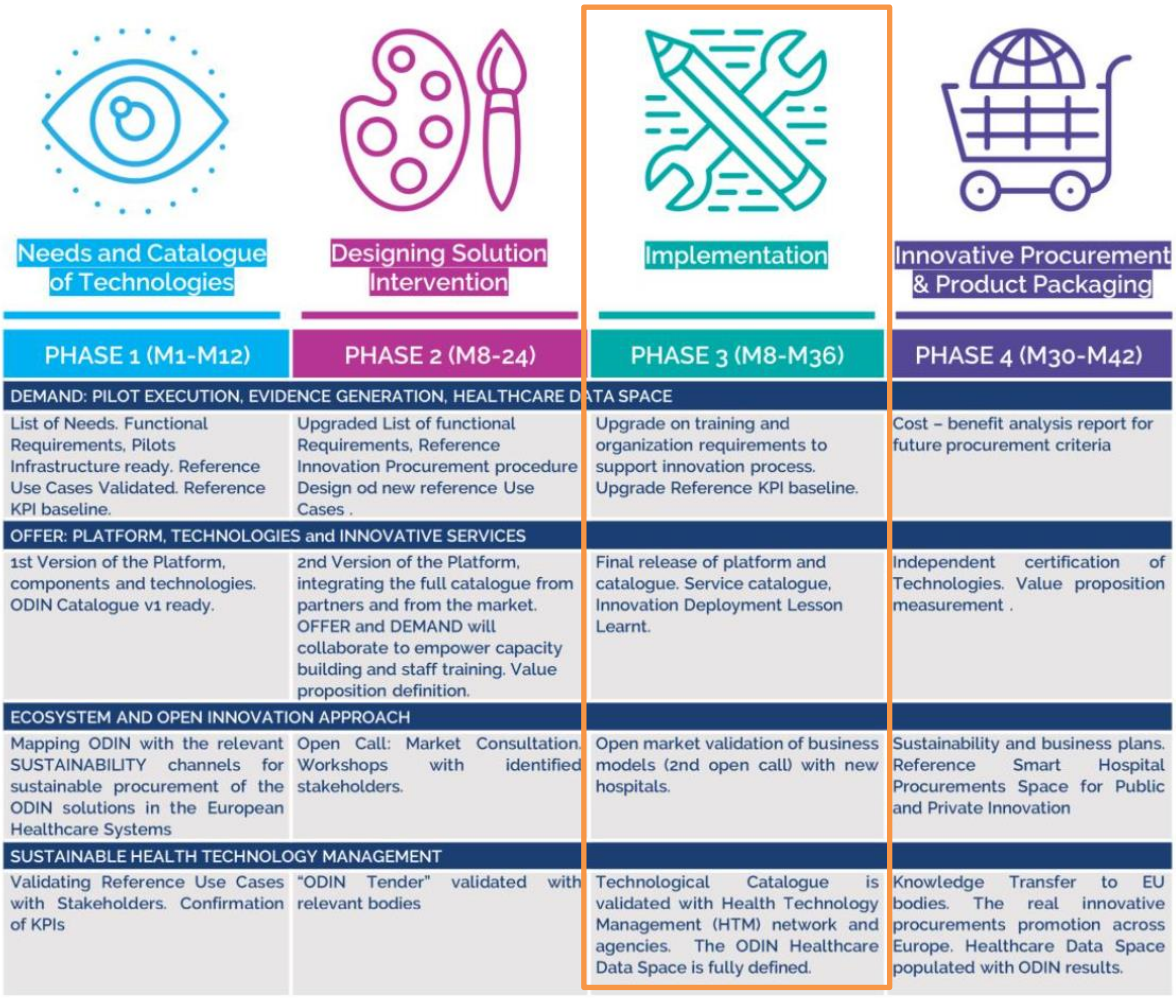

Figure 20: The ODIN Innovation Procurement Phases

During the first phase, the project identifies the demand in terms of pilots' execution, evidence generation, and deployment of healthcare data space by compiling a list of needs, functional requirements, and preparing and validating reference use cases. An initial version of the platform, its compartments and technologies shall be prepared. The definition of the technology is a crucial step of the process, as it will allow to determine which privacy enhancing and safeguarding mechanisms need to be implemented, and how, i.e., deployment of an anonymization strategy. As the definition of technology is still ongoing, the document will be continuously updated, and identified risks and needs regarding the deployed technology will be described, and mitigation strategies will be proposed.

In a second phase of the project, new infrastructure/technology, according to the identified demand, will be integrated to the platform. Pilot solutions for eWorkers, eRobots, eLocation will be offered. The first round of open calls will ensure the necessary capacity building and training through market consultation and workshops with identified stakeholders. For the carry-out of the workshops and interviews, which are related to collection of data, the previous version of the data

management plan offered privacy-compliant approaches and mitigation strategies. The ethics management questionnaires, completed by pilot owners, additionally provided details regarding the recruitment of participants.

During the third phase of the project, the platform and catalogue shall be released. The technological catalogue will be validated with the Health Technology Management (HTM) network and agencies, in order to ensure a sustainable management of the technology involved. As such, the ODIN Healthcare Data Space shall be fully defined. A second open call will also take place in order to launch the market validation of the ODIN-designed business models by new hospitals.

Health providers' requirements and interests are two-fold: On one hand, there is a direct interest to ease access to data collected by healthcare providers to citizens and third parties, who could develop services out of it. On the other hand, there is a strict need to respect and preserve the privacy and the personal data of the participants. While the project will support the access to open data by third parties, it will abide to strict personal data protection policy in line with the EU General Data Protection Regulation (GDPR) and other applicable norms, including the upcoming European Health Data Space Regulation. Personal data protection compliance is part of the project's requirements and will guide the architecture design. Personal data protection principles will determine and limit the data sharing. ODIN is committed to proactively ensure full compliance with the GDPR through a set of ad hoc policies, mechanisms, and tools.

Moreover, the project commits to strictly stick to the principle of data minimization by avoiding the collection and processing of any unnecessary personal data. Personal data can be kept in a form which permits identification of data subjects for no longer than is reasonable, proportionate, and necessary for the purposes for which the personal data are processed.

The core providers of (sensitive) personal data in the framework of the project will be the Pilot owners, which will act as data controllers for any data to be provided to the consortium from their infrastructure and/or network and are the key responsible organizations to ensure compliance with ethical and data protection requirements throughout the design and implementation of the processing activities (e.g.: DPIA, data subject right protection, etc.). The implementation phase of the pilots will enable the controllers to validate the means and purposes of any processing to be performed and to identify whether any other partner should be granted access to the data. The project is already in discussions regarding the adoption of the relevant Data Sharing Agreements, which will be reported on the final iteration of this deliverable.

As was analysed in detail in the previous version of the Data Management Plan, there is a set of general instructions that Data Controllers must take into consideration under the individual Controllers and Controller to Controller data transfer scenarios. Said guidelines mainly focus on purpose limitation, data minimization, the obligation to ensure data subjects are adequately informed and can effectively exercise their rights, as well as the adoption and implementation of adequate technical and organizational measures in order to ensure personal data remains protected at all times, which may also be accompanied by audits.

Similarly, data Processors under the Controller to Processor data transfer scenario are bound to follow additional guidelines adapted to this scenario. In this context, it is highly important that the Processors provide sufficient guarantees to implement technical and organizational measures, as well as to refrain from engaging sub-processors without the Controller's prior authorization. They shall additionally take all appropriate measures to assist the Controller to fulfil their obligations and shall act only in the name of and in accordance with the Controller's instructions. Finally, the Processor shall return all personal data to the Controller once the relevant task has been concluded and shall delete all copies.

Contractual activities envisioned:

In the context of the ODIN project, work is underway to develop a Unilateral Commitment tool that will ease compliance with requirements established data processing agreements. The final iteration of the deliverable will report on these developments.

6.3 Project ethical risk assessment and Data Protection Impact Assessment (DPIA)

A DPIA is a process designed to describe the data-processing, assess its necessity and proportionality, and help manage the risks to the rights and freedoms of natural persons by assessing them and determining the measures to address them.

The DPIA is required when the processing is "likely to result in a high risk to the rights and freedoms of natural persons". "The rights and freedoms" of data subjects primarily concerns the rights to data protection and privacy but may also involve other fundamental rights such as freedom of speech, freedom of thought, freedom of movement, prohibition of discrimination, right to liberty, conscience, and religion.

The previous version of the Data Management Plan thoroughly described WP29's Guidelines regarding the need to perform a DPIA, the need to consult with the Supervisory Authorities, the common criteria on the methodology for performing it, as well as certain lessons learned by national legislations. Finally, the previous version of the deliverable thoroughly explains the criteria that need to be considered in order to determine compliance with the GDPR requirements.

In view of the above, it is worth highlighting anew that the WP29 confirms that a **single DPIA could be used to assess multiple processing operations that are similar in terms of nature, scope, context, purpose, and risks**. This might be the case where similar technology is used to collect the same sort of data for the same purposes. This may also be applicable to similar processing operations implemented by various data controllers. In those cases, a reference DPIA should be shared or made publicly accessible, measures described in the DPIA must be implemented, and a justification for conducting a single DPIA must be provided. Moreover, the said DPIA should set out which party is responsible for the various measures mentioned, while each data controller should express their needs and share useful information without either compromising secrets or disclosing vulnerabilities.

6.3.1 ODIN Ethical Risk Assessment:

The Grant Agreement (Page 142) requires the evaluation of the ethics risks associated with the ODIN project and the development of an opinion on whether a DPIA is required. In order to determine the potential ethical impact of the ODIN project, the key ethical principles were originally identified in the previous version of this deliverable, including confidentiality and privacy, beneficence, justice and respect for persons, transparency and sustainability.

In order to better understand potential risks with regards to privacy, it is necessary to review a **summary of personal data collected by ODIN Partners**, as follows:

- Robotnik: no personal data collected
- SERMAS: Personal data: WP7 UC 1: information concerning material and equipment consumptions and purchases for a yet to be defined medical procedure, data from the patients who undergo that procedure (sociodemographic, hospital intake and leave dates,

procedure outcomes, successive hospital intakes; WP7 UC7: Video image of a hospital area (either emergency service or a surgical area); geographical position of equipment and personnel/patients (RFID).

- M&S: WP9 contact list
- MySphera: no personal data collected nor processed
- THL: no personal data collected nor processed
- CERTH: WP5 T5.2, video (hospital areas), depth information (derived from video)
- FORTH: WP6 retrospective (data that have already been collected by patients from the clinical partners serving as data providers) or prospective patient data for generation of AI models to fulfill the pilots requirements (Prospective are new data from new patients of the clinical partners (i.e., patients for whom no data has been collected and thus provided as part of the retrospective data)).
- UMCU: UC3: patient data from patients that are using the Luscii app
- INETUM: At this stage, there is no data collection, processing or sharing envisioned on behalf of INETUM. UPM: T2.4: participant data through interviews and workshops; t7.1 pilot data from pilots (questionnaires); T10.3: data from open call submissions (participant forms); T3.3.: user data (credentials); T4.6: metric data (logs, IP address); WP5: interaction to users with social robots.
- CUB: WP7/WP8/WP9. Questionnaires including medical data and economic data, interviews, sleep-related data obtained through monitoring devices. More information will be provided in the final iteration of the Data Management Plan, as the definition of the useful datasets is ongoing.
- PEN: de-identified patient data from WP6 and pilots for analytic and AI model generation.
- UoW: Pilot representative information
- UCBM: WP7: data collected by y clinical staff for registration of consent to participate in the experimental studies (name, surname, age, patient status, pathology, consent to participate, etc.) in accordance with the provisions of the UCBM Ethics Committee and the current laws about processing of personal data and conducting clinical studies (WP7). Furthermore, there will be other data collected automatically by robotic/AI systems during the carrying out of the validation activities foreseen in the ODIN project (i.e. WP7). They include, for example, trajectories traveled during navigation tasks and performed during grasping tasks by the robot, patient ID, face recognition data, annotation of the level of food intake by the patient, score obtained by the patient during the execution of tasks rehabilitation, compliance with oxygen intake prescriptions (WP5 and WP7).
- SSSA: WP5: data concerning robotics activities. Data will be generated into the WP5 also data coming from other technical WPs will be processed. Data will be used in real-time or wired/wireless transmitted for being stored into Hospital's ICT infrastructure. So far, Data to be considered into robotics activities regard: 1) Data coming from cameras and used for human awareness, robot navigation, human detection and tracking, social interaction models, monitoring and security, human action and behavioral recognition, human-robot interaction, etc.; 2) Data coming from sensors for localization of devices and robots that could be transported by people or wearables for cognitive performance monitoring and user's state estimation (stress, cognitive load, sleep quality, etc..); 3) Sensitive data of patients and workers that are transmitted/processed through robotic modules that come from human-machine interfaces installed in the robots (for accessing to services or registrations) or coming from the Hospital's ICT infrastructure (other WP's).

- MEDEA: No personal data collected nor processed
- UDGA: Partner representative information
- MDT: Partner representative information (WP2, WP8, WP10), interview results (WP2), data collected through workshops and open calls (WP10)
- MUL: IoT Data from tagging devices (WP7); Data related to clinical staff and patients (WP7); Electronic health record data (WP7) towards development and implementation of AI solutions (UC2, UC6; UC7). No personal data is envisaged within the those datasets but only machine generated data.

Key outcomes:

Identified risk factors requiring the performance of a DPIA by the controllers include the collection and processing, by some of the ODIN project partners, of sensitive health data (EHR data). The project's nature and the technologies used to fulfil the DoA requirements intrinsically fall under the "innovative use or applying new technological or organizational solutions" mentioned by EDPB. This element is sufficient to form an informed opinion regarding the need to perform a DPIA.

6.3.2 ODIN Opinion regarding need to perform DPIA:

Based on the results of the ODIN Ethical Risk Assessment, the guidance from EDPB, and the express dispositions on this subject found in the Grant Agreement Section 5.1.1.1 (Page 310), every Data Controller in the ODIN project will be required to perform a DPIA. The results of the performed DPIAs will be shared with the project's Ethical and Trusted Data Manager and with the Legal and Ethics Board for discussion, evaluation, and documentation. The results of this DPIA will be reported in the final iteration of this DMP.

6.4 Consent forms

Even though it is not only the explicit consent of participating individuals in medical research that should constitute a legal basis for a data processing (Art. 6(1)(b-d) GDPR, Art. 9(2) GDPR), the "ethical requirement" of an informed consent should be met either way.[8] The statements according to the ethical standards and bio-ethics conventions primarily aim to protect individuals against being included in medical research activities against their will and/or without their knowledge. Hence, sufficiently informing participating individuals about their engagement in scientific (or medical) research activities is imperative. On a more theoretical note, the ethical requirement of "informed consent" is to be distinguished from the consent as a legal basis for lawful data processing in Art. 6(1)(a) GDPR.

[8] EDPB Document on response to the request from the European Commission for clarifications on the consistent application of the GDPR, focusing on health research, paragraph 7, p.4. Available here: https://edpb.europa.eu/sites/default/files/files/file1/edpb_replyec_questionnaireresearch_final.pdf.

The provisions of the Oviedo Convention[9], along with those of the Declaration of Helsinki[10], outline the essential information for prospective research participants to obtain an informed consent, mainly focusing on maintaining the research subjects' privacy, while ensuring a free and informed consent, providing all required information to participants of age. The participants rights, especially with regards to information and access to the clinical study data, are of utmost importance, along with the protection of vulnerable individuals. The previous version of the deliverable analyses those requirements more in depth.

6.5 Findings

All the researchers involved in ODIN will comply with and follow the principles outlined by the GDPR, European and International Instruments in the fields of data protection, and ethical provision on protection of individuals with regard to the processing of personal data and on the free movement of such data. Medical data will be deposited in anonymised form which is allowed by the informed consent signed by study participants from the user groups. All participants are made aware if their collected will be shared with other research collaborators; nevertheless, they are ensured that their personal data is kept confidential at all times. The databases that hold confidential data on participating subjects sit on a secure network and do not have an internet (HTTP) connection so as not to compromise the data. Furthermore, technical procedures are in place to monitor what data is entered and exported to ensure there is no breach of this. Measures will be taken to ensure data security at each pilot site.

The project invites the pilots and takes itself into consideration the best practices developed in the context of the other LSPs projects.

6.6 Main Principles and Concepts of Data Management

All partners are invited to consider the following information throughout their project-related activities, as were previously discussed in detail in the first version of the Data Management Plan:

6.6.1 Guidelines for GDPR Compliant Deployment of AI, IoT and Robotics in Pilots

As has already been explained, the management and exchange of information within ODIN pilots, i.e. the hospitals, has been recognized to be deeply heterogeneous, due to the implementation of various functionalities, different data representations, user interfaces, terminologies, etc. The interoperability, which enables the heterogenous structures to interfere, is achieved through health data exchange standards and protocols, common middlewares, and specific standards for

[9] Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine, Oviedo, 4.IV.1997. Available here: https://rm.coe.int/168007cf98.

[10] WMA DECLARATION OF HELSINKI – ETHICAL PRINCIPLES FOR MEDICAL RESEARCH INVOLVING HUMAN SUBJECTS. Available here: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/.

the management of domains and servers.[11] As such, the ODIN project aims at developing and/or integrating technologies to reconstruct the surrounding environment and retrieve useful information to monitor the quality of the environment, model the human behavior, enhance human-robot collaboration and increases cognition capabilities from ODIN platform technologies.[12] In this context, there will be 3 types of communication architectures: Robotics, AI and IoT, as demonstrated below:

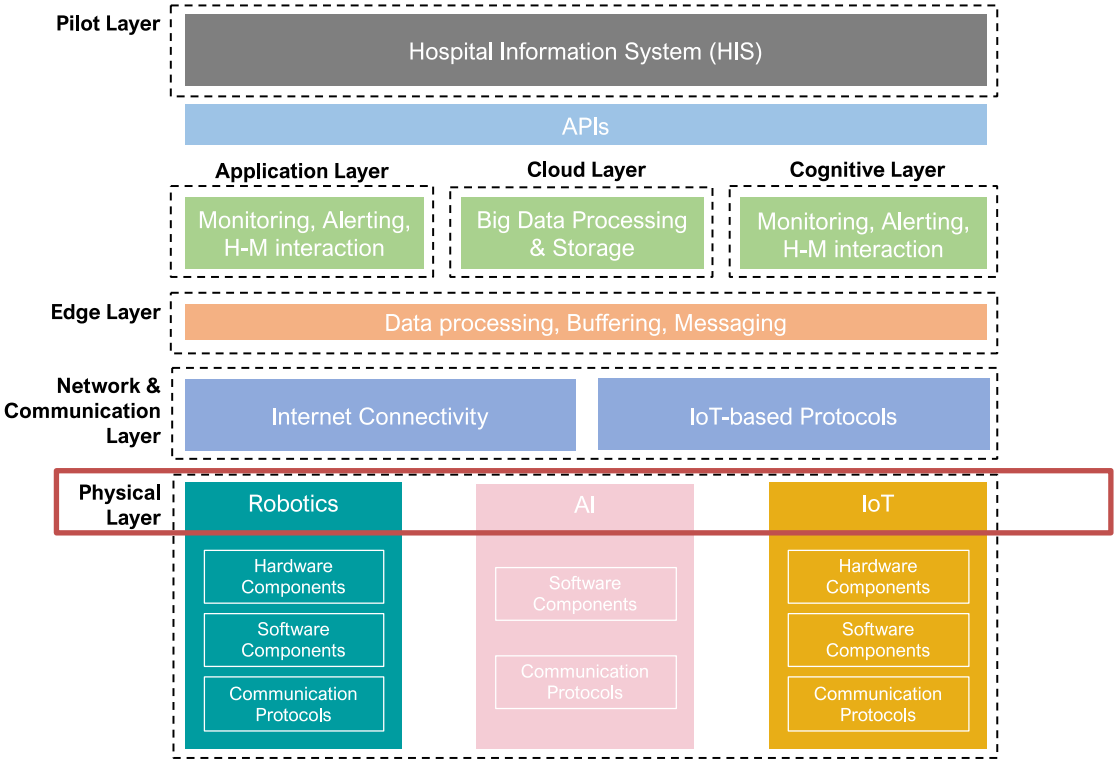

Figure 21: ODIN Platform blueprint (Source: D2.2)

The architecture of the physical layer is multi-layered and, partially, very complex. On top, the architectures are not strictly distinct and can merge, creating for example an Internet of Robotics Things (IoRT) infrastructure, combining IoT and Cloud Robotics technology. Its accelerated development is enabled by Artificial Intelligence of Things (AIoT) to improve humans-machines interactions and enhance data management and analytics. Such human- robot collaborative environments need to ensure the safety of human beings, but also to guarantee their rights and freedoms.

Such Large-Scale-Pilots (LSPs) have been long examined, tested and validated in diverse application domains by the research community. In view of this, a number of guidelines have been

[11] ODIN_D2.2_Hospital_Requirements_Report_v.1.1_revTHL: section 2.2.4.

[12] ODIN_D5.1 - Context_awareness_human_modelling_v05; Section 4.1.

established as best practice to guide the consortium and the pilots in deploying different solutions, as were previously reported[13].

The above-mentioned guidelines focus on data minimization, purpose limitation, personal data transfer minimization, storage and retention limitations, as well as employing all adequate technical and organizational measures, including anonymization and pseudonymization techniques. Designating a DPO and ensuring partners' DPOs are also involved in the procedure are also recognized as best practices. Similarly, organizing training activities, performing a DPIA and adopting and enforcing a data protection policy, an access rights policy, as well as a policy for updating the firmware of IoT are also crucial. Of course, abiding by GDPR requirements and data protection principles, as well as benefiting from of online commitment tools to ensure that all partners located in other jurisdictions are committed to respect the same level of data protection (i.e., privacypact.com) are highly relevant.

Importantly, following the developments around the COVID-19 crisis, new guidelines were introduced. For reference, on April 21st, 2020, the European Data Protection Board (EDPB) approved the Guidelines 3/2020 on the processing of data concerning health for the purpose of scientific research in the context of the COVID 19 outbreak.[14]

6.7 Data Protection Principles

Data management is an ongoing process and the current data management plan, as outlined above, is meant to be a live document. In view of this, the present deliverable builds upon the work already performed in its previous iteration and recognizes the ongoing need to respect the following principles, as were originally identified:

A. Lawfulness and Fairness, focusing on the identification of a valid legal basis, as per Art. 6 (1) and 9 GDPR, and the processing of data in a way that is not unjustifiably detrimental, unlawfully discriminatory, unexpected or misleading to the data subject;

B. Transparency and provision of sufficient information to data subjects;

C. Purpose Limitation and Presumption of Compatibility of any Further Processing purposes;

D. Data Minimization and Storage Limitation, also referring to national obligations and guidelines;

E. Accuracy and, where required, updates of the data;

F. Integrity and Confidentiality, ensuring appropriate security against unauthorised or unlawful processing and accidental loss, destruction or damage, using appropriate technical or organisational measures. Said security measures must be designed in accordance with the nature of the data as sensitive data related to the data subjects' health;

[13] The list of principles is taken by S. Ziegler and others, *Personal Data Protection for Internet of Things deployment: Lessons learned from the European Large-Scale Pilots of Internet of Things*, February 2020, pp. 30-31.

[14] https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202003_healthdatascientificresearchcovid19_en.pdf.

G. **Accountability,** ensuring not only that adequate measures are adopted but also that compliance can be demonstrated.

The ODIN project demonstrates its commitment towards compliance with the data protection rules and GDPR regulations and commitment to promoting EU fundamental rights by dedicating a Work Package (WP8) to Legal, Ethical and Standardization Aspects for Sustainability. However, each partner remains responsible for *its actions, for compliance with the GDPR and safeguarding EU fundamental rights.*

6.8 Ethical Principles

ODIN as a project, its consortium and the actions related to the project will comply with ethical and legal principles, standards and regulation. This includes undertaking activities in compliance with ethical principles and applicable international, EU and national law. The most important guiding principles are outlined in this section. The deliverable D8.2 "Policy, Legal and Ethics Framework" provides a comprehensive mapping of the most relevant initiatives, which should be drawn to support the current document.

Potential ethical issues will mainly concern data protection. Within ODIN, any research involving human subjects, data processing, and sensitive data processing will conform to applicable legislation and regulations both on European level, as well as to complementary obligations of the countries where the activities will be carried out.

The Good Clinical Practice guidelines[15] are in agreement with the Declaration of Helsinki. Work done by partners and its beneficiaries will also conform to relevant EU legislation, such as:

- The Charter of Fundamental Rights of the EU (specially Art.3: right to the integrity of the person; and Art. 8: protection of personal data)
- Directive 2001/20/EC of the European Parliament and of the Council of 4 April 2001 on the approximation of the laws, regulations and administrative provisions of the Member States relating to the implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use;
- EU General Data Protection Regulation 2016/679;
- Treaty on the European Union (TEU): Article 6;
- EU Charter of Fundamental Rights of 7 December 2000;
- Medical Device Regulation (EU) 2017/745 for the implementation of the system in the public health environment in a secure setting.

In order to protect the privacy rights of participants, a number of best practice principles will be followed. There are two basic components to the ethical standards: (i) **informed consent** and (ii)

[15] COMMISSION DIRECTIVE 2005/28/EC of 8 April 2005 laying down principles and detailed guidelines for good clinical practice as regards investigational medicinal products for human use, as well as the requirements for authorisation of the manufacturing or importation of such products. Available here: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32005L0028&from=EN.

independent ethical oversight.[16] Leveraging on *International Ethical Guidelines for Health-related Research Involving Humans*[17], ODIN's consortium has identified and implemented additional ones, in the light of the specific identified risks for the project, and its needs. These include:

1) Informed Consent, that is prior, freely given and specific. Researchers must provide all relevant information to the participants beforehand, as well as all clarification where required.
2) Approval by Research Ethics Committees, as reported in Deliverable D11.1.
3) Scientific and Social Value, conducting studies that are scientifically sound, build on an adequate prior knowledge base, and are likely to generate valuable information.
4) Purpose Limitation.
5) Data minimization, not retaining any ancillary data obtained and anonymizing all personal data when possible.
6) Use of Data obtained from the Online Environment and Digital Tools in Health-related Research, focusing on the adoption of privacy-protecting measures and the mitigation of risks that could result from combining data from multiple sources and their subsequent use and publication.
7) Reimbursement and Compensation for Research Participants, such as travel costs, and compensated reasonably for their inconvenience and time spent, whether monetary or non-monetary, following the approval of the local ethics committee.
8) Recruitment of Affiliated Participants and the protection of their privacy, confidentiality, ensuring their non-discrimination.
9) Privacy and Confidentiality, specifically with regards to the measures to prevent re-identification of the data subjects.
10) Data Sharing, in accordance with legal requirements.
11) Vulnerable Persons and Groups, placing additional safeguards for their protection.

6.9 Data Subject Rights

Recital 59 GDPR provides that modalities for facilitating the exercise of the data subject's rights, such as mechanisms to request and obtain free of charge access to, request rectification or erasure of personal data shall be set. Further, the data controller should also make such exercise of data subjects' rights possible electronically and should be obliged to respond to requests from the data subject without undue delay.

[16] EDPS A Preliminary Opinion on data protection and scientific research, p.14. Available here: https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf.

[17] International Ethical Guidelines for Health-related Research Involving Humans, Prepared by the Council for International Organizations of Medical Sciences (CIOMS) in collaboration with the World Health Organization (WHO), Geneva 2016. Available here: https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf.

All participants are made aware of their right to withdraw from the research activities without providing any reason for their withdrawal and that they retain this right at all times. If any of the research activities they are participating in involves audio recording or electronic note taking, they will be notified that they can ask the interviewer to stop or delete all or a portion of the recorded material at any time. Participants can also request that content is erased retrospectively. Invited individual participants are briefed about this right through the informed consent process. Withdrawal can be also expressed orally.

Following, below are the data subjects' rights, as per the provisions of the GDPR, which the ODIN's consortium commits to respect and guarantee, as were previously analyzed:

a. Right to access their personal data that is collected and/or processed;
b. Right to information about all relevant aspects surrounding the collection and processing of their personal data, unless the data was not collected directly by the data subjects and if it "proves impossible or would involve a disproportionate effort, in particular for processing for scientific research purposes when the conditions of Article 89 are satisfied or when this is likely to render impossible or seriously impair the achievement of the objective of that processing".;
c. Right to rectification of their personal data;
d. Right to object to the processing of their personal data;
e. Right to erasure of their personal data;
f. Right to restriction of processing of their personal data;
g. Right to data portability of their personal data in a "structured, commonly used and machine-readable format";
h. Rights related to Automated Individual Decision-making and Profiling, focusing in particular to the data subjects' right to not be subject to automated Individual Decision-making and Profiling where that may have a legal or similar effect for them.

7 Conclusion and Future Plans

The current deliverable includes the best available information on data processes at the project level and at the pilot level in the current stage of development of the ODIN project. As is evident, pilots have been considering data management and ethics requirements from the start of the project and have developed strategies in order to ensure data remains secure and confidential. In particular, they have heavily focused on the use of strict anonymization and pseudonymization procedures in order to ensure that data subjects cannot be identified.

This deliverable is the result of a collaborative work with representatives of the different WPs and of the different pilot sites. It is to be understood as a living document that will be constantly updated during the next steps of the project and thanks to the contribution of the different participants to the project. Particular attention will be also devoted to the issues arising from data sharing in the context of multisite research. The data management plan of the different pilots will also be constantly monitored and updated. The updated versions of the deliverable will be made available according to what has been agreed in the DoA.

Important implications from the work done can be already highlighted:

- Data (personal and non-personal) play an increasingly important role in the medical research domain and e-health;
- The pilots, as data controllers, also have to guarantee the flow of data to the platform. Further work will be required to define and design the Data Sharing Agreements (DSA);
- Non-personal data have also to be taken into consideration especially as far as IPR and licensing are concerned;
- Where possible and appropriate, data must be shared in a FAIR manner in order to comply with the relevant principles;
- Additional training and webinars must be provided to partners in order to best guide them and assist them with designing their FAIR data and IPR strategies respectively;
- This document is understood by the Consortium as a living document; it will be constantly updated through the involvement of the pilots in order to best reflect the project's status at each stage.

Appendix A Data Management Questionnaire

PART A - DATA SUMMARY

1. What is the purpose of your data generation/collection/processing and how is it related to the objectives of the project?

.....
.....
.....
.....

2. What types and formats of data will you generate/collect/process within the project? Please list below.

.....
.....
.....
.....

3. Are you in charge of making decisions about what data to be collected/ processed, how and for what purpose?

☐ Yes, namely .....

☐ No

4. Do you process data under another partner's behalf/instructions?

☐ Yes, namely .....

☐ No

5. Do you currently or will you in the future share data with other partners inside the project? If yes, please list below:
   - to whom,
   - for which purpose and task,
   - whether the data will be anonymized/ pseunonymized or raw, and
   - whether you have an agreement in place with the respective party.

   .....
   .....
   .....
   .....

6. Will you be reusing data for further purposes beyond the project? If yes, please specify what data and for which purposes.

.....
.....
.....

7. Will you re-use any existing datasets? If so, please fill out the table below.

	Please provide your answers in this column:
Dataset(s) name	*What is the name of the used dataset(s)?*
Dataset(s) description	*Please provide a short description of the dataset(s).*
Personal Data	*Does the dataset include personal data? If yes, please specify the type of personal data.*
Purpose	*What is the purpose for which you use/ process the dataset(s)?*
Data format	*What format(s) are your dataset(s)?*
Data Storage	*Where will you store the dataset(s)?*
Main Data Source	*What is the main source of the dataset(s)?*
Data Ownership	*Who owns the dataset(s)?*
Country of Origin	*Where does the dataset come from?*
Restrictions on the use	*Are there any restrictions for the use of the datasets?*
Access	*Who has access to the datasets? Please include other partners and/or work packages which will also access the datasets.*
Retention Period	*How long will you keep the datasets?*
Licence	*Under which licence did you obtain access to the datasets?*
WP and task	*For which work package and which task do you need to use the datasets?*
Additional Comments	*Please add here any additional comments.*

PART B – FAIR DATA

MAKING DATA FINDABLE:

1. Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism? If so, please describe.

……
……
……
……

2. What naming conventions do you follow?

……
……
……
……

3. Will search keywords be provided that optimize possibilities for re-use?

……
……
……
……

4. Do you provide clear version numbers?

……
……
……
……

5. What metadata will be created, if any?

……
……
……
……

MAKING DATA OPENLY ACCESSIBLE:

1. Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.

……
……
……
……

2. How will the data be made accessible?

……
……
……
……

3. What methods or software tools are needed to access the data?

……
……
……
……

4. Is documentation about the software needed to access the data included? If so, please explain.

……
……
……
……

5. Is it possible to include the relevant software? Please elaborate.

……
……
……
……

6. Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.

……
……
……
……

7. Have you explored appropriate arrangements with the identified repository? If so, please explain.

……
……
……
……

8. If there are restrictions on use, how will access be provided?

……
……
……
……

9. Is there a need for a data access committee? Why or why not?

……
……
……
……

10. Are there well described conditions for access (i.e. a machine readable license)? If so, please explain.

……
……
……
……

11. How will the identity of the person accessing the data be ascertained?

……
……
……
……

MAKING DATA INTEROPERABLE:

1. Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?

☐ Yes, namely ……

☐ No

2. What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

……
……
……
……

3. Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability? If so, please explain.

……
……
……
……

4. In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

……
……
……
……

INCREASE DATA RE-USE (THROUGH CLARIFYING LICENSES):

1. How will the data be licensed to permit the widest re-use possible?

……
……
……
……

2. Will the data be made available for re-use? If so, when? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

……
……
……
……

3. Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.

☐ No, because……

☐ Partially, namely ……
☐ Yes.

4. How long is it intended that the data remains re-usable?

……
……
……
……

5. Are data quality assurance processes described?

☐ No
☐ Yes, in ……

PART C – ALLOCATION OF RESOURCES

1. Are there additional costs for making data FAIR in your project?

……
……
……
……

2. How will these be covered?

……
……
……
……

3. Who will be responsible for data management in your project?

……
……
……
……

4. Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?

……
……
……
……

PART D – DATA SECURITY

1. What technical and organisational measures have you put in place to ensure data security? (i.e., anonymisation techniques, pseudonymisation, tokenization, etc)

……
……
……
……

2. Is the data safely stored in certified repositories? Please describe

……
……
……
……

3. How long are you storing the data?

……
……
……
……

PART E – ETHICAL ASPECTS

6. Are there any ethical or legal issues that can have an impact on data sharing? If so, please describe.

……
……
……
……

7. Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data? If so, please describe.

……
……
……
……

PART F – INTELLECTUAL PROPERTY RIGHTS

1. Are you bringing any intellectual property (IP) to the ODIN project and/or licencing any IP to the project partners towards enabling the completion of the project's objectives? If yes, please specify.

☐ No

☐ Yes, namely ……

2. Have you or do you foresee developing, or being involved in the development of new IP in the scope of the project? If yes, please specify.

☐ No

☐ Yes, namely ……

3. In case of mutual development of IP with other project's partners, what would be your requirements in regard to protection of the developed IP?

……
……
……
……

4. If the project produces IP-protection-eligible results, how will you exploit those results?

……
……
……
……

5. Will you enable free access rights (Open Access/ Open Source) after the end of the project? If so, for how long?

……
……
……
……