Project Title:	**Sensing and predictive treatment of frailty and associated co-morbidities using advanced personalized models and advanced interventions**
Contract No:	690140
Instrument:	Collaborative Project
Call identifier:	H2020-PHC-2014-2015
Topic:	PHC-21-2015: Advancing active and healthy ageing with ICT: Early risk detection and intervention
Start of project:	1 January 2016
Duration:	36 months

Deliverable No: D8.12

Data Management Plan (vers a)

Due date of deliverable:	M6 (1st July 2016)
Actual submission date:	30th June 2016
Version:	1.0

Lead Author:	Kosmas Petridis (Hypertech)
Lead partners:	Hariklia Zografou (Hypertech), Vasilis Megalooikonomou, Yannnis Elloul, Christos Makris, Andreas Kanavos, Konstantinos Deltouzos (UoP), Christina Karamanidou (CERTH), Roberto Orselli, Carlo Mancuso (Smartex), Javier Montesa (Brainstorm), Cristiana Degano, Luca Bianconi (Gruppo SIGLA), Ioanna Petridou (Materia Group), Marina Kotsani (INSERM)

Horizon 2020
European Union funding
for Research & Innovation

Change History

Ver.	Date	Status	Author (Beneficiary)	Description
0.3	10/06/2016	draft	Kosmas Petridis, Hariklia Zografou (Hypertech)	First draft datasets definition.
0.5	17/06/2016	draft	Kosmas Petridis, Hariklia Zografou (Hypertech), Vasilis Megalooikonomou, Yannnis Elloul, Christos Makris, Andreas Kanavos, Konstantinos Deltouzos (UoP), Christina Karamanidou (CERTH), Roberto Orselli, Carlo Mancuso (Smartex), Javier Montesa (Brainstorm), Cristiana Degano, Luca Bianconi (Gruppo SIGLA), Ioanna Petridou (Materia Group), Marina Kotsani (INSERM)	First draft deliverable report.
0.8	23/06/2016	draft	Kosmas Petridis, Hariklia Zografou (Hypertech), Vasilis Megalooikonomou, Yannnis Elloul, Christos Makris, Andreas Kanavos, Konstantinos Deltouzos (UoP), Christina Karamanidou (CERTH), Roberto Orselli, Carlo Mancuso (Smartex), Javier Montesa (Brainstorm), Cristiana Degano, Luca Bianconi (Gruppo SIGLA), Ioanna Petridou (Materia Group), Marina Kotsani (INSERM)	Updated deliverable report sent for internal review.
1.0	30/06/2016	final	Kosmas Petridis, Hariklia Zografou (Hypertech)	Deliverable finalised taking into account internal review's comments.

EXECUTIVE SUMMARY

The deliverable “D8.12-Data Management Plan (vers a)” reports the first version of the Data Management Plan (DMP). It summarizes the scientific, technological and personal data expected to be collected or generated during the FrailSafe lifecycle but also the specific measures to be adopted for each distinguished dataset (standards and metadata, sharing, archiving and preservation, ethical issues etc.). Finally, the deliverable illustrates the envisaged strategy to achieve open access to FrailSafe research data and results.

DOCUMENT INFORMATION

Contract Number:	H2020-PHC–690140	**Acronym:**	FRAILSAFE
Full title	Sensing and predictive treatment of frailty and associated co-morbidities using advanced personalized models and advanced interventions		
Project URL	http://frailsafe-project.eu/		
EU Project officer	Mr. Ramón Sanmartín Sola		

Deliverable number:	8.12	**Title:**	Data Management Plan (vers a)
Work package number:	8	**Title:**	Dissemination and Exploitation

Date of delivery	**Contractual**	01/07/2016 (M6)	**Actual**	30/06/2016
Status	Draft ☐		Final ☒	
Nature	Report ☒	Demonstrator ☐	Other ☐	
Dissemination Level	Public ☒	Consortium ☐		
Abstract (for dissemination)	This deliverable reports the first version of the Data Management Plan (DMP). It summarizes the data expected to be collected or generated during the FrailSafe lifecycle but also the specific measures to be adopted for each distinguished dataset. Finally, the deliverable illustrates the envisaged strategy to achieve open access to FrailSafe research data and results.			
Keywords	FrailSafe, Data Management Plan, DMP, open research data, frailty			

Contributing authors (beneficiaries)	Hariklia Zografou (Hypertech) Vasilis Megalooikonomou, Yannnis Elloul, Christos Makris, Andreas Kanavos, Konstantinos Deltouzos (UoP) Christina Karamanidou (CERTH) Roberto Orselli, Carlo Mancuso (Smartex) Javier Montesa (Brainstorm) Cristiana Degano, Luca Bianconi (Gruppo SIGLA) Ioanna Petridou (Materia Group) Marina Kotsani (INSERM)			
Responsible author(s)	Kosmas Petridis		**Email**	k.petridis@hypertech.gr
	Beneficiary	Hypertech	**Phone**	+30 2310 262201

TABLE OF CONTENTS

TABLE OF FIGURES

1 Introduction

The deliverable “D8.12-Data Management Plan (vers a)” reports the first version of the **Data Management Plan (DMP)** of the FrailSafe project.

In Horizon 2020, the emphasis on **data management and open research data** has been increased compared to earlier framework programmes. In this respect, the deliverable summarizes the scientific, technological and personal data expected to be collected or generated during the FrailSafe lifecycle but also the specific measures to be adopted for each distinguished dataset (standards and metadata, sharing, archiving and preservation, ethical issues etc.), following the respective template provided by the European Commission. Finally, the deliverable illustrates the envisaged strategy to achieve open access to FrailSafe research data and results.

The FrailSafe DMP is a living document and is expected to be **updated as the project evolves**, in order to reflect the **progress performed** and the **results generated** throughout the project execution. Clearly, as the present deliverable was prepared during the first stages of FrailSafe (M6), it can only reflect the intentions of the consortium towards developing the overall DMP strategy. Therefore, an updated version of the DMP will be available in M24, while the final version will be delivered in M36, documenting the detailed descriptions of all datasets collected, also clearly depicting which of them will be made publicly available and under which management framework.

2 General Principles

The establishment of the FrailSafe DMP underlines an appreciation of the project's responsibility to manage relevant data in an appropriate manner. All FrailSafe partners have to collect, store and manage data in line with local laws and to treat data in line with the guidelines of this document.

Several principles have to be used while dealing with research data. Amongst these are:

- Data protection and privacy has to be respected and appropriate solutions for data storage and handling must be established.
- Open access to data should be the main principle for public-funded projects.
- Data should be discoverable, accessible and interoperable to specific quality standards.
- Integrity of the research depends on the quality of data.

The following subsections describe the measures adopted by the FrailSafe consortium to achieve compliance with these principles.

2.1 EU Commission Guidelines

The FrailSafe DMP is based on the guidelines of the EU Commission regarding the openness of the data generated from projects funded by the H2020 framework [1]. According to these guidelines the scientifically-oriented data, which is going to be generated by the FrailSafe project will be formed so that it could be easily **discoverable, accessible, assessable and intelligible, usable** beyond the original purpose of its collection and usage but also **interoperable** to appropriate quality standards.

2.2 Security

The datasets foreseen to be collected through FrailSafe are of high value and may contain sensitive data. Special care should be taken to prevent any unauthorized data access. This is another key aspect of the FrailSafe DMP and all data repositories used by the project will include effective protection.

A holistic security approach will be followed, in order to protect the pillars of information security (confidentiality, integrity, availability). The security approach will consist of a methodical assessment of security risks followed by their impact analysis. This analysis will be performed on the personal information and data processed by the FrailSafe system, their flows and any risk associated to their processing.

Moreover, the pilot sites will apply monitored and controlled procedures related to the data collection, its integrity and protection. The data protection and privacy of personal information will include protective measures against infiltration as well as physical protection of core parts of the systems and access control measures.

2.3 Personal Data Protection

FrailSafe activities will involve human participants to monitor human activity and behaviour analysis. Therefore, it is clear that in some cases personal data will have to be collected. Such data will be protected in accordance with the EU's Data Protection Directive 95/46/EC [2] on the protection of individuals with regard to the processing of personal data. In this direction, data will also be carefully anonymized in order to be furtherly used inside the project or made available publicly.

Further information on how personal data collection and handling should be approached inside FrailSafe, is provided in the deliverable “D9.9: Ethics, Safety and mHealth Barriers”.

3 Foreseen FrailSafe Datasets

In order to meet the requirements of the DMP according to the Pilot of Open Access of the Horizon 2020 programme, a "dataset identification" template was first drafted. This was circulated to all project partners so as to collect all relevant information concerning the datasets, which are planned to be developed in the course of the FrailSafe. On the basis of all partners' feedback, the preliminary FrailSafe DMP has been established, as described in the next subsections.

3.1 WWBS Metrics

WWBS metrics	
Data identification	
Naming prefix	FS01.WWBSMetrics
Data set description	The data will be collected from subjects enrolled using the sensorized vest. Three (3) different kind of sensors will be used: • **ECG** electrodes placed on the middle of the chest to monitor heart rate and ECG similar to DI. • **Breathe** sensors placed on the middle of the chest to monitor breathe rate. • **IMUs** (Inertial Measurement Unit) placed on the arms to monitor activity, posture, instability/falls, step counter. The raw data are the signals from the sensors. The electronics acquisition device saves and sends the data wirelessly. These signals will be used to daily monitor the subject.
Project activities	
Partners involved	Smartex, CERTH, UoP, Gruppo SIGLA
WPs and Tasks	WP3
Standards and metadata	
Metadata and documentation	The dataset will be accompanied with detailed documentation of its record fields and metadata.
Standards and format	The raw data saved in the electronics device will be in a proprietary format (.wwsx).
Data sharing	
Access policy / Dissemination level	Sensitive personal data will be handled only by the local clinical research personnel bound by local confidentiality rules. This data will not be transferred, merged or exchanged. All other dataset's collected data will be anonymized and will contain no identifying information. Data will then pass to be used by the members of the FrailSafe consortium. A carefully selected subset of the data is

	expected to be provided publicly for research and educational purposes. The IPR and Joint Ownership and Exploitation agreements that will be developed later in the project will be taken into account in this respect.
Embargo periods	None
Archiving and preservation	
Data storage: Where? For how long?	The raw data will be saved in a microSD card inside the electronics device. The capacity of the microSD will be enough to store at least one week of acquisition. Part of this data will be regularly transferred wirelessly to the mobile device, present at the place of metrics recording. All data, saved and streamed, will be then uploaded to the FrailSafe cloud infrastructure. Anonymized data stored in the cloud will be encrypted. The dataset will be preserved at least until the end of the project duration.
Ethics	
Ethical issues	The dataset will be obtained in accordance to the local ethics requirements. Any personal information regarding the participants will be treated as sensitive personal data (as defined in deliverable D9.9) and kept strictly private. Recorded data will be anonymized with no personal identifiers and no means to link these to personal identifiers – hence falling outside the scope of legislation concerning personal data.

3.2 Device Metrics

Device metrics	
Data identification	
Naming prefix	FS02.DeviceMetrics
Data set description	The data will be collected from subjects enrolled in the trial process. Four (4) different devices will be used: • **Dynamometer,** to be used for strength evaluation. • **Mobilograph,** for arterial stiffness evaluation. • **BP monitor,** to be used for blood pressure recording. • **Scale,** to evaluate weight and bioelectrical impedance.
Project activities	
Partners involved	Smartex, CERTH, UoP, Materia,

	INSERM
WPs and Tasks	WP2, WP3
Standards and metadata	
Metadata and documentation	The dataset will be accompanied with detailed documentation of its record fields and metadata.
Standards and format	Formats vary according to the device and the measurement. For instance, the dynamometer and mobilograph will communicate with the portable devices using special software and for the mobilograph a whole set of results will be transferred centrally.
Data sharing	
Access policy / Dissemination level	Sensitive personal data will be handled only by the local clinical research personnel bound by local confidentiality rules. This data will not be transferred, merged or exchanged. All other collected data will be anonymized and will contain no identifying information. Data will then pass to be used by the members of the FrailSafe consortium. A carefully selected subset of the data is expected to be provided publicly for research and educational purposes. The IPR and Joint Ownership and Exploitation agreements that will be developed later in the project will be taken into account in this respect.
Embargo periods	None
Archiving and preservation	
Data storage: Where? For how long?	Depending on the device, the collected data will be transferred either wirelessly or by hand to the mobile device, present at the place of metrics recording and will be then uploaded to the FrailSafe cloud infrastructure. Anonymized data stored in the cloud will be encrypted. The dataset will be preserved at least until the end of the project duration.
Ethics	
Ethical issues	The dataset will be obtained in accordance to the local ethics requirements. Any personal information regarding the participants will be treated as sensitive personal data (as defined in deliverable D9.9) and kept strictly private. Recorded data will be anonymized with no personal identifiers and no means to link these to personal identifiers – hence falling outside the scope of legislation concerning personal data.

3.3 Location Detection

Location detection	
Data identification	
Naming prefix	FS03.LocationDetection
Data set description	**Indoor location dataset** This dataset will be collected from the indoor localization system, which utilizes Bluetooth beacons to track the older person movements. Each record in the dataset will contain the following fields: • **User ID:** A unique ID characterizing the older person participating in the monitoring process. It must be noted that the dataset will pass through an anonymization process where the actual participants' names will be removed and the dataset instances will be referenced as numeric IDs. • **Facility ID**: A unique ID characterizing the facility in which the monitoring process takes place, e.g. a specific home or indoor area. This ID is linked with further information regarding the facility, such as a top view of its rooms. This can be useful for analyzing the recorded coordinates of the person movements inside the actual context of the facility. • **Timestamp:** The date and time of the recording. Recordings will take place regularly, at small time intervals. The duration of the time intervals depends on the level of granularity required for the data analysis. It could be e.g. 10 seconds or some minutes. • **Room ID**: A unique ID characterizing the room in which the person is at the time of the recording. The room ID is related to the specific facility that the monitoring takes place. **Outdoor location dataset** This dataset will be collected from the outdoor location system, which utilizes GPS tracking methods for monitoring the user movements. Each record in the dataset will contain the following fields: • **User ID**: A unique ID characterizing the older person participating in the monitoring process. Again, the

	dataset will pass through an anonymization process where the actual participants' names will be removed and the dataset instances will be referenced as numeric IDs. • **Timestamp:** The date and time of the recording. Recordings will take place regularly, at small time intervals. The duration of the time intervals depends on the level of granularity required for the data analysis. It could be e.g. 10 seconds or some minutes. • **Coordinates:** The geographical coordinates, i.e. latitude and longitude, of the person at the time of the recording.
Project activities	
Partners involved	CERTH
WPs and Tasks	WP3
Standards and metadata	
Metadata and documentation	Metadata include any information regarding the facilities and rooms for the indoor location dataset. E.g. facility-related metadata are the following (not limited to these, depending on the available information): • **Facility ID:** A unique ID of a facility. This ID will be linked to the same field in the indoor location dataset records. • **Facility type:** The type of the facility, e.g. home, hospital, etc. Similarly, room-related metadata include the following: • **Room ID:** A unique ID, to be linked to the same field in the indoor location dataset records. • **Room type**: The type of the room, e.g. living-room, bathroom, etc. Moreover, the indoor and outdoor datasets will be accompanied with detailed documentation of their record fields and their metadata.
Standards and format	The raw data will be recorded by the mobile device in CSV format, where each line represents a record and each comma-separated column is a different field of the record. Once the data is transferred to the FrailSafe cloud server, it will be stored in appropriate databases.
Data sharing	
Access policy / Dissemination level	Sensitive personal data will be handled only by the local clinical research personnel bound by local confidentiality rules. This data will not be transferred,

	merged or exchanged. All other collected data will be anonymized and will contain no identifying information. Data will then pass to be used by the members of the FrailSafe consortium. A carefully selected subset of the data is expected to be provided publicly for research and educational purposes. The IPR and Joint Ownership and Exploitation agreements that will be developed later in the project will be taken into account in this respect.
Embargo periods	None
Archiving and preservation	
Data storage: Where? For how long?	The raw data will be collected in a microSD card inside the mobile device, i.e. the smartphone, both for the indoor and the outdoor location datasets. The capacity of the microSD card is enough to store one week of acquisition. After the one-week monitoring period of one user, the data will be transferred to the FrailSafe cloud and stored in the FrailSafe database. Anonymized data stored in the cloud will be encrypted. The data will persist in the FrailSafe database at least throughout the duration of the project.
Ethics	
Ethical issues	The dataset will be obtained in accordance to the local ethics requirements. Any personal information regarding the participants will be treated as sensitive personal data (as defined in deliverable D9.9) and kept strictly private. Recorded data will be anonymized while the data collection will comply with the geolocation data protection guidelines reported in deliverable D9.9 (Section 6.5).

3.4 Games

Games	
Data identification	
Naming prefix	FS04.Games
Data set description	Games will contain various datasets, as described in the following paragraphs. **Session dataset** This dataset contains the time programming data and summary of the

	game projected or completed: • **Session ID:** A unique ID characterizing the game session recorded. • **Game ID:** The ID of the game that has been played on the session. • **User ID:** A unique ID characterizing the user playing the game (anonymization process will be applied). • **Expected timestamp:** A timestamp indicating the date when the user is expected to make the session. • **Execution timestamp:** The date and time when the user started the session. • **Completed:** A boolean indicating if the user has completed the session or not. • **Score:** The average score of the session. • **Results ID:** An id referring to the results dataset. **Game dataset** • **Game ID:** A unique ID identifying the game. • **Title:** The name of the game. • **Description:** A short description to the users about the exercises programmed on the game. • **Timestamp:** The date and time when the game was created. • **Doctor ID:** A reference to the user (Doctor), who has customized the game. **Script dataset** This dataset contains the different movements than the players must perform in the scripts, including the timing and the order of the event: • **Pose ID:** An ID referring to a pose. • **Game ID:** An ID referring to the game that the position belongs to. • **Order:** An integer meaning the position of that movement inside the full script. • **GoTime:** A float parameter that defines the available seconds for the player to reach the next target position. • **DoTime:** A float parameter that defines the time to stay in the

	position. **Posture dataset** This dataset will contain all the data involved in the configuration of a pose: • **Pose ID:** A unique ID characterizing the pose. • **Pose description:** A string describing the pose. • **Rate:** The rate of the pose, an integer defined by the creator of the pose. • **Joint data:** A set of values describing the joint position. Head, pitch and roll. And a second set of values describing the tolerance of these values. This second set of values make the position more or less rigid, more or less easy to reach. • **Joint weights:** A set of integers, one per joint, expressing the value of each one for the average pose validation. **Results dataset** Dataset that stores the paths to the files, which contain the raw data of the completed games: • **Session ID:** An ID identifying the session that data is referencing. • **Game ID:** An ID identifying the game that data is referencing. • **Script ID:** An ID identifying the script that data is referencing. • **Results File:** A reference to the file that contains the game results of a played script. • **Raw File:** A reference to the file that contains the whole data that sensors have recorded in a played script. **Cognitive games dataset** This dataset will contain the data collected through the use of the cognitive assessment games. For each session played by the user, the following data will be recorded: ○ **User ID:** A unique ID characterizing the user playing the game (anonymization process will be applied). ○ **Session ID**: A unique ID characterizing the game session recorded.

	○ **Timestamp:** The date and time when the user played the specific game session. ○ **Game statistics**: A set of statistic information regarding the performance of the user in this game session. Such information include: ○ The time needed for completing the whole game. ○ The time needed for completing specific tasks of the game. ○ The steps taken by the user in order to fulfil the game objectives. ○ The percentage of objectives successfully met by the player during the game session, over all required objectives. This information will be more detailed and specific, according to the specific characteristics and gameplay of the cognitive game. ○ **Cognitive state:** The cognitive state of the user at the time of the game session, automatically evaluated after the completion of the game.
Project activities	
Partners involved	BRAINSTORM, CERTH
WPs and Tasks	WP5
Standards and metadata	
Metadata and documentation	The games datasets will be accompanied by detailed documentation of the records and their fields.
Standards and format	Regarding the cognitive games dataset, the data collected during the playing of the game are stored in XML format. An XML file containing the data is initially stored locally and then is serialized and sent to the FrailSafe server, which deserializes it and stores the data in a MySQL database. Regarding the rest of the games, the format of the data stored on the mobile devices will be Ascii files. Each data set will have a different data structure. The poses and user data will be XML data files and the logs ofthe different games will be stored in CSV format.
Data sharing	
Access policy / Dissemination level	Sensitive personal data will be handled only by the local clinical research personnel bound by local confidentiality rules. This data will not be transferred, merged or exchanged. All other collected data will be

	anonymized and will contain no identifying information. Data will then pass to be used by the members of the FrailSafe consortium. A carefully selected subset of the data is expected to be provided publicly for research and educational purposes. The IPR and Joint Ownership and Exploitation agreements that will be developed later in the project will be taken into account in this respect.
Embargo periods	None
Archiving and preservation	
Data storage: Where? For how long?	The game data will initially be collected locally in a microSD card inside the tablet. The capacity of the microSD card is enough to store the data collected during the one-week monitoring period for one patient. After the user's monitoring period, the data will be transferred to the FrailSafe cloud and stored in the FrailSafe database. Anonymized data stored in the cloud will be encrypted. The data will persist in the FrailSafe database at least throughout the duration of the project
Ethics	
Ethical issues	The dataset will be obtained in accordance to the local ethics requirements. Any personal information regarding the participants will be treated as sensitive personal data (as defined in deliverable D9.9) and kept strictly private. Recorded data will be anonymized with no personal identifiers and no means to link these to personal identifiers – hence falling outside the scope of legislation concerning personal data.

3.5 Clinical Assessments

Clinical assessments	
Data identification	
Naming prefix	FS05.ClinicalAssessments
Data set description	The clinical assessment includes: i) Face to face clinical assessment; ii) Self-administrated questionnaires / text sampling; iii) Phone-follow up assessment, as it will be described in the next paragraphs. Participants will form four groups: i) start up group; ii) main group; iii) short and

	long term evaluation group; iv) control group. **Face to face clinical assessment** • **Sensitive personal data:** name, date of birth, address, telephone numbers, e-mail. • **Demographic/General data:** Date of entry in the study, initials, year of birth, gender, education, profession, living condition, family status, leisure activities, physical activity use of alcohol and smoking habits. • **Past and present medical history:** Comorbidities, effects of comorbidities on ADL, medications, previous hospitalizations, previous fractures (number and anatomic location), present irregular pulse, sleep pattern. • **Clinical measurements**: Height/weight measurement, BMI, body fat/lean mass, waist/chest circumference, blood pressure measurements (sitting and standing), arterial rigidity evaluation. • **Conventional evaluation of frailty**: Fried's criteria for frailty. • **Balance and gait evaluation**: Lower limb strength: raise from the chair 5 times without helping by the arms; balance: single foot standing, gait speed: “Time up and go test”, speed for 4 meters straight walking. • **Sensory evaluation:** Vision, audition. • **Nutritional assessment:** MNA scale (short version); if score <11 then MNA extended version. • **Activities of daily living:** Katz index, Lawton IADL scale. • **Cognitive-emotional evaluation:** scales: MMSE, MoCa, Geriatric depression Scale (15-GDS), anxiety visual analogue scale • **Quality of life, Self-Health rating, Pain evaluation**: Visual analogue scale. **Self-administrated questionnaires / text sampling,** which are discussed in detail in the next table (see Section 3.6)**:** • **Social interaction**: 34-item self-

	administered questionnaire with both open and close questions. • **Data collection of written texts**: Collection of previous typed or handwritten text, type or handwrite or dictate a major life event, type or handwrite or dictate the description of a standard picture. • **Big five personality trait:** Big 5 test (self evaluating test). **Phone-follow up assessment** • **Record of events**: Number of falls and fractures (including anatomic location), date and cause of hospitalizations and/ or death.
Project activities	
Partners involved	UoP, INSERM, Materia, Gruppo SIGLA
WPs and Tasks	WP2, WP4, WP6
Standards and metadata	
Metadata and documentation	The clinical assessment dataset will be accompanied by detailed documentation of the records and their fields, as well as a description of the medical assessment statistics used to perform the clinical research.
Standards and format	The clinical web platform of FrailSafe will be used. Once the data is transferred to the FrailSafe cloud facilities, it will be stored in appropriate databases.
Data sharing	
Access policy / Dissemination level	Sensitive personal data will be handled only by the local clinical research personnel bound by local confidentiality rules. This data will not be transferred, merged or exchanged. All other dataset's collected data will be anonymized and will contain no identifying information. Data will then pass to be used by the members of the FrailSafe consortium. A carefully selected subset of the data is expected to be provided publicly for research and educational purposes. The IPR and Joint Ownership and Exploitation agreements that will be developed later in the project will be taken into account in this respect.
Embargo periods	None
Archiving and preservation	
Data storage: Where? For how long?	The dataset will be securely stored -via the clinical web platform- in the FrailSafe cloud facilities into appropriate databases. Anonymized data stored in

	the cloud will be encrypted. The data will persist in the FrailSafe database at least throughout the duration of the project.
Ethics	
Ethical issues	The data collection and persistence will comply with the data protection guidelines reported in D9.9 (Section 6) with the aim of, at same time, keeping the maximum level of security and privacy of the data and allowing the successful performance of the clinical research and of the other tasks of the project.

3.6 Self-administered Questionnaires for Social Media Sensing

Self-administered questionnaires for social media sensing	
Data identification	
Naming prefix	FS06.SocialMediaSensing
Data set description	The main idea is to measure social interaction of ageing people as well as social and behavioral parameters that emotionally characterize their scripts. The dataset will be collected with use of questionnaires. The plan is to collect data and to investigate techniques that connect text features and multiple values of the Big Five personality traits with symptoms of frailty. The training classification phase aims at predicting / characterizing frailty based on the written scripts. The self-filled questionnaires are: 1. **Social interaction** (measured by the number of incoming/outgoing phone calls/sms, emails, use of social networks etc) and other social and behavioural parameters (through linguistic analysis of text appearing in chat sessions or other write-ups, monitoring their location throughout the day) while respecting privacy and without becoming invasive. The outcome of the analysis of this plethora of data will be a formal and quantitative definition of a frailty metric that will be based on the aforementioned sensing dimensions. This questionnaire is administrated once, during or after the first clinical assessment in order to investigate the social interaction behavior of

	each participant. 2. **Data collection of written texts** will be given to the participants to fill in a second time, except if the participant is unable to write. More specifically, participants are asked for previous text, are asked to think of a major life event (prompts for life events are available such as weddings, child's birth, professional achievements etc), are asked to describe in written text an attached picture. The timing of collection of the written texts coincides with the clinical assessments. The natural language analysis tool, which will be developed in terms of T4.5 will be able to detect signs of cognitive deficiencies in written text. 3. **Big five personality trait:** There are works that connect mental disorders with the Big Five personality traits and works that try to exploit this information employing social media. We will extend existing approaches in predicting personality traits by sketching the user profile. According to it, the human personality is described as a vector of 5 values of traits: i. **Openness:** This trait features characteristics such as curious, original, intellectual, creative and open to new ideas; ii. **Conscientiousness:** Common features of this dimension include organized, systematic, punctual, achievement, oriented and dependable; iii. **Extraversion:** This trait includes characteristics such as outgoing talkative, sociable and enjoying social situations; iv. **Agreeableness:** This personality dimension includes attributes such as affable, tolerant, sensitive, trusting, kind and warm; v. **Neuroticism:** Individuals high in this trait tend to experience anxious, temperamental and moody. This personality characterization will be automatically extracted and taken into account in the frailty metric. In addition it will help in improving the

	services proposed by our approach to the elder according to his personality and his current emotional state.
Project activities	
Partners involved	UoP, INSERM, Materia, Gruppo SIGLA
WPs and Tasks	WP2, WP4, WP6
Standards and metadata	
Metadata and documentation	The social media sensing and processing datasets will be accompanied by detailed documentation of the records and their fields, as well as a description.
Standards and format	The questionnaires web platform -that is part of the clinical web platform- will be used. Once the data is transferred to the FrailSafe cloud facilities, it will be stored in appropriate databases.
Data sharing	
Access policy / Dissemination level	Sensitive personal data will be handled only by the local clinical research personnel bound by local confidentiality rules. This data will not be transferred, merged or exchanged. All other dataset's collected data will be anonymized and will contain no identifying information. Data will then pass to be used by the members of the FrailSafe consortium. A carefully selected subset of the data is expected to be provided publicly for research and educational purposes. The IPR and Joint Ownership and Exploitation agreements that will be developed later in the project will be taken into account in this respect.
Embargo periods	None
Archiving and preservation	
Data storage: Where? For how long?	The answers to the questionnaires will be securely stored -via the clinical web platform- into the FrailSafe cloud facilities. These answers will be stored into appropriate databases, which could be securely accessed by the Analysis modules for producing new reasoned data, as foreseen by the project design. Anonymized data stored in the cloud will be encrypted. The data will persist in the FrailSafe database at least throughout the duration of the project.
Ethics	
Ethical issues	The data collection and persistence will comply with the data protection guidelines reported in D9.9 (Section 6) with the aim of, at same time, keeping

	the maximum level of security and privacy of the data and allowing the successful performance of the clinical research and of the other tasks of the project.

3.7 Data Analysis

Data analysis	
Data identification	
Naming prefix	FS07.DataAnalysis
Data set description	During the (offline and online) analysis, older people data collected will be analyzed and knowledge will be extracted (analysis data). This analysis data will assist the medical professionals on their diagnosis and will provide them with advanced decision making capabilities. Additionally this kind of data will lead to a better understanding of frailty and to the extraction of new frailty metrics.
Project activities	
Partners involved	UoP
WPs and Tasks	WP4
Standards and metadata	
Metadata and documentation	The analysis data/results derived from offline as well as online data analysis will be accompanied by detailed documentation of the records and their fields, as well as a description.
Standards and format	The analysis data/results will be transferred to the FrailSafe cloud facilities, where it will be stored in appropriate databases.
Data sharing	
Access policy / Dissemination level	Sensitive personal data will be handled only by the local clinical research personnel bound by local confidentiality rules. This data will not be transferred, merged or exchanged. All other dataset's collected data will be anonymized and will contain no identifying information. Data will then pass to be used by the members of the FrailSafe consortium. A carefully selected subset of the data is expected to be provided publicly for research and educational purposes. The IPR and Joint Ownership and Exploitation agreements that will be developed later in the project will be taken into account in this respect.

Embargo periods	None
Archiving and preservation	
Data storage: Where? For how long?	The dataset will be securely stored in the FrailSafe cloud facilities into appropriate databases. Specifically, the offline data analysis results will be stored directly in the FrailSafe cloud. The online data analysis results will be stored locally in the older people's devices (e.g. smartphone, mini-pc) and will be transferred periodically to the FrailSafe cloud. Anonymized data stored in the cloud will be encrypted. The data will persist in the FrailSafe database at least throughout the duration of the project.
Ethics	
Ethical issues	The data collection and persistence will comply with the data protection guidelines reported in D9.9 (Section 6) with the aim of, at same time, keeping the maximum level of security and privacy of the data and allowing the successful performance of the clinical research and of the other tasks of the project.

3.8 Virtual Patient Models

Virtual patient models	
Data identification	
Naming prefix	FS08.VirtualPatientModels
Data set description	This dataset will cover the Virtual Patient Models produced in the project's framework resulting from processing of the original data/measurements. These profiles and models of frailty patients will be created within the FrailSafe project as they are described in "T4.3: Dynamic user profiling models", based on the data analyses of T4.1.
Project activities	
Partners involved	CERTH, UoP, Hypertech
WPs and Tasks	WP4
Standards and metadata	
Metadata and documentation	The dataset will be accompanied with detailed documentation of its contents and of all the variables involved in the FrailSafe Virtual Patient Models.
Standards and format	The dataset would be in the form of XML-based representations of the parameters involved in the FrailSafe Virtual Patient

	Models. Guidelines for Virtual Human Modelling derived from the VERITAS project and the VUMS cluster [3] will be used, along with related XSD and XML specifications. The adoption and extension of the existing representation format (OWL or UsiXML) developed in the context of the VERITAS project will also be investigated. Furthermore, the clinical component of the models could be based on the format of electronic health records such as the openEHR framework [4].
Data sharing	
Access policy / Dissemination level	Sensitive personal data will be handled only by the local clinical research personnel bound by local confidentiality rules. This data will not be transferred, merged or exchanged. All other dataset's collected data will be anonymized and will contain no identifying information. Data will then pass to be used by the members of the FrailSafe consortium. A carefully selected subset of the data is expected to be provided publicly for research and educational purposes. The IPR and Joint Ownership and Exploitation agreements that will be developed later in the project will be taken into account in this respect.
Embargo periods	None
Archiving and preservation	
Data storage: Where? For how long?	The dataset will be securely stored in the FrailSafe cloud facilities. Anonymized data stored in the cloud will be encrypted. The data will persist in the FrailSafe database at least throughout the duration of the project.
Ethics	
Ethical issues	The data collection and persistence will comply with the data protection guidelines reported in D9.9 (Section 6) with the aim of, at same time, keeping the maximum level of security and privacy of the data and allowing the successful performance of the clinical research and of the other tasks of the project.

3.9 Educational and Training Material

Educational and training material	
Data identification	
Naming prefix	FS09.EducationalTrainingMaterial
Data set description	A dataset of educational and training content will be generated during the FrailSafe project lifecycle to support older people and clinicians. Specifically, the dataset will include educational material related to ageing diseases and frailty as well as material related to the training of trial participants. This content will also be made available to attract interest to specific target groups and achieve increased and active involvement to the FrailSafe activities. For instance, the dataset would be useful, among others, in the training of graduate and undergraduate students of computer engineering and informatics, neuroinformatics, and medical informatics, as well as students and training physicians in medicine (neurology, gerontology, etc) to better understand frailty and related comorbidities.
Project activities	
Partners involved	All
WPs and Tasks	WP7, WP8
Standards and metadata	
Metadata and documentation	The dataset will be accompanied with detailed documentation of its contents.
Standards and format	Existing common formats for documents, videos, images and presentations will be utilized (e.g. pdf, doc, png).
Data sharing	
Access policy / Dissemination level	The dataset is expected to be widely open and publicly available. The IPR and Joint Ownership and Exploitation agreements that will be developed later in the project will be taken into account in this respect.
Embargo periods	None
Archiving and preservation	
Data storage: Where? For how long?	The dataset will be stored in the FrailSafe open repository and will be preserved at least until the end of the project duration.
Ethics	
Ethical issues	The dataset will include no personal information regarding the participants. This information -as stated in deliverable

	D9.9- will be treated as sensitive personal data and kept strictly private. Presented data (recorded, processed etc.) will be anonymized with no personal identifiers and no means to link these to personal identifiers – hence falling outside the scope of legislation concerning personal data.

4 Open Repository

FrailSafe will support an **open repository** to publish the foreseen public part of each dataset, as described in Section 3 above. For this purpose, the **FrailSafe wiki/repository** will be used, which is already available as part of the FrailSafe website and was presented in detail in deliverable “D8.5: Project Web Presence”.

This repository currently supports a **private/logged area**, restricted to consortium members only, enabling file sharing and exchange among the project partners. It allows users to navigate through the available folders and files, add new files and its metadata or download the existing ones (Figure 1).

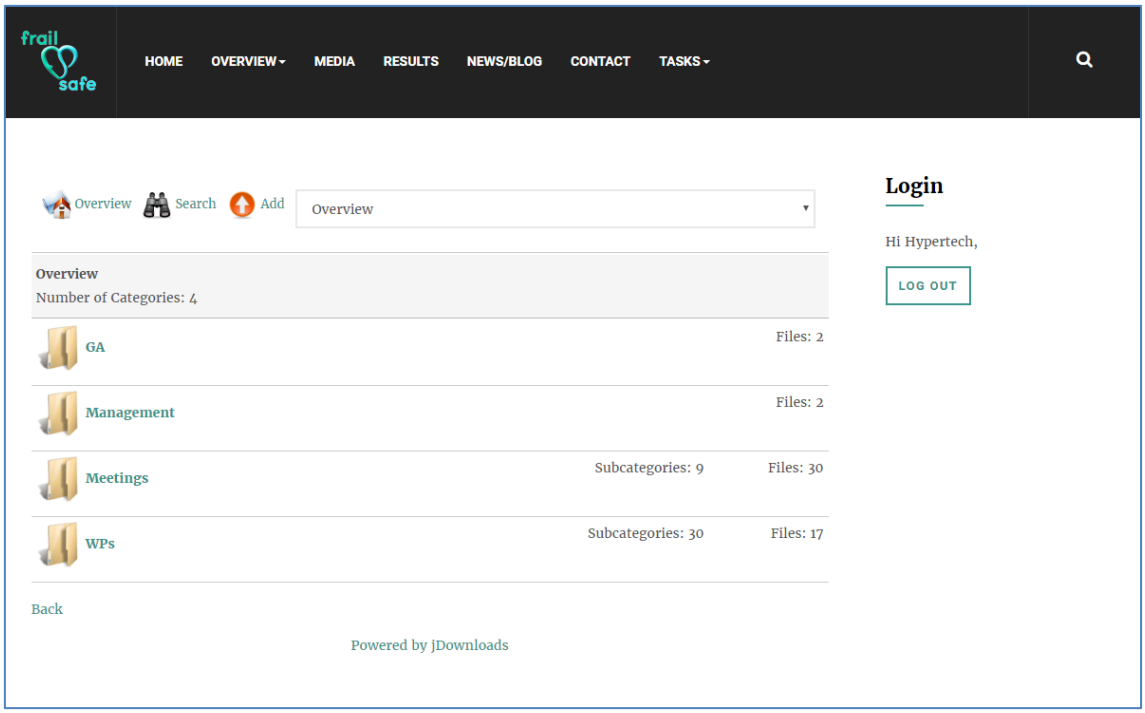

Figure 1 FrailSafe repository

As already mentioned, for the time being, all folders and files are kept private. However, the repository also provides the ability to publish selected folders or files to the public. This publc part will actually serve as the **open repository of FrailSafe,** as soon as the first datasets are made available.

The **public datasets** of FrailSafe will be published through the respository. They will aggregate suitable **metadata descriptions** and provide links to respective download sections to the interested public, as well as centralized data management functionalities to project partners (see deliverable D8.5 for more).

References

[1] European Commission, “Guidelines on Data Management in Horizon 2020”, version 2.1: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf. Last updated: 15 February 2016.

[2] European Parliament and Council, Directive 95/46/EC “on the protection of individuals with regard to the processing of personal data and on the free movement of such data”: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31995L0046&from=en

[3] VERITAS project and VUMS cluster: http://veritas-project.eu/2010/06/vums-cluster-website-online/. Assessed 2015.

[4] OpenEHR standard: http://www.openehr.org/