STREAMLINED **GER**IATRIC AND **ON**COLOGICAL EVALUATION BASED ON IC **TE**CHNOLOGYFOR HOLISTIC PATIENT-ORIENTED HEALTHCARE MANAGEMENT FOR OLDER MULTIMORBID PATIENTS

HORIZON 2020 PROGRAMME – TOPIC H2020-SC1-BHC-24-2020
Start date : 01/04/2021 - Duration : 60 months

GERONTE DATA MANAGEMENT PLAN

Lead Beneficiary : 3-DIAK

Involved Beneficiaries : All

Primary author(s)	Cecile Bacles | UBx; Paul Davis | DCU ; Marije Hamaker | DIAK
Deliverable Type	Report
Dissemination Level	Public
Due Date	30/09/2021 (MONTH 6)
Pages	26
Document version	V1.2
Project Acronym	GERONTE
Project Title	Streamlined **Ger**iatric and **On**cological evaluation based on IC **Te**chnology for holistic patient-oriented healthcare management for older multimorbid patients
Grant Agreement Number	945218
Project Coordinator	Université de Bordeaux Prof. Pierre SOUBEYRAN

GERONTE Consortium

Number	Participant Name	Short Name	Country
1	Université de Bordeaux	UBx	FR
	LTP1 : CHU Bordeaux	CHU Bx	
	LTP2 : Institut Bergonié	IB	
2	Katholieke Universitat Leuven	KUL	BE
3	Stichting Diakonessenhuis	DIAK	NL
4	Oslo University Hospital	OUS	NO
5	University College Dublin	UCD	IE
6	E-Seniors	ESE	FR
7	MyPL SAS	MyPL	FR
8	Università Luigi Bocconi	BOC	IT
9	International Society of Geriatric Oncology	SIOG	CH
10	Dublin City University	DCU	IE

Disclaimer and acknowledgement

Published with a Creative Commons Attribution 4.0 International Licence, (CC BY 4.0).
This report has been carried out under grant agreement No. 945218 from the European Union's Horizon 2020 research and innovation programme. The sole responsibility for the content of this report lies with the GERONTE consortium. It does not necessarily reflect the opinion of the European Union. The European Commission is not responsible for any use that may be made of the information contained therein.

Contributors

Contributor	Participant
Cecile Bacles	UBx – participant 1
Paul Davis	DCU - participant 10
Marije Hamaker	DIAK – participant 3

Formal Reviewers

Contributor	Participant
Pierre Soubeyran	UBx - participant 1
Consortium Board	All

History of Changes

Version	Date	Author	Description of change
V1.0	10/08/2021	Cecile Bacles [UBx]	First draft
V1.1	16/08/2021	Marije Hamaker [DIAK]	Review from Data Manager
V1.1	10/09/2021	Paul Davis [DCU]	Review from Exploitation Manager
V1.2	15/09/2021	Cecile Bacles [UBx]	Implementation of comments from all participants for approval by Consortium Board

Table of Content

Executive Summary

The present document corresponds to the deliverable D8.2 “Data Management Plan” (DMP) of the GERONTE project, which has received funding from the European Union’s Horizon 2020 Programme under Grant Agreement N° 945218. It contains information about the main elements of the Data Management policy that will be used by the Consortium with respect to the project research data. It will include information about the production and management of the research data throughout the project life and the data curation after the projects end.

As GERONTE participates in the Horizon 2020 Open Research Data Pilot, it is intended to identify every opportunity for generating Findable Accessible Interoperable and Re-usable (FAIR) data to maximise GERONTE’s impact. This document has been prepared following the format of the EC H2020 DMP template. Its version is annotated for the use of participants under Societal Challenge 1 (V2.0 of 15/02/2018) and will include analysis of the most relevant aspects of the data management policy.

The document is divided into seven sections, with each section corresponding to a key area of the DMP scheme: I) a general introduction describing the framework of the DMP; II) Data Summary; III) FAIR data; IV) Allocation of resources; V) Data security; VI) Ethical aspects; and VII) Other issues.

This is an initial version of the DMP produced within the first six months of the project as a preliminary discursive document for data management in GERONTE approved by GERONTE consortium board on 20/09/2021. It is intended to be a living document in which information can be made available through updates as the implementation of the project progresses and datasets are produced. The research data types referred to in this document are to be drafted during the first stage of the project (as of 15/09/2021). This document can only reflect the intentions of the project partners toward developing the overall project’s datasets. Updates will be prepared for submission at least at the end of each periodic and the final reporting period (M18, M36, M48, M60).

Attainment of the objectives and explanation of deviations

The objectives related to this deliverable have been achieved in full and as described in Annexe 1 (Description of the Action Part A) of the Grant Agreement N°945218

Justification for delay in deliverable submission

The objectives related to this deliverable have been achieved on time and as scheduled in Annexe 1 (Description of the Action Part A) of the Grant Agreement N°945218

Glossary

DMP	Data Management Plan
DOI	Digital Object Identifier
CRA	Clinical Research Associate
e-CRF	Electronic-Case Report Form
Diamond open access	A peer-reviewed research output, generally a journal-article, is immediately provided in open access mode as published, in a publication process that is both free of charge to the author and to the reader
Gold open access	A peer-reviewed research output, generally a journal-article, is immediately provided in open access mode as published. Some journals require a fee as a one-off payment by authors for this option.
Green open access	Self-archiving: the published article or the final peer-reviewed manuscript is archived by the author in repository of their choice and is available free of charge to the reader. Most journals operate with an embargo-period, typically 12 months, before the article can be self-archived.
ICT	Information Communication Technology
PROMs	Patient-Related Outcome Measures
Open access	The practice of providing online access to scientific information that is free of charge to the reader and that is re-usable under specified conditions
Open data	Research data that is made openly accessible for any third party to be mined, exploited, reproduced and disseminated free of charge for the user.
QKPI	Qualitative Key Performance Indicators
SOP	Standard Operating Procedures

1. Introduction

1.1. GERONTE and its objectives

GERONTE is a 5-year research and innovation action (April 2021 to March 2026) funded by the European Union within the framework of the H2020 Research and Innovation programme, in response to the health-societal challenge topic SC1-BHC-24-2020 "*Healthcare interventions for the management of the elderly multimorbid patient*". The overall aim of GERONTE is to improve quality of life - defined as well-being on three levels: global health status, physical functioning and social functioning- for older multimorbid patients, while reducing the overall costs of care. To this end, GERONTE will co-design, test, and prepare for deployment an innovative cost-effective patient-centred holistic health management system, hereafter referred to as the GERONTE intervention. The GERONTE intervention will rely on an ICT (information and communication technology) based application for real-time collection and integration of standardised clinical and home patient-reported data. The GERONTE intervention will be demonstrated in the context of care of multimorbid patients having cancer as a dominant morbidity, and be adaptable to any other combination of morbidities.

Objectives

O1: INFORMATION - gather the stakeholders and data needed for patient-centred and multi-actor complex decision-making process and management

O2: **TOOLS -** develop ICT tools for the GERONTE intervention to be implemented

O3: METHODS - develop socio-economic methods for evaluating the impacts of the implementation of the GERONTE intervention

O4: **DEMONSTRATION** - demonstrate in 16 study sites from three EU countries the feasibility and effectiveness of the GERONTE intervention

O5: **REPLICATION -** develop recommendations for the replication of GERONTE best practices in all European health systems

O6: ENGAGEMENT - engage all stakeholders by co-designing the GERONTE intervention

Expected Impacts

In line with the EU-directive on cross-border healthcare, GERONTE expects to deliver a robust care pathway for the management of multimorbid patients, an ICT tool allowing for the easy adoption across Europe, and socio-economic methods and QKPIs for the objective evaluation of its costs savings and quality of life improvement.

1.2. Rationale

This Data Management Plan (DMP) describes the data management life cycle for the datasets to be collected, generated and processed by GERONTE. The DMP outlines the handling of **research data**[1] during the project, and how and what parts of the datasets will be made available after the project has

[1] *refers to information, in particular facts or numbers, collected to be examined and considered as a basis for reasoning, discussion, or calculation. In a research context, examples of data include statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings and images. The focus is on research data that is available in digital form. Users can normally access, mine, exploit, reproduce and disseminate openly accessible research data free of charge*

been completed. This includes an assessment of when and how data can be shared without disclosing directly or indirectly identifiable information from study participants or compromising exploitation for partnering SMEs. The guiding principles are to ensure timely open access to research data and associated publications whenever possible with traceability from design, production to publication (Fig 1.1a and Fig 1.1b)

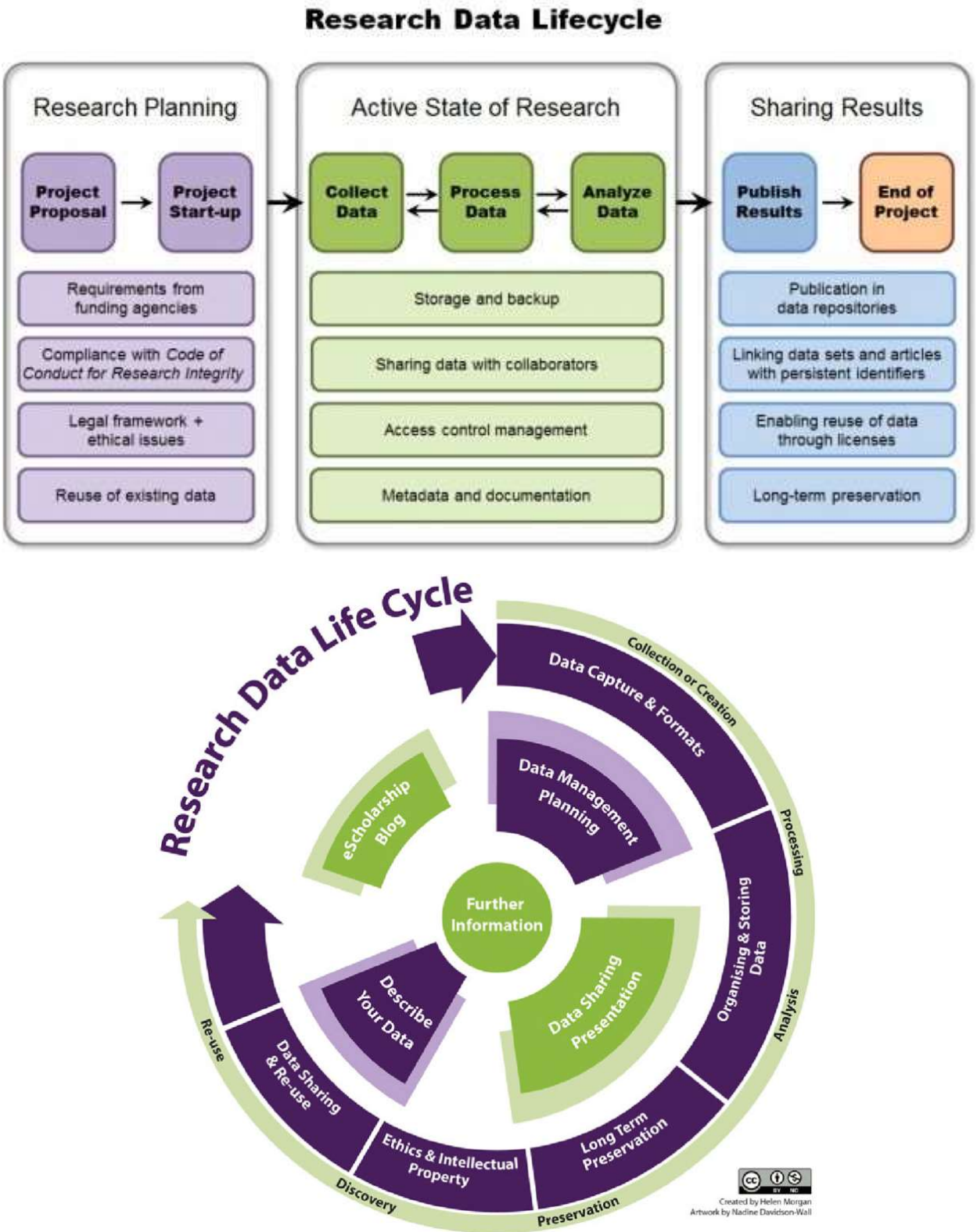

Fig1.1a Two visuals of Research Data Lifecycle (reproduced from DTU AIS Bibiometrics and Data Management, licenced under CCO 1.0; and University of Queensland library, used as examples of guiding principles for GERONTE research data management

Fig1.1b: GERONTE decision chart for open science

This document will cover two types of research data:

1. the 'underlying data' (the data needed to validate the results presented in scientific publications), including the associated metadata (i.e. metadata describing the research data deposited),
2. any other data (for instance curated data not directly attributable to a publication, or raw data), including the associated metadata, as specified and within the boundaries laid down in the DMP – that is, according to the individual judgement of responsible partners.

GERONTE will distinguish collected existing research data and, newly generated research data. The former is meant as existing data produced by various sources, which will be systematically collected and stored together in appropriate repositories. In that case; GERONTE will produce metadata for this type of data, describing the availability of the datasets included in the repository that will be created. Generated data is data created *de novo* as part of the project.

In this regard, GERONTE will share the datasets that are linked to scientific papers that will be published in the scope of the project. There will be raw data as well as analysed, processed and generated or collected data. The shared data will be useful to the readers of the scientific articles that will be published, to validate the conclusions, to reproduce the experiments and finally to go further than the GERONTE project results and broaden its impact in the long-term.

1.3. Stakeholders' engagement

The participation of end-users and stakeholders is central to GERONTE and is intended to take various shapes, some of which may lead to major input to data generated or collected through surveys, interviews, webinars, questionnaires, workshops or participation in small-scale pilots or clinical trials. This deliverable should be read in conjunction with D8.6 GERONTE Ethics Management Plan for a detailed assessment of measures to ensure the protection of the rights of stakeholders participating

in the project (in consideration of personal data, research involving human subjects, and research with volunteers in social sciences and humanities), which is a priority to the GERONTE consortium.

1.4. Intended Readership

This initial document is primarily intended for use internally in the project, to provide guidance on data management to project partners and participants responsible for data collection, storage, transfer and processing. Forthcoming updates will be of use to external actors to gain knowledge of what data has been generated and how to access it. It can also serve as a tool to create awareness on the different topics of Open Science and the FAIR principles.

Throughout the life of the project, Data Management will be coordinated by the GERONTE **Data Manager**, Dr Marije Hamaker (DIAK- participant 3) with contribution on intellectual property rights protection from the GERONTE **Exploitation Manager**, Dr Paul Davis (DCU – participant 10) and general assistance from the GERONTE **Project Manager**, Dr Cecile Bacles (UBx – coordinator). For any comment and/or suggestions, please contact the GERONTE Project Manager: Dr Cecile Bacles, cecile.bacles@u-bordeaux.fr

1.5. Relationship with other deliverables

This document complements the following deliverables and internal documents

- D8.1. Standard Operating Procedures (SOP) which provide the GERONTE consortium which all relevant information for internal communication, project management and reporting;
- PEDR Plan for exploitation and dissemination of results generated by the GERONTE project which details plans for publication and contractual agreements on dissemination, exploitation and intellectual property protection;
- D8.6. Ethics Management Plan which compiles all identified ethical and regulatory issues and describes measures taken by project partners to best address them according to national and EU legislation and recommendations;
- GERONTE consortium agreement which provides rules for general and specific data exchange.

2. Data Summary

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

The GERONTE project has six main research objectives (see section §1.1). The purpose of the data collection/generation aligns with these objectives and is implemented in six corresponding work-packages (WP):

- **The purpose of data collection/generation in WP1 (led by DIAK- participant 3) is to achieve O1: INFORMATION,** that is to define the minimum requirements for stakeholders (whom to involve) and data needed (what to characterise) for patient-centred and multi-actor complex decision-making process and management for improved multimorbid patient care using the GERONTE intervention.

- **The purpose of data collection/generation in WP2 (led by MyPL- participant 7) is to achieve O2: TOOLS,** that is to develop the algorithms and design the front-end and back- end software architecture for data access by the different users (from patient to health professionals) in the GERONTE intervention pathway.

- **The purpose of data collection/generation in WP3 (led by BOC-participant 8) is to achieve O3: METHODS,** that is to gather the necessary information relating to socio-economic parameters associated to the GERONTE care pathway to perform its economic evaluation, to develop relevant QKPI for the management of multimorbid patients, and define best-practices.

- **The purpose of data collection/generation in WP4 (led by KUL- participant 2) is to achieve O4: DEMONSTRATION** that is to analyse the feasibility and effectiveness of the GERONTE intervention against standard practices in real-life across 16 study sites in three EU countries.

- **The purpose of data collection/generation in WP5 (led by DCU- participant 10) is to achieve O5: REPLICATION** that is to gather the necessary information to evaluate the socio-economic conditions for the replication of GERONTE best practices across European health systems.

- **The purpose of data collection/generation in WP6 (led by ESE- participant 6) is to achieve O6: ENGAGEMENT,** that is to inform the design and pilot implementation of the GERONTE digital tools with the participation of various stakeholders (patients, carers and health professionals).

2.2. Specification of types and formats of data collected/generated.

Specify if existing data is being re-used (if any?)
Specify the origin of the data?
State the expected size of the data (if known)?

The types, formats and origin of the data generated or collected during the GERONTE project will vary and result from a range and combination of activities including, among others, reviews of the scientific literature, qualitative interviews or questionnaires, clinical trials or technical experiments (table 2.1). Later updates of this initial DMP will provide a refined and more detailed description (including expected size) of the data actually generated and collected.

Table 2.1 Types of research data expected to be generated or collected in the implementation of the GERONTE project.

WP	Task	Description	Datatype(s) *e.g. Collection, Dataset, Event, Image, Software, Sound, Text, WorkFlow, Other*	Origin e.*g. Analysis, Experiment, Observation, Simulation, Survey, Other*	Re-use from previous research *e.g. clinical research substudy, database search, other*	Lead
1	1	Core group of health professionals for multi-disease decision making	Dataset	survey		UCD
1	2	Core list of multimorbidity indicators for multi-disease decision making	Dataset	survey		UCD
1	3	Core list of intrinsic capacity parameters	Dataset	survey	Literature review	UCD
1	3	Protocol for intrinsic capacity evaluation	Text	Analysis	Literature review	DIAK
1	4	Symptoms and PROMs list for patient status monitoring	Dataset	Survey	Literature review	OUS
1	5	Self-management recommendations	Text	Analysis	Literature review	DIAK
2	1	Digital application user interface design	Software, Dataset, Image, Sound	Survey, Other		MyPl
3	1	Review of cost-effectiveness of service intervention	Text	Analysis	Literature review	BOC
3	2	Protocol for economic evaluation	Text	Analysis	Literature review	BOC
3	4	Dashboard of socio-economic indicators for the management of multimorbidity	Dataset	Survey, Analysis	Clinical research, Literature review	BOC
3	5	Alignment of existing care pathways with the GERONTE model	Workflow	Survey, Observation, Analysis	Literature review	BOC
4	4	Clinical trial eCRFs (demographics, medical history, clinical data, biology, treatments, quality and life and safety data)	Database	Observation		UBx
4	5	Clinical trials summary statistics reports	Dataset, Text	Analysis		UBx
5		Business model	Process Map, Text	Survey, CSV Models		DCU
5		Economic evaluation	Dataset	Analysis		BOC
6	3	Small scale pilot of digital application	Dataset, Text	Survey, Other (ticketing system)		MyPl
6	5	User Support Toolkit	Video, Text	Other		MyPl

3. FAIR data

3.1. Making data findable, including provisions for metadata

Outline the discoverability of data (metadata provision)
Outline the identifiability of the data and refer to standard identification mechanisms.
Do you make use of persistent and unique identifiers such as Digital Object Identifiers?
Outline naming conventions used
Outline the approach towards search keyword
Outline the approach for clear versioning
Specify standard for metadata creation (if any) If there are no standards in your discipline describe what metadata will be created and how?

The naming convention used for all datasets created (either from de novo generation or collection) as a contribution to the GERONTE project outputs is as follows:

GERDATNNN. <extension>

Where

NNN: is a three-digit number referring to the entry in the GERONTE dataset log (annexe 2)

<extension>: is the type of data file (e.g. csv, txt, pdf etc…)

The **Findable** principle aims to facilitate the discovery of data by humans and computer systems and requires the description and indexing of data and metadata.

Metadata refers to “data about data”, i.e., it is the information that describes the data that is being published with sufficient context or instructions to be intelligible for other users. Metadata must allow a proper organization, search and access to the generated information and can be used to identify and locate the data via a web browser or web-based catalogue.

Two types of metadata will be considered within the frame of the GERONTE project: that corresponding to the project publications, and that corresponding to the published research data. In the context of data management, metadata will form a subset of data documentation that will explain the purpose, origin, description, time reference, creator, access conditions and terms of use of a data collection.

The metadata that would best describe the data depends on the nature of the data. For research data generated in GERONTE, it is difficult to establish global criteria for all data, since the nature of the initially considered datasets will be different, so that the metadata will be based on generalised metadata schema as the one used in ZENODO, which includes elements such as:

- Title: free text
- Creator: Last name, first name
- Date
- Contributor: It should always provide information referred to the EU funding and to the GERONTE project itself; mainly, the terms "European Union (EU)" and "Horizon 2020", as well as the name of the action, acronym and the grant number
- Subject: Choice of keywords and classifications

- Description: Text explaining the content of the data set and other contextual information needed for the correct interpretation of the data.
- Format: Details of the file format
- Resource Type: text, image, audio, etc.
- Identifier: DOI
- Access rights: closed access, embargoed access, restricted access, open access.
- Additionally, a readme.txt file could be used as an established way of accounting for all the files and folders comprising the project and explaining how all the files that make up the data set relate to each other, what format they are in or whether particular files are intended to replace other files,

In ZENODO, the chosen centralised repository as GERONTE's main open data archive, versioning is clearly handled (version, date) and a DOI representing all versions, is automatically created on registration.

3.2. Making data openly accessible

Specify which data will be made openly available? If some data is kept closed provide a rationale for doing so?
Specify how the data will be made available
Specify what methods or software tools are needed to access the data? Is document about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?
Specify where the data and associated metadata, documentation and code are deposited
Specify how access will be provided in case there are any restriction

The **Accessible** principle encourages the long-term storage of data and metadata and the facilitation of their access and/or download, specifies the conditions of access (open or restricted access) and use (licence).

In GERONTE, all of the data associated with scientific publication (journal article, conference contribution etc.) will be made openly available and linked to the original publication via its DOI unless there is a specific reason not to publish the data, which will then be reviewed by the appropriate project governing bodies (see also section $4 – data management responsibilities)

Other data may also be made available in open access unless:

- Data is obtained with the permission of third parties (re-use), but the third parties have not agreed to make the data available in open access.
- Data that compromises the protection of a partner(s) intellectual property
- Data is personal or sensitive and cannot be fully anonymised.

Data made available in open access will at a minimum be deposited in a trusted certified centralised repository, based on a known and sustainable business model, namely ZENODO (https://zenodo.org/) via the ZENODO GERONTE community (https://zenodo.org/communities/geronte/)

Authors may also choose to archive files in other archives, for instance if it is required by their institutional or national practice.

Table 3.1: Additional archiving for GERONTE consortium participants

Participant	Additional institutional or national archiving option for making data safely, securely and openly accessible (if any)
1-UBx	HAL with cross-reference to OpenAIRE for H2020 projects: https://hal.archives-ouvertes.fr/
2-KUL	RDR data repository: https://www.kuleuven.be/rdm/en/rdf (available from end 2021)
3-DIAK	*None*
4-UCD	*None*
5-OUS	OUS open access publishing agreements (in Norvegian): https://ehandboken.ous-hf.no/document/16887#23
6-ESE	*None*
7-MYPI	*None*
8-BOC	*None*
9-SIOG	*None*
10-DCU	Doras – Institutional open access repository for research outputs: https://doras.dcu.ie

This will guarantee long term preservation and access and will ensure the following:

- Metadata and data will be retrievable by their identifier using a standardized communications protocol.
- The protocol for information retrieval on the web will be open, free, and universally implementable.
- The protocol will allow for an authentication and authorisation procedure for access to data, where necessary. Metadata will be publicly accessible and licensed under public domain. No authorisation will ever be necessary to retrieve it.
- **M**etadata will be accessible, even when the data will no longer be available.

All GERONTE datasets will be made available no later than the publication of the research findings and will be openly available with no restriction ("open access" in ZENODO), unless restriction is deemed necessary. If data linked to publication is restricted (Intellectual Property Protection) or embargoed (due to the embargo of the related scientific publication), the metadata will be still accessible to indicate the existence of the data while preserving their protection.

GERONTE data is not in proprietary format and its access will not require any commercial software.

3.3. Making data interoperable

Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability

Specify whether you will be using standard vocabulary for all data types present in your dataset, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontology?

The **Interoperable** principle can be broken down into: downloadable, usable, intelligible, and combinable with other data, by humans and machines. Data must be easy to combine with other data sets, both by humans and by computer systems.

Free and open formats will be preferred in GERONTE at least for data dissemination and archiving. Where proprietary formats are not open or interoperable, exports in interoperable formats will be used. For example, in the case of tabular data sharing, we will:

- Use an open and independent format (*e.g.* CSV rather than Excel), and follow good practice for publishing tabulated files (*e.g.* unique information per cell), see https://www.w3.org/TR/2016/NOTE-tabular-data-primer-20160225/.
- Contextualize the data: indicate links to other data (earlier or more recent versions, additional data, etc.), and links to publications (articles citing the data, data papers).

All specifications required to access the data will be inserted in the data repository.

To comply with this principle, the following will be ensured:

- Metadata and data will use a formal, accessible, shared, and broadly applicable language for knowledge representation, preferably following MeSH, i.e. ontological standard provided by the Medical Subject Headings of the National Library of Medicine (https://www.nlm.nih.gove/mesh/meshhome.html). The ability to combine GERONTE data with other data will be achieved by using a standard/controlled vocabulary for naming columns in the file (glossary, ontology to name columns).
- Metadata and data will use vocabularies that follow FAIR principles.
- Metadata and data will include qualified references to other metadata and data. Each referenced external piece of metadata will be qualified by a resolvable URL.

3.4. Increase data re-use (through clarifying licences)

Specify how the data will be licensed to permit the widest reuse possible
Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed
Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project. If the re-use of some data is restricted, explain why.
Describe data quality assurance processes
Specify the length of time for which the data will remain re-usable

This principle highlights the characteristics that make the data reusable for future research or other purposes (education, innovation, replication/transparency of science).

In GERONTE:

- Data and metadata will be richly described with a plurality of accurate and relevant attributes based on ZENODO's enrichments. Semantic interoperability will be ensured by using the ontological standard provided by the Medical Subject Headings of the National Library of Medicine (https://www.nlm.nih.gove/mesh/meshhome.html)
  - Data and metadata will be released with a clear and accessible data usage licence.
    - License is one of the mandatory terms in ZENODO's metadata, and will refer to an open definition license. Level of access will be precisely defined for each dataset, in accordance with the terms and conditions of the GERONTE Consortium Agreement regarding IP protection.
    - Data downloaded by the users will be subject to the license specified in the metadata by the uploader.
  - Data and metadata will be associated with detailed provenance.
    - All data and metadata uploaded will be traceable to a registered ZENODO user.
    - Metadata will describe the original authors of the published work.
  - Data and metadata will meet domain-relevant community standards, with ZENODO not being a domain-specific repository.

GERONTE data, that meets criteria for open access (see §3.2 above) will be made available as soon as the scientific papers to which it is linked have been published, with no unjustified delay. If publisher imposes an embargo, the data will be deposited as soon as the embargo period ends. Metadata are not submitted to such embargo.

The licence Creative Commons Attribution 4.0 International (CC BY 4.0) will be preferred whenever possible. It means that users are free to share - copy and redistribute the material in any medium or format and adapt - remix, transform, and build upon the material - for any purpose, even commercially. Users must give appropriate credit, provide a link to the licence, and indicate if changes were made. They may do so in any reasonable manner, but not in any way that suggests the licensor endorses them or their use.

The general quality assurance process will be clarified in subsequent versions of this DMP. As a first step, quality is already achievable by the peer-review publication standard policy. Specific to the clinical trials, is their prior registration in the centralised database clinicaltrials.gov hosted by the U.S. National Library of Medicine and eCRFs (electronic Case Report Form), fully validated and compliant with FDA CFR 21, part 11 (Criteria for electronic records – electronic signatures) will be designed and set up for sites to enter data electronically (see also §5 Data security).

Open data will remain re-usable for at least 20 years, as to the current policy of the ZENODO repository.

4. Allocation of resources

Estimate the costs for making your data FAIR.
Describe how you intend to cover these costs
Clearly identify responsibilities for data management in your project
Describe costs and potential value of long-term preservation

Direct costs incurred for making the data FAIR are included in the project's budget partly with provision made by each beneficiary partner for fees for publication in Gold Open Access of at least one scientific journal article, as well as dedicated staff time. Such costs are being kept to a minimum by choosing to deposit in free-to-use repositories, datasets or scientific article manuscripts whenever possible, especially when the appropriate journal of choice for publication does not offer an option for immediate open access, including:

1) Archiving GERONTE open datasets, manuscripts, in ZENODO, under the GERONTE ZENODO community page, linked to GitHub for code.
2) For Researchers at French public organisations such as UBx (Participant 1), and in compliance with the national plan for open science, also deposit authored contribution (published or unpublished) in the free-to-use common open-access national repository HAL (https://hal.archives-ouvertes.fr/), with a crossed-linked reference to OpenAIRE for H2020 funded contributions. As coordinator of the GERONTE project, UBx is committed to deposit GERONTE outputs on HAL to ensure their long-term preservation and accessibility.

Should publishers not offer a gold or diamond open-access option for journals selected for peer-reviewed scientific publications, journals rules will be checked to ensure that authors may post Author's accepted post-print or pre-print version in the chosen open repositories (with or without embargo) to ensure immediate accessibility.

Responsibilities for data management in GERONTE is shared at different levels.

Clinical trial data management for GERONTE trials FRONE in France and TWOBE in Belgium and the Netherlands falls under legal responsibility of the trial sponsors, respectively UBx (coordinator) linked third party, Institut Bergonié, and KUL (participant 2). The Clinical Trial Data Manager will be responsible for creation and management of the trials' e-CRFs.

For data other than clinical trial data, it is the responsibility of every beneficiary partner generating, collecting or re-using data from third-party sources, to curate (including backup and storage) their authored datasets according to their own national and institutional regular practice during the production stage, and to deposit them in the corresponding repositories.

It is also the responsibility of every beneficiary partner to inform GERONTE data manager of data generated by them by filling the dataset descriptor template (annexe 1) and the GERONTE dataset log (annexe 2), made available to GERONTE consortium partners via GERONTE secured online collaborative platform (see D8.1 SOP).

Overall monitoring and regular updates of the data management plan fall under the responsibility of GERONTE data manager with general assistance from GERONTE project manager. It is expected that updates of the data management plans will be released at the end of each periodic (including final) reporting period (M18, M36, M48, M60).

- This does not preclude obligation of GERONTE beneficiary partners to follow the rules concerning the (joint-) ownership, dissemination and exploitation of the GERONTE project results set in GERONTE Consortium Agreement, signed by all the beneficiary partners (section 8).
- This does not preclude obligation of WP leaders to report at Project Steering Committee at least every 6-month on progress, intellectual property protection and dissemination, and preparation of the content and timing of press releases or joint publications by the consortium or proposed by the Funding Authority as established in GERONTE Consortium Agreement, signed by all the beneficiary partners (section 6) and detailed in GERONTE PEDR.

Potential value and resources necessary for long-term preservation of data, other than that already made available on identified centralised repositories, beyond the end of the project and accounting for partners regular practice, will be reviewed at Project Steering Committee and described in a later update of the DMP.

5. Data security

Address data recovery as well as secure storage and transfer of sensitive data

GERONTE Consortium partners must refer to section 4.6 (Data Exchange) of the GERONTE consortium agreement signed by all beneficiary partners for the general and specific rules applicable to the secure storage and transfer of sensitive data, in accordance with the Data Protection Laws. This includes secured storage, recover and transfer of personal data. (see also section 6 below).

Specific measures in place for the secured storage and transfer of sensitive data are described in the GERONTE ethics management plan (D8.6). For instance, clinical trial data security (including secured collection, storage and transfer) will be dealt independently and centrally with the creation and maintenance of eCRFs for clinical trials FRONE and TWOBE by their respective sponsors, UBx-LTP, Institut Bergonié and KUL as described section §4.6.2 of the GERONTE consortium agreement signed by all beneficiary partners. Both trials will be registered in the centralised database clinicaltrials.gov hosted by the U.S. National Library of Medicine and eCRFs (electronic Case Report Form), fully validated and compliant with FDA CFR 21, part 11 (Criteria for electronic records – electronic signatures), and will be designed and set up for sites to enter data electronically. Programmed and manual queries on the data completed in the eCRF will be raised centrally by GERONTE clinical Trial Data Manager, and manual queries may be raised at the national level by CRAs. Self-Evident Correction plans will be developed and endorsed by site PIs. GERONTE participants will be able to request access to the anonymised project data via a specific query form which will be reviewed for approval by the Trial Steering Committee.

Long-term data safe storage beyond the end of the project of research data- that has been made open- is ensured by ZENODO, as a certified repository where all data files are stored in CERN Data Centres, primarily Geneva, with replicas in Budapest. Data files are kept in multiple replicas in a distributed file system, which is backed up to tape on a nightly basis. In addition, beneficiary partners responsible for production and curation of research data will follow regular institution practices. By default, the following recommendation will apply as a minimum security standard: Storage of both data in

production (live data) and final data, in at least two different locations, secured and compliant with EU law on privacy and digital security, with regular, if possible automated, backup.

6. Ethical aspects

To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former.

The ethical aspects relating to the collection and storage of research data for the GERONTE project are addressed in deliverable D8.6 Ethics Management Plan (WP8 Task 8.7 Ethics Management), excluding clinical trial data management (eCRFs) which is the responsibility of trials sponsors in France and Belgium (respectively, UBx linked third party, Institut Bergonié; and KUL) under WP4 task 4.4 clinical data trial management and collection, both link with WP9 (Ethics requirements) of the ethics review. To facilitate identification of potential ethics issues and means to address them, data monitoring tools include dedicated sections that will help both the GERONTE data manager and GERONTE ethics manager to coordinate with WP leaders at project steering committee (see Annexe 1 and 2).

7. Other

Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any).

GERONTE consortium partners responsible for management of datasets at their own institution, will follow national and institution policies for research data management (see table 7.1) as well as Horizon 2020 guidelines on FAIR data management. If no institution or national policy available, the standard policy will be; to follow the guidelines on FAIR data management in Horizon 2020, which available in the online manual, and, to adopt the following practice (as used herein) for each generated dataset:

1. Make a data management plan and store it online securely and safely– (for instance using ARGOS).
2. Store your live data securely and safely
3. Before sharing, enquire about ethics and intellectual property limitations (with support from your research office, when applicable)
4. Store your final data for 10+ years, making it publicly available in a centralised archive
5. Inform your research support office or librarian (when applicable) where your data is published or stored
6. Reference the relevant funding and data sources for the publications it underpins.

Table 7.1 : Institutional or National Research Data Management Guidelines of GERONTE participants

Participant	Country	Data Management Policy or Guidelines (if any)
UBx	FR	Second National Plan for Open Science (Plan S V2– 2021) https://www.ouvrirlascience.fr/second-national-plan-for-open-science/
KUL	BE	KU Leuven RDM policy https://wwww.kuleuven.be/rdm/en/rdm-policy/policy
DIAK	NL	N/A
OUS	NO	(In Norvegian) https://ehandboken.ous-hf.no/folder/37, https://www.iou.no/english/services/it/research/sensitive-data/index.html
UCD	IE	https://hub.ucd.ie/usis/!W_HU_MENU.P_PUBLISH?p_tag=GD-DOCLAND&ID=227
ESE	FR	N/A
MyPL	FR	N/A
BOC	IT	https://www.unibocconi.eu/wps/wcm/connect/Bocconi/SitoPubblico_EN/Navigation+Tree/Home/faculty+and+research/research/grants+office/resources/; https://www.unibocconi.it/wps/wcm/connect/Bocconi/SitoPubblico_IT/Albero+di+navigazione/Home/Chi+siamo/Statuto+regolamenti+e+documenti/Regolamento+protezione+dati+GDPR/
SIOG	CH	N/A
DCU	IE	https://dcu.libguides.com/dmp

Annexe 1: Template for Research Dataset Descriptor

	Dataset ID *(Please refer to Dataset Log)*			**GERDATNNN**	
	Creator	Name	Surname	**DATE**	YYYY-MM-DD
		ORCID			
PARTNER SERVICES AND RESPONSIBILITIES					
Partner in charge of the data collection/generation	*Fill in beneficiary name – specify if there are several partners collaborating, add their names*				
Partner in charge of the data analysis (if different)	*Fill in beneficiary name – specify if there are several partners collaborating, add their names*				
Partner in charge of the data storage (if different)	*Fill in beneficiary name – specify if there are several partners collaborating, add their names*				
WP and Task	*The data is generated/collected within activities of task X of WP Y*				
ETHICS & SECURITY COMPLIANCE CHECK *(Ethics Management Plan)*					
Personal Data - GDPR	*YES/NO*				
Humans - Research in SSH	*YES/NO*				
Humans- Research with human subjects	*YES/NO*				
If YES, Ethics Management Log	*YYYY-MM-DD*				
INTELLECTUAL PROPERTY CHECK *(PEDR)*					
Data (joint-)Ownership	*Fill in beneficiary name – specify if there are several partners collaborating, add their names*				
Project Steering Committee Report	*YYYY-MM-DD*				
DATA ARCHIVING DESCRIPTION*(DMP)*					
2. DATA SUMMARY					
Please complete all that apply, choosing from available lists or add options if applicable.					
Purpose of the Data		Purpose	Objectives	*Field Description (for information only)*	
			1	*State the purpose of the data collection/generation, indicating the relation with the objectives of the project. Add additional objectives if necessary.*	
			2		
			3		
Type and Format of Data		Form	Format	*Describe the type of data used or generated within the project, specifying the form and format of the data.*	
	Text	type_of_text	type of format	*Form: Field or laboratory notes, survey responses Format: in plain text, (txt), HTML, XLM, PDF/A …*	
	Numeric	Type_of_table	Type of format	*Tables, row counts, measurements, genotypes etc…- in .XLSX, .CSV …*	
	Audiovisual	image		*Images, sound recordings, video - in .JPEG, .JPG, .PNG, .TIFF, AIFF, WAVE, .MP3, .MP4…*	
	Simulated				

	model			*Please state the model, model type and computer code - and specify output datatype and format.*
	model type			
	computer code			
	data type			
	format			
	Discipline specific information	Discipline	format	*Specify domain specific information if any*
	Instrument specific information	Equipment	format	*Equipment output (specify equipment and format).*
	Reused-Data (rd)	YES/NO	If yes, write explanation	*Indicate if you re-use existing data (generated outside the GERONTE project). If so, explain how.*
Data Origin		*Define and describe the origin/source of your data. Data can be gathered from different sources.*		
	Observational			*Data captured in real time - often not reproducible i.e. sensor readings, images, telemetries, sample data*
	Experimental			*Data from lab equipment, often reproducible, but with high costs or large infrastructures - i.e. imaging data, high-throughput sequencing data etc….*
	Simulation			*Data generated by computational models where model and metadata are equally important to output data - i.e. climate models, economic models, materials models,…*
	Derived/Compiled			*Data coming from analysis or compilation. Reproducible but with high costs - i.e. the results of text and data mining, compiled databases..*
	Reference or Canonical (links)		*write the reference or canonical*	*Collection or conglomeration of smaller (peer-reviewed) datasets published and curated - i.e. chemical structures, gene sequence databanks, spatial data portals..*
	Dataset is:	*Fixed/Growing/Revisable*		*Fixed: never change after being collected or generated. Growing : new data may be added, but the old data is never changed or deleted. Revisable : new data may be added, and old data may be changed or deleted.*
	Quantity		in MB/GB	*of each experiment*
			in MB/GB	*overall*

			In case not just digital archiving is required, indicated quantities of other form of storage.

3 FAIR DATA
ZENODO central archive: GERONTE Community information
3.1 FAIR Data Making data findable -METADATA

	Title	*Title of the research data as general descriptor for viewers on ZENEDO (and other archives if any)*	
	Description	*Text explaining the content of the data set and other contextual information needed for the correct interpretation of the data*	
	Identifier (DOI)	*DOI will be automatically generated by ZENODO on registration of the research data set – please report here if known or leave blank*	
	Date	*Date of archive created on the ZENODO GERONTE community if known or leave blank*	
	Contributor	*The GERONTE project leading to this [result/application/other] has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 945218.*	*It can provide information referred to the EU funding and to the GERONTE project itself;. If other funding apply indicate*
	Subject	*Indicate search keywords* / *Describe approach*	*Indicate the approach to keywords generation, indexing and tagging. Use recognised standard: e.g. Medical subject Headings*: *https://www.nlm.nih.gov/mesh/meshhome.html*
	Standards or procedures for metadata creation applied	*Indicate*	*Some references: MODA, EMMO, DataCite Metadata Schema, Open Archives Initiative Object Reuse and Exchange, ISAtools … If there are no standards in your discipline, describe what type of metadata will be created and how.*

	Information on documentation generated that make the data understandable	*Indicate*		*Any information generated that help understand how the data was generated (e.g. template of survey form, methods description…)*
3.2 FAIR data Making data openly accessible – ACCESS RIGHTS				
	Data openly available	*indicate ownership*		*Indicate ownership of the data, if it is openly available or can be made openly available.*
	Data restricted	*users*	*Reasons,*	*Indicate if data access is restricted, to what users, and explain the reasons. +*
	Repository for deposit of data, metadata, documentation and code	*indicate open or private, list additional (institutional, or national, or other) repositories used as well or instead of ZENODO GERONTE community for specified closed access*		*Indicate the (open or private) repositories in which the data, metadata, documentation and code are stored and/or those in which they will be stored in the future.*
	Access conditions	*indicate*	*Explain access request conditions*	*Indicate if there are limitations and restrictions to access the data, and if they are linked to a specific timeframe. Explain how access will be provided after these restrictions are lifted and can be provided if requested (e.g. Data access committee upon submission of a request describing the purpose for reuse of the data)*
3.3 FAIR data - Making data interoperable				
	Data interoperability assessment	*indicate*		*Assess the level of interoperability of the dataset. Indicate data and metadata vocabularies, standards and methodologies followed to facilitate interoperability. Indicate if open standards are used, and (if you know) the range of utilization of proprietary SW and methodologies used to generate and manage the data.*
	SW documentation and other information needed	*indicate*		*Indicate any specific SW documentation that is needed to access the data, or additional information that is needed to understand the data (i.e. abbreviations, supplementary notes).*
3.4 FAIR data– Increase data re-use (through clarifying licences)				
	Quality assurance process	*how assured*	*how controlled & documented*	*Explain how quality of the data is assured, how the consistency and quality of data collection/generation is controlled and documented (e.g. peer-review, replication etc…)*
	Licence for reuse	*Indicate*		*If applicable, define data licensing approach for the dataset wide reuse. Indicate the chosen licence tools. The default licence will be CC-BY 04*

	Timing of data availability for re-use (incl. indications on embargo)	*define*	*indicate*	*If applicable, define the timeframe for making data available for re-use. Indicate any embargo period if required.*
	Link to other research data	*Indicate URL of linked material*		*If applicable indicate cross-reference(URL of DOI) of linked research outputs (e.g. journal article linked to datasets and other supplementary material – and vice-versa)*
	Data usability by Third Parties (after the end of the project)	*Indicate if applicable*		*Indicate any limitation to the use of the data by Third Parties, after the end of the project.*
	Restrictions to data re-use	*indicate*	*explain*	*Indicate and explain any restriction to the re-use of data (i.e. confidentiality agreements, other issues).*
	Length of time of data reusability	*Indicate (if applicable for depository other than ZENODO GERONTE community)*		*Indicate the time limit for the data re-usability, if any.*
4 ALLOCATION OF RESOURCES				
	Costs estimates for making data FAIR	*Estimate if applicable*	*describe*	*Estimate the costs (if any) for making your data FAIR (findable, accessible, interoperable and reusable) and describe how you intend to cover these costs (i.e. institute dedicated resources, dedicated part of the project budget …).*
	Data Management Responsibilities	*identify*		*Identify responsibilities for data management of this dataset (within your research team and organisation and within the project if applicable).*

Annexe 2: Template for dataset Log

GERDATID	DATASET DESCRIPTOR FILLED DATE	CONTACT PERSON NAME	CONTACT PERSON ORCID	DATE ARCHIVED (ZENODO)	DOI (GENERATED BY ZENODO)	OTHER ARCHIVE (1)	OTHER ARCHIVE (2)	OTHER ARCHIVE (3)	LINKED DATA DESCRIPTION (1)	LINKED DATA URL (1)	LINKED DATA DESCRIPTION (2)	LINKED DATA URL (2)	LINKED DATA DESCRIPTION (3)	LINKED DATA URL (3)	COMMENT
GERDAT001	yyyy-mm-dd														
GERDAT002															
GERDAT003															
GERDAT004															
GERDAT005															
GERDAT006															
GERDAT007															
GERDAT008															
GERDAT009															
GERDAT010															
GERDAT011															
GERDAT012															
GERDAT013															
GERDAT014															
GERDAT015															
GERDAT016															
GERDAT017															
GERDAT018															
GERDAT019															
GERDAT020															
…															

www.geronteproject.eu **contact@geronteproject.eu**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 945218 . The sole responsibility for the content of this project lies with the authors. It does not necessarily reflect the opinion of the European Union. The European Commission is not responsible for any use that may be made of the information contained therein.