D1.3

ETHICAL ISSUE AND DATA MANAGEMENT PLAN.

2022.09.27 | xy

emilio
PERSONAL ASSISTANT

AAL
PROGRAMME

Project Number	aal-2021-8-120-CP
Project Acronym	emilio
Duration	01-02-2022 –31-07-2024
Coordinator	IRCCS INRCA
Document ID	
Release Number /Date	V1.1/July 2022
Document Type	Project Deliverable
Original Due Date	July 2022
Dissemination Level	Public
Leading partner	
Contributing Partners	
Reviewed by	

Consortium


INRCA
Istituto di Ricovero e Cura a Carattere Scientifico


SOLVING TEAM


Erdmann Solution
Erdmann Design


ict factory


Magicview
Connected Products


point


vulpia


ISS
INSTITUTE OF SPACE SCIENCE


Transilvania University of Brasov

Content

List of figures

List of Tables

Executive summary

This document presents the ethical principles and the data management standards that will be adopted in the Emilio project throughout the project implementation. This deliverable will guide the project execution as a common direction and ethical protocol for all the partners with particular attention to end user involvement in the different phases of WP2.
Chapter 1 introduces the scope of the document

Chapter 2 describes the general ethical framework and the existing regulations, declarations and conventions under which the project will be carried out.

Chapter 3 is focused on the exploration of the general principles in the use of AI in health, considering different facets: the ethics and governance of artificial intelligence for health (WHO), the emerging trends in the use of AI in clinical care, the collection and use of health data, the ethics guidelines for trustworthy ai (EC) with an in-depth overview of the European White Paper on Artificial Intelligence.

Chapter 4 examines the Eu regulation on data protection in accordance with the European law and the national legislations of the four involved countries Italy, Belgium Switzerland and Romania on data protection and privacy.

Chapter 5 is a thorough description of how the Emilio project will comply with the ethical principles and gives information about data that will be collected and how will be handled and managed by project partners.

The architecture of the Emilio infrastructure has been defined in Deliverable 3.1 but the project is still in the 1st phase and important designing activities are still ongoing so this deliverable will be a dynamic document and shall be subject to amendments when necessary.

1. Introduction

1.1 Purpose of the document

The aim of the document is to draw the boundaries of the framework for the management of the use of data handled by the project, considering that - for the type of data, i.e. health data - there are legislative and ethical rules to be observed in order to guarantee data security.

The content seeks to define a general framework of ethical principles and data management standards to be referred to and respected, in particular with regard to data protection and privacy according to European legislation and the national laws of the four countries involved (Italy, Belgium, Switzerland and Romania), as well as general principles in the use of AI in healthcare and other digital services (such as virtual voice assistants) used in the Emilio project.

2. Legislation and general ethical principles

The goal of Emilio project is the develop an ICT platform to manage a comprehensive set of web services, supporting various use cases to increase Comfort, Vitality, and Safety of elderly clients who live in an assisted facility such as: access to telemedicine service, medication adherence, fall detection, automatic evaluation of conditions, and control of home automation. To this end, an IoT (Internet of Things) infrastructure is deployed in the premise that observes the client's Daily Activities and interacts vocally with the client when needed.
The "primary target group" is made of elderly people who are expected to be not always familiar with new technologies. For this reason Emilio consortium will give high importance to the ethical aspects of the project with the aim of ensure the adequate protection of the privacy and the personal rights of all the users.

2.1 General Ethical framework

The common framework used by the consortium will be represented by the "four principles" approach postulated by Tom Beauchamp and James Childress in their textbook Principles of Biomedical Ethics. It recognizes four basic moral principles, which are to be judged and weighed against each other, with attention given to the scope of their application. The four principles are:
Beneficence
The principle of beneficence is in medical ethics the obligation of physician to act for the benefit of the patient. Applied to Emilio it means that the proposed ICT solution should benefit the participant according to his or her own conception of the good.
Nonmaleficence
Nonmaleficence is in medical ethics the obligation of a physician not to harm the patient. This simply stated principle supports several moral rules – do not kill, do not cause pain or suffering, do not incapacitate, do not cause offense, and do not deprive others of the goods of life. The practical application of nonmaleficence in Emilio will imply that the operation of the ICT platform should not harm the participant, or put him or her under unacceptable risk, included risks to privacy.
Autonomy
The philosophical underpinning for autonomy, is that all persons have intrinsic and unconditional worth, and therefore, should have the power to make rational decisions and moral choices, and each should be allowed to exercise his or her capacity for self-determination. The principle of autonomy does not extend to persons who lack the capacity (competence) to act autonomously.
People that will be involved in Emilio project are autonomous with cognitive and functional abilities to make decisions so, participation in the study and in the general operation of the system should be based upon a process of informed consent, and the participants right to control his or her personal information will be respected at all times.
Justice.
Justice is generally interpreted as fair, equitable, and appropriate treatment of persons. Of the several categories of justice, the one that is most pertinent to clinical ethics is distributive justice. Distributive justice refers to the fair, equitable, and appropriate distribution of health-care resources determined by justified norms that structure the terms of social cooperation. The consortium will avoid any bias based on gender, culture, nationality, or other sources of social prejudice (this includes fair selection of the subjects for the user trials). Benefits of the study will be shared with the involved communities (this includes publication of the results of the study).

2.1.1. Declaration of Helsinki

The Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects is a set of ethical principles regarding human experimentation developed originally in 1964 for the medical community by the World Medical Association (WMA). It is considered as the cornerstone document on human research ethics. The current revision as updated during 64th WMA General Assembly in Fortaleza, Brazil, October 2013 is the accepted basis for clinical study ethics, and must be fully followed and respected by all engaged in research on human beings. Any exceptions must be justified and stated in the protocol.
The fundamental principle is respect for the individual, his or her right to self-determination and the right to make informed decisions regarding participation in research, both initially and during the course of the research. The investigator's duty is solely to the patient or volunteer and while there is always a need for research, the participant's welfare must always take precedence over the interests of science and society, and ethical considerations must always take precedence over laws and regulations. The recognition of the increased vulnerability of individuals and groups calls for special vigilance. It is recognized that when the research participant is incompetent, physically or mentally incapable of giving consent, then allowance should be considered for surrogate consent by an individual acting in the participant's best interest, although his or her consent should still be obtained if at all possible.
Research should be based on a thorough knowledge of the scientific background, a careful assessment of risks and benefits, have a reasonable likelihood of benefit to the population studied and be conducted by suitably trained investigators using approved protocols, subject to independent ethical review and oversight by a properly convened committee
Ethical principles extend to publication of the results and consideration of any potential conflict of interest ed informed consent.

2.1.2. The European Charter of Fundamental Rights

The Charter of Fundamental Rights of the European Union brings together the most important personal freedoms and rights enjoyed by citizens of the EU into one legally binding document. The Charter was declared in 2000, and came into force in December 2009 along with the Treaty of Lisbon.
The Charter contains some 54 articles divided into seven titles. The first six titles deal with substantive rights under the headings: dignity, freedoms, equality, solidarity, citizens' rights and justice, while the last title deals with the interpretation and application of the Charter.
The European Charter of Fundamental Rights contains several principles which can be relevant in the context of research. These principles form the basis of important ethics guidelines but also support the conduct of research. The most relevant articles are mentioned in **Errore. L'origine riferimento non è stata trovata.**.

Article	Content
Art 3 - Right to the integrity of the person	Everyone has the right to respect for his or her physical and mental integrity.

	In the fields of medicine and biology, the following must be respected in particular: – The free and informed consent of the person concerned, according to the procedures laid down by law. – The prohibition of eugenic practices, in particular those aiming at the selection of persons. – The prohibition on making the human body and its parts as such a source of financial gain. – The prohibition of the reproductive cloning of human beings.
Art. 7 – Respect for private and family life	Everyone has the right to respect for his or her private and family life, home and communications.
Art. 8 - Protection of personal data	Everyone has the right to the protection of personal data concerning him or her. Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified. Compliance with these rules shall be subject to control by an independent authority
Art. 13 - Freedom of the arts and sciences	The arts and scientific research shall be free of constraint. Academic freedom shall be respected.
Art 25 - The rights of the elderly	The Union recognises and respects the rights of the elderly to lead a life of dignity and independence and to participate in social and cultural life.

Table 1 Most relevant article of Charter of Fundamental Rights

2.1.3. EU Policy recommendations for Responsible Research and Innovation in Health and Ageing

Responsible Research and Innovation (RRI) is a newly emerging approach to govern science and innovation. Currently, its main champions are European funding agencies.
Hence Health and ageing are recognised as major societal challenges for the EU and together they constitute a large section of public spending. Included in this, is funding for innovations that are necessary to cope with demographic ageing. In the public sector, the ageing trend in the EU is projected to require age-related expenditure to rise by 1.4 percentage points of GDP by 2060 compared to 2013.
The project focuses on ICT-based solutions (applications, products, services) for an ageing society, a growing sector combining two essential aspects for RRI: it is research and technology intensive and it raises significant social and ethical dilemmas that need to be addressed during product development.
A Framework for pursuing RRI has been designed as output of the responsible industry project to provide researchers and innovator with strategic options and recommendations for pursuing responsible practices in research and innovation, and improving ethical acceptability, social desirability and quality of their devices, products and services.

The market penetration of ICT solutions for an ageing society is still slow and limited Nonetheless, it is expected that this market will grow if the end-user acceptance of ICT technologies increases. Often end users are not convinced that ICT solutions will effectively improve their quality of life and wellbeing.

In order to meet their needs and expectations, the ICT systems/products should be ethically acceptable, affordable, accessible, reliable and easy to use. The implementation of RRI concepts in the industry of ICT for an ageing society could impact positively on these aspects. Increased involvement of stakeholders and the general public in the research and innovation process would bring improved matching of ICT products with societal needs, greater acceptability and increased quality of these products. More generally, an enhanced consideration of societal needs and ethical aspects from the industry could translate into economic benefits.

RRI is intended to go beyond the practices already adopted by some industries (such as corporate social responsibility initiatives, dialogue and stakeholder engagement practices, etc), since it focuses on early adoption of responsible practices along the research and innovation value chain and alignment of R&I outcomes to the needs of end-users and consumers. Figure present the overview schema of RRI framework.

The full implementation of the RRI framework is out of the objective of Emilio project but the partners will work to increase awareness of RRI principles in the consortium.

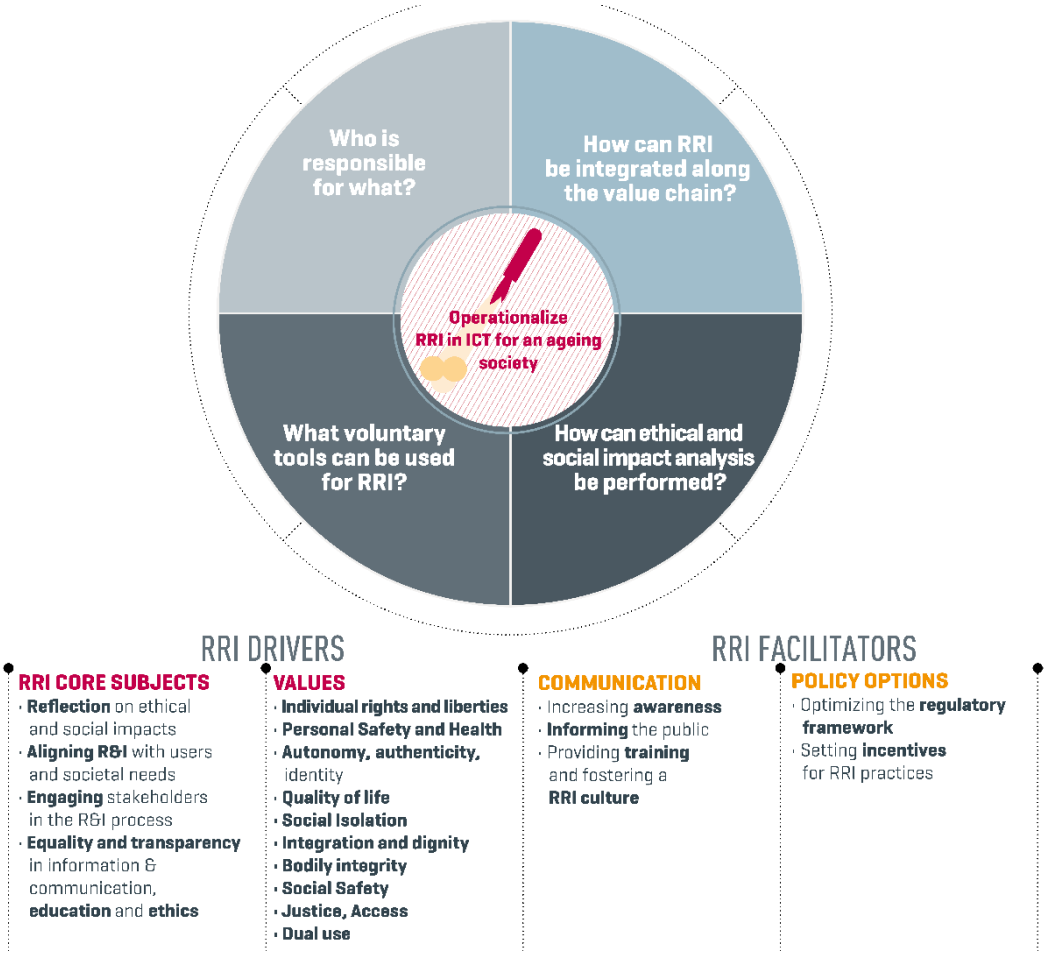

Figure 1 Overview scheme of the RRI framework (extracted from: Implementing Responsible Research and Innovation in ICT for an

ageing society -Executive Brief)

2.1.4. AAL Guidelines for Ethics, Data Privacy and Security

The AAL guidelines for Ethics, Data Privacy and Security has the aim to provide the participants to AAL funded project and the AAL Community of stakeholders with guidelines for ethics, data privacy and security regarding digital solutions for the Active and Healthy Ageing (AHA) domain, fostering two main aspects: to be compliant with existing regulations, standards, etc and to aim for ethical excellence.

An approach that promotes ethical excellence in all stages of the development can leverage the trust of citizens and organisations, fostering the adoption of AAL solutions and services. To this aim, AAL propose a two-fold complementary model of ethical excellence, that integrates compliance (to principles, regulations and standards) with the ethical dialogue.

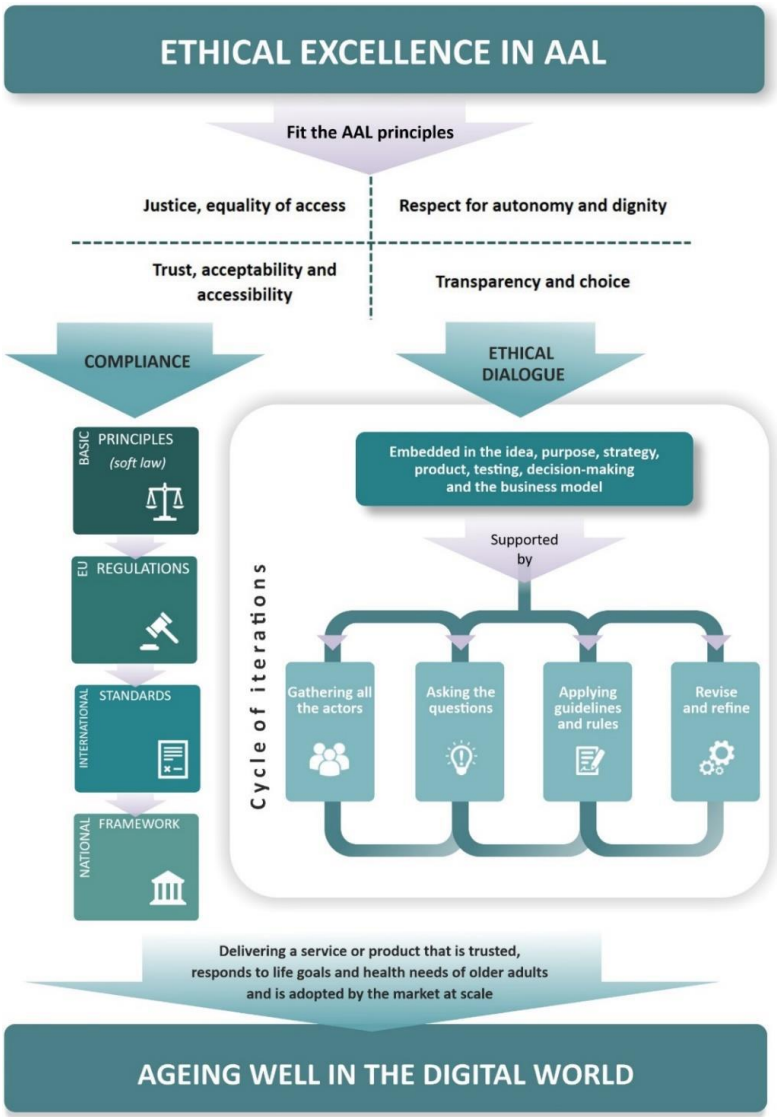

Figure 2 AAL ethical excellence model (from AAL Guidelines for Ethics, Data Privacy and Security)

The ethical dialog is an iterative process that should organized in three main phases corresponding to the different stages of development:

- conceptualisation and (co) creation;
- development& testing;
- market entry & scale-up.

The model is structured to be applied in an iterative way for all the project lifetime. It has three phases:

1) Technology in context

To start the ethical dialogue it is necessary to have a good picture about the technology and the surrounding where it will be used. All the aspects that are relevant for the solution, that often is based on a combination of different technologies and services, each of which has its own ethical aspects should be considered in the ethical dialogue about a solution.

2) Dialogue

To assure the efficacy of the dialogue it is important to include the perspectives of all the different actors that may have different expectations about the effects of the technology. The aim is to get all hopes and worries to the table. The impact is effected by the expectations and the expectations are influenced by the personal values of the different stakeholders.

3) Action opportunities

Analysing the results of the dialogue it is necessary to put everything in the perspective of the final scope that is answer to the question: “What can we do to make it better?”

The actions to be taken in this phase should be tailored to reach three main objectives:

- Ethics by DESIGN (e.g. engineering, design);
- Ethics by CONTEXT (e.g. agreements, policy);
- Ethics by INDIVIDUAL (e.g. behaviour, awareness).

AAL made available a toolkit that includes three forms to implement the different phases of the dialogue.

The correct implementation of the Ethical Dialogue represent an important step to fulfil legal compliance and to drive to ETHICAL EXCELLENCE.

3. General principles in the use of AI in health

3.1 WHITE PAPER on Artificial Intelligence - A European approach to excellence and trust (EC 27)

Artificial Intelligence is developing fast. It will change our lives by improving healthcare (e.g. making diagnosis more precise, enabling better prevention of diseases), increasing the efficiency of farming, contributing to climate change mitigation and adaptation, improving the efficiency of production systems through predictive maintenance, increasing the security of Europeans, and in many other ways that we can only begin to imagine. At the same time, Artificial Intelligence (AI) entails a number of potential risks, such as opaque decision-making, gender-based or other kinds of discrimination, intrusion in our private lives or being used for criminal purposes. Against a background of fierce global competition, a solid European approach is needed, building on the European strategy for AI presented in April 2018. To address the opportunities and challenges of AI, the EU must act as one and define its own way, based on European values, to promote the development and deployment of AI.

The European Commission supports a regulatory and investment oriented approach with the twin objective of promoting the uptake of AI and of addressing the risks associated with certain uses of this new technology.

The EU guidelines are contained in a White Paper[1] published in February 20202, the purpose of which is to to define policy options for achieving these goals. It does not address the development and use of AI for military purposes. The Commission invites Member States, other European institutions and all interested parties to react to the contents of the white paper options and to contribute to future Commission decision-making in this area.

Given the major impact that AI can have on our society and the need to build trust, it is vital that European AI is grounded in European shared values and fundamental rights such as human dignity and privacy protection. Furthermore, the impact of AI systems should be considered not only from an individual perspective, but also from the perspective of society as a whole. The use of AI systems can have a significant role in achieving the Sustainable Development Goals, and in supporting the democratic process and social rights.

With its recent proposals on the European Green Deal, Europe is leading the way in tackling climate and environmental-related challenges. Digital technologies such as AI are a critical enabler for attaining the goals of the Green Deal. Given the increasing importance of AI, the environmental impact of AI systems needs to be duly considered throughout their lifecycle and across the entire supply chain, e.g. as regards resource usage for the training of algorithms and the storage of data.

A common European approach to AI is necessary to reach sufficient scale and avoid the fragmentation of the single market. The introduction of national initiatives risks to endanger legal certainty, to weaken citizens' trust and to prevent the emergence of a dynamic European industry.

The White Paper presents policy options to enable a trustworthy and secure development of AI in

[1] "White Paper on Artificial Intelligence: a European approach to excellence and trust" 19 February 2020 https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

Europe, in full respect of the values and rights of EU citizens. The main building blocks of this White Paper are:

- The policy framework setting out measures to align efforts at European, national and regional level. In partnership between the private and the public sector, the aim of the framework is to mobilise resources to achieve an 'ecosystem of excellence' along the entire value chain, starting in research and innovation, and to create the right incentives to accelerate the adoptionof solutions based on AI, including by small and medium-sized enterprises (SMEs).
- The key elements of a future regulatory framework for AI in Europe that will create a unique 'ecosystem of trust'. To do so, it must ensure compliance with EU rules, including the rules protecting fundamental rights and consumers' rights, in particular for AI systems operated in the EU that pose a high risk. Building an ecosystem of trust is a policy objective in itself, and should give citizens the confidence to take up AI applications and give companies and public organisations the legal certainty to innovate using AI. The Commission strongly supports a human-centric approach based on the Communication on Building Trust in Human-Centric AI and will also take into account the input obtained during the piloting phase of the Ethics Guidelines prepared by the High-Level Expert Group on AI.

As with any new technology, the use of AI brings both opportunities and risks. Citizens fear being left powerless in defending their rights and safety when facing the information asymmetries of algorithmic decision-making, and companies are concerned by legal uncertainty. While AI can help protect citizens' security and enable them to enjoy their fundamental rights, citizens also worry that AI can have unintended effects or even be used for malicious purposes. These concerns need to be addressed. Moreover, in addition to a lack of investment and skills, lack of trust is a main factor holding back a broader uptake of AI.
That is why the Commission set out an AI strategy on 25 April 2018 addressing the socioeconomic aspects in parallel with an increase in investment in research, innovation and AI-capacity across the EU. It agreed a Coordinated Plan with the Member States to align strategies. The Commission also established a High-Level Expert Group that published Guidelines on trustworthy AI in April 2019[2].

The Commission published a Communication welcoming the seven key requirements identified in the Guidelines of the High-Level Expert Group:

- Human agency and oversight,
- Technical robustness and safety,
- Privacy and data governance,
- Transparency,
- Diversity, non-discrimination and fairness,
- Societal and environmental wellbeing, and
- Accountability.

In addition, the Guidelines contain an assessment list for practical use by companies. During the second half of 2019, over 350 organisations have tested this assessment list and sent feedback. The High-Level Group is in the process of revising its guidelines in light of this feedback and will finalise

[2] https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top

this work by June 2020. A key result of the feedback process is that while a number of the requirements are already reflected in existing legal or regulatory regimes, those regarding transparency, traceability and human oversight are not specifically covered under current legislation in many economic sectors.
On top of this set of non-binding Guidelines of the High-Level Expert Group, and in line with the President's political guidelines, a clear European regulatory framework would build trust among consumers and businesses in AI, and therefore speed up the uptake of the technology. Such a regulatory framework should be consistent with other actions to promote Europe's innovation capacity and competitiveness in this field. In addition, it must ensure socially, environmentally and economically optimal outcomes and compliance with EU legislation, principles and values.
This is particularly relevant in areas where citizens' rights may be most directly affected, for example in the case of AI applications for law enforcement and the judiciary.

Developers and deployers of AI are already subject to European legislation on fundamental rights (e.g. data protection, privacy, non-discrimination), consumer protection, and product safety and liability rules. Consumers expect the same level of safety and respect of their rights whether or not a product or a system relies on AI. However, some specific features of AI (e.g. opacity) can make the application and enforcement of this legislation more difficult. For this reason, there is a need to examine whether current legislation is able to address the risks of AI and can be effectively enforced, whether adaptations of the legislation are needed, or whether new legislation is needed. Given how fast AI is evolving, the regulatory framework must leave room to cater for further developments. Any changes should be limited to clearly identified problems for which feasible solutions exist.

Member States are pointing at the current absence of a common European framework. The German Data Ethics Commission has called for a five-level risk-based system of regulation that would go from no regulation for the most innocuous AI systems to a complete ban for the most dangerous ones. Denmark has just launched the prototype of a Data Ethics Seal. Malta has introduced a voluntary certification system for AI. If the EU fails to provide an EU-wide approach, there is a real risk of fragmentation in the internal market, which would undermine the objectives of trust, legal certainty and market uptake. A solid European regulatory framework for trustworthy AI will protect all European citizens and help create a frictionless internal market for the further development and uptake of AI as well as strengthening Europe's industrial basis in AI.

3.2 Ethics and governance of artificial intelligence for health (WHO)

Artificial Intelligence (AI) refers to the ability of algorithms encoded in technology to learn from data so that they can perform automated tasks without every step in the process having to be programmed explicitly by a human[3]. WHO recognizes that AI holds great promise for the practice of

[3] A specific definition of AI in a recommendation of the Council on Artificial Intelligence of the OECD (4) states:
An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy.

public health and medicine and it recognizes that, to fully reap the benefits of AI, ethical challenges for health care systems, practitioners and beneficiaries of medical and public health services must be addressed.

Whether AI can advance the interests of patients and communities depends on a collective effort to design and implement ethically defensible laws and policies and ethically designed AI technologies. There are also potential serious negative consequences if ethical principles and human rights obligations are not prioritized by those who fund, design, regulate or use AI technologies for health. AI's opportunities and challenges are thus inextricably linked.

AI can augment the ability of health-care providers to improve patient care, inform the decisions of health policy-makers or allocate resources within health systems. To unlock this potential, health-care workers and health systems must have detailed information on the contexts in which such systems can function safely and effectively.

AI can also empower patients and communities to assume control of their own health care and better understand their evolving needs. To achieve this, patients and communities require assurance that their rights and interests will not be subordinated to the powerful commercial interests of technology companies or the interests of governments in surveillance and social control. It also requires that the potential of AI to detect risks to patient or community health is incorporated into health systems in a way that promote the human autonomy as per as respect the dignity and does not move humans from the centre of health decision-making.

AI technologies are also changing where people access health care. AI technologies for health are increasingly distributed outside regulated health-care settings, including at the workplace, on social media and in the education system. With the rapid proliferation and evolving uses of AI for health care, including in response to the COVID-19 pandemic, government agencies, academic institutions, foundations, nongovernmental organizations and national ethics committees are defining how governments and other entities should use and regulate such technologies effectively. Ethically optimized tools and applications could sustain widespread use of AI to improve human health and the quality of life, while mitigating or eliminating many risks and bad practices. To date, there is no comprehensive international guidance on use of AI for health in accordance with ethical norms and human rights standards.

AI systems must be carefully designed to reflect the diversity of socioeconomic and health-care settings and be accompanied by training in digital skills, community engagement and awareness-raising. Systems based primarily on data of individuals in high-income countries may not perform well for individuals in low- and middle-income settings. Country investments in AI and the supporting infrastructure should therefore help to build effective health-care systems by avoiding AI that encodes biases that are detrimental to equitable provision of and access to healthcare services.

The guidance document[4], produced jointly by WHO's Health Ethics and Governance unit in the department of Research for Health and by the department of Digital Health and Innovation, endorses a set of six key ethical principles. WHO hopes that these principles will be used as a basis for governments, technology developers, companies, civil society and inter-governmental organizations to adopt ethical approaches to appropriate use of AI for health:

1. **Protecting human autonomy:** Use of AI can lead to situations in which decisionmaking power could be transferred to machines. The principle of autonomy requires that the use of AI or other

[4] Ethics and governance of artificial intelligence for health: WHO guidance ISBN 978-92-4-002920-0 (electronic version) ISBN 978-92-4-002921-7 (print version)- © World Health Organization 2021

computational systems does not undermine human autonomy. In the context of health care, this means that humans should remain in control of health-care systems and medical decisions. Respect for human autonomy also entails related duties to ensure that providers have the information necessary to make safe, effective use of AI systems and that people understand the role that such systems play in their care. It also requires protection of privacy and confidentiality and obtaining valid informed consent through appropriate legal frameworks for data protection.

2. **Promoting human well-being and safety and the public interest**. AI technologies should not harm people. The designers of AI technologies should satisfy regulatory requirements for safety, accuracy and efficacy for well-defined use cases or indications. Measures of quality control in practice and quality improvement in the use of AI over time should be available. Preventing harm requires that AI not result in mental or physical harm that could be avoided by use of an alternative practice or approach.
3. **Ensuring transparency, explainability and intelligibility**. AI technologies should be intelligible or understandable to developers, medical professionals, patients, users and regulators. Two broad approaches to intelligibility are to improve the transparency of AI technology and to make AI technology explainable. Transparency requires that sufficient information be published or documented before the design or deployment of an AI technology and that such information facilitate meaningful public consultation and debate on how the technology is designed and how it should or should not be used. AI technologies should be explainable according to the capacity of those to whom they are explained.
4. **Fostering responsibility and accountability**. Humans require clear, transparent specification of the tasks that systems can perform and the conditions under which they can achieve the desired performance. Although AI technologies perform specific tasks, it is the responsibility of stakeholders to ensure that they can perform those tasks and that AI is used under appropriate conditions and by appropriately trained people. Responsibility can be assured by application of "human warranty", which implies evaluation by patients and clinicians in the development and deployment of AI technologies. Human warranty requires application of regulatory principles upstream and downstream of the algorithm by establishing points of human supervision. If something goes wrong with an AI technology, there should be accountability. Appropriate mechanisms should be available for questioning and for redress for individuals and groups that are adversely affected by decisions based on algorithms.
5. **Ensuring inclusiveness and equity.** Inclusiveness requires that AI for health be designed to encourage the widest possible appropriate, equitable use and access, irrespective of age, sex, gender, income, race, ethnicity, sexual orientation, ability or other characteristics protected under human rights codes. AI technology, like any other technology, should be shared as widely as possible. AI technologies should be available for use not only in contexts and for needs in high-income settings but also in the contexts and for the capacity and diversity of Low and/or Middle Income Countries (LMIC). AI technologies should not encode biases to the disadvantage of identifiable groups, especially groups that are already marginalized. Bias is a threat to inclusiveness and equity, as it can result in a departure, often arbitrary, from equal treatment. AI technologies should minimize inevitable disparities in power that arise between providers and patients, between policy-makers and people and between companies and governments that create and deploy AI technologies and those that use or rely on them. AI tools and systems should be monitored and evaluated to identify disproportionate effects on specific groups of people. No technology, AI or otherwise, should sustain or worsen existing forms of bias and discrimination.
6. **Promoting AI that is responsive and sustainable.** Responsiveness requires that designers,

developers and users continuously, systematically and transparently assess AI applications during actual use. They should determine whether AI responds adequately and appropriately and according to communicated, legitimate expectations and requirements. Responsiveness also requires that AI technologies be consistent with wider promotion of the sustainability of health systems, environments and workplaces. AI systems should be designed to minimize their environmental consequences and increase energy efficiency. That is, use of AI should be consistent with global efforts to reduce the impact of human beings on the Earth's environment, ecosystems and climate. Sustainability also requires governments and companies to address anticipated disruptions in the workplace, including training for health-care workers to adapt to the use of AI systems, and potential job losses due to use of automated systems.

3.3 Emerging trends in the use of AI in clinical care

In the health care, the use of AI in medicine raises notions of AI replacing clinicians and human decision making. The prevailing sentiment is, however, that AI is increasingly improving diagnosis and clinical care, based on earlier definitions of the role of computers in medicine and regulations in which AI is defined as a support tool (to improve judgement).

Several important changes imposed by the use of AI in clinical care extend beyond the provider–patient relationship. Four trends described here are the evolving role of the patient in clinical care; the shift from hospital to home-based care; use of AI to provide "clinical" care outside the formal health system; and use of AI for resource allocation and prioritization. Each of these trends has ethical implications.

Several important changes imposed by the use of AI in clinical care extend beyond the provider–patient relationship. Four trends described here are: the evolving role of the patient in clinical care; the shift from hospital to home-based care; use of AI to provide "clinical" care outside the formal health system; and use of AI for resource allocation and prioritization. Each of these trends has ethical implications such as the principal ones listed below.

The evolving role of the patient in clinical care: AI could eventually change how patients self-manage their own medical conditions, especially chronic diseases such as cardiovascular diseases, diabetes and mental problems. Patients already take significant responsibility for their own care, including taking medicines, improving their nutrition and diet, engaging in physical activity, caring for wounds or delivering injections. AI could assist in self-care, including through conversation agents (e.g. "chat bots"), health monitoring and risk prediction tools and technologies designed specifically for individuals with disabilities

The shift from hospital to home-based care: Telemedicine is part of a larger shift from hospital- to home-based care, with use of AI technologies to facilitate the shift. They include remote monitoring systems, such as video-observed therapy for tuberculosis and virtual assistants to support patient care. Even before the COVID-19 pandemic, over 50 health-care systems in the USA were making use of telemedicine services. COVID-19, having discouraged people in many settings from visiting health-care facilities, accelerated and expanded the use of telemedicine in 2020, and the trend is expected to continue. The shift to home-based care has also partly been facilitated by increased use of search engines (which rely on algorithms) for medical information as well as by the growth in the number of text or speech chatbots for health care, the performance of which has improved with

improvements in natural language processing, a form of AI that enables machines to understand human language. The use of chatbots has also accelerated during the COVID-19 pandemic.

Use of AI to extend "clinical" care beyond the formal health-care system: AI applications in health are no longer exclusively used in health-care systems (or home care), as AI technologies for health can be readily acquired and used by nonhealth system entities. This has meant that people can now obtain health-care services outside the health-care system.

Use of AI for resource allocation and prioritization: AI is being considered for use to assist in decision-making about prioritization or allocation of scarce resources. Prognostic scoring systems have long been available in critical care units.

Several AI tools for population and public health can be used in public health programmes. For example, new developments in AI could, after rigorous evaluation, improve identification of disease outbreaks and support surveillance. Several concerns about the use of technology for public health surveillance, promotion and outbreak response must, however, be considered before use of AI for such purposes, including the tension between the public health benefits of surveillance and ethical and legal concern about individual (or community) privacy and autonomy

3.4 The collection and use of health data

The various types of AI technology include machine-learning applications such as pattern recognition, natural language processing, signal processing and expert systems. Machine learning, which is a subset of AI techniques, is based on use of statistical and mathematical modelling techniques to define and analyse data. Many machine-learning approaches are data-driven. They depend on large amounts of accurate data, referred to as "big data", to produce tangible results.

AI could improve the delivery of health care, such as prevention, diagnosis and treatment of disease, and is already changing how health services are delivered in several high-income countries. The possible applications of AI for health and medicine are expanding continually, although the use of AI may be limited outside HIC because of inadequate infrastructure. The applications can be defined according to the specific goals of use of AI and how AI is used to achieve those goals (methods). In health care, usable data have proliferated as a result of collection from numerous sources, including wearable technologies, genetic information generated by genome sequencing, electronic health-care records, radiological images and even from hospital rooms.

The collection, analysis and use of health data, including from clinical trials, laboratory results and medical records, is the bedrock of medical research and the practice of medicine. Over the past two decades, the data that qualify as health data have expanded dramatically. They now include massive quantities of personal data about individuals from many sources, including genomic data, radiological images, medical records and nonhealth data converted into health data.

The various types of data, collectively known as "biomedical big data", form a health data ecosystem that includes data from standard sources (e.g. health services, public health, research) and further

sources (environmental, lifestyle, socioeconomic, behavioural and social). Fig. 2. Health data ecosystem

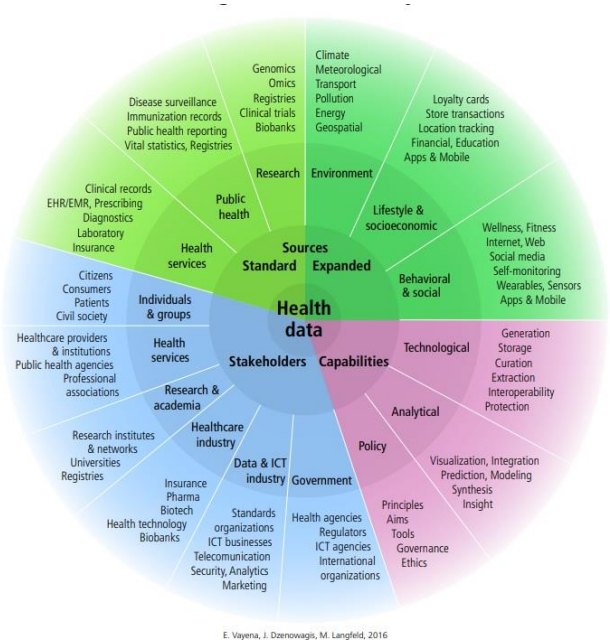

Figure 3 Health data ecosystem Surce: Vayena E, Dzenowagis J, Langfeld M. Evolving health data ecosystem. Geneva: World Health Organization; 2016 (https://www.who.int/ehealth/resources/ecosystem.pdf?ua=1, accessed 17 April 2021)

Thus, there are many more sources of health data, entities that wish to make use of such data and commercial and non-commercial applications. The development of a successful AI system for use in health care relies on high-quality data for both training the algorithm and validating the algorithmic model. The potential benefits of biomedical big data can be ethically important, as AI technologies based on high-quality data can improve the speed and accuracy of diagnosis, improve the quality of care and reduce subjective decision-making. The ubiquity of health data and the potential sensitivity of health care to data indicate possible benefits. Health care is still lagging in the adoption of data science and AI as compared with other sectors and individuals informed of the potential benefits of the collection and use of such data might support use of such data for their personal benefit or that of a wider group. Several concerns may undermine effective use of health data in AI-guided research and drug development. Concern about the use of health data is not limited to their use in AI, although AI has exacerbated the problem. One concern with health data is their quality, especially with those from LMIC (see above). Furthermore, training data will always have one or more systemic biases because of under-representation of a gender, age, race, sexual orientation or other characteristic. These biases will emerge during modelling and subsequently diffuse through the resulting algorithm.

A second major concern is safeguarding individual privacy. The collection, use, analysis and sharing of health data have consistently raised broad concern about individual privacy, because lack of privacy may either harm an individual (such as future discrimination on the basis of one's health status) or cause a wrong, such as affecting a person's dignity if sensitive health data are shared or broadcast to others. There is a risk that sharing or transferring data leaves them vulnerable to cyber-theft or accidental disclosure. Recommendations generated by an algorithm from an individual's

health data also raise privacy concerns, as a person may expect that such “new” health data are private, and it may be illegal for third parties to use “new” health data. Measures to collect data or track an individual’s status and to construct digital identities to store such information have accelerated during the COVID-19 pandemic.

A third major concern is that health data collected by technology providers may exceed what is required and that such excess data, so-called “behavioural data surplus”, is repurposed for uses that raise serious ethical, legal and human rights concerns. The uses might include sharing such data with government agencies so that they can exercise control or use punitive measures against individuals. Such repurposing, or “function creep”, is a challenge that predates but is heightened by the use of AI for health care. A fourth concern with biomedical big data is that it may foster a divide between those who accumulate, acquire, analyse and control such data and those who provide the data but have little control over their use.

3.5 Ethics guidelines for trustworthy ai (EC)

The European Commission to support the implementation of this vision about Artificial Intelligence established the High-Level Expert Group on Artificial Intelligence (AI HLEG), an independent group whose objective was, , among other initiatives, the preparation of AI Ethics Guidelines.
The group identifies Trustworthy AI as foundational ambition, “since human beings and communities will only be able to have confidence in the technology’s development and its applications when a clear and comprehensive framework for achieving its trustworthiness is in place”.
Similarly to what happen in other technologically complex areas that focus on reliability Trustworthy AI hence concerns not only the trustworthiness of the AI system itself, but requires a holistic and systemic approach, encompassing the trustworthiness of all actors and processes that are part of the system’s socio-technical context throughout its entire life cycle.
Trustworthy AI has **three components**, which should be met throughout the system's entire life cycle:

1. **lawful**, complying with all applicable laws and regulations
2. **ethical**, ensuring adherence to ethical principles and values
3. **robust**, both from a technical and social perspective since, even with good intentions, AI systems can cause unintentional harm.

Each component in itself is necessary but not sufficient for the achievement of Trustworthy AI. Ideally, all three components work in harmony and overlap in their operation.
Lawful AI : AI systems operate in contests widely regulated by a plenty of European and National laws. In addition to general horizontally applicable rules, various domain-specific rules exist that apply to particular AI applications (such as for instance the Medical Device Regulation in the healthcare sector). The law provides both positive and negative obligations, which means that it should not only be interpreted with reference to what cannot be done, but also with reference to what should be done and what may be done.
The Guidelines do not directly deal with the first component of Trustworthy AI (lawful AI), but proceed on the assumption that all legal rights and obligations that apply to the processes and

activities involved in developing, deploying and using AI systems remain mandatory and must be duly observed.

The first part of the guidelines is dedicated to the identification of the ethical principles and their correlated values that must be respected in the development, deployment and use of AI systems.

• Develop, deploy and use AI systems in a way that adheres to the ethical principles of: respect for human autonomy, prevention of harm, fairness and explicability. Acknowledge and address the potential tensions between these principles.
• Pay particular attention to situations involving more vulnerable groups such as children, persons with disabilities and others that have historically been disadvantaged or are at risk of exclusion, and to situations which are characterised by asymmetries of power or information, such as between employers and workers, or between businesses and consumers.2
• Acknowledge that, while bringing substantial benefits to individuals and society, AI systems also pose certain risks and may have a negative impact, including impacts which may be difficult to anticipate, identify or measure (e.g. on democracy, the rule of law and distributive justice, or on the human mind itself.) Adopt adequate measures to mitigate these risks when appropriate, and proportionately to the magnitude of the risk.

Table 2: Summarises the Key guidance about identification of ethical principles connected to trustworthy AI

The second part provides guidance on how Trustworthy AI can be realised, by identifying a list of seven requirements that AI systems should meet. **Errore. L'origine riferimento non è stata trovata.** The requirements are listed.

1 Human agency and oversight	Including fundamental rights, human agency and human oversight
2 Technical robustness and safety	Including resilience to attack and security, fall back plan and general safety, accuracy, reliability and reproducibility
3 Privacy and data governance	Including respect for privacy, quality and integrity of data, and access to data
4 Transparency	Including traceability, explainability and communication
5 Diversity, non-discrimination and fairness	Including the avoidance of unfair bias, accessibility and universal design, and stakeholder participation
6 Societal and environmental wellbeing	Including sustainability and environmental friendliness, social impact, society and democracy
7 Accountability	Including auditability, minimisation and reporting of negative impact, trade-offs and redress.

Table 3: List of requirements for trustworthy AI

The implementation of these requirements, requires the use of both technical and non-technical methods in all stages of an AI system's life cycle. AI systems are continuously evolving and acting in a dynamic environment so the realisation of Trustworthy AI is therefore a continuous process, as depicted in Figure 1.

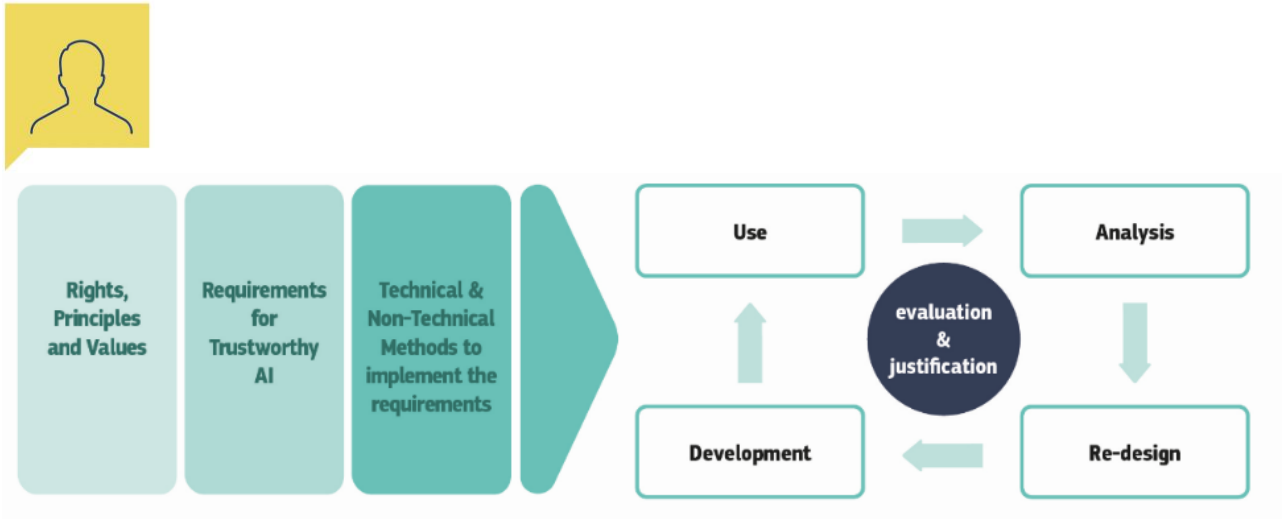

Figure 4: Realising Trustworthy AI throughout the system's entire life cycle (from "Ethics guidelines for trustworthy AI)

3.5.1. Technical methods

The Guidelines identify a set of technical methods that can be incorporated in the design, development and use phases of an AI system to ensure Trustworthy AI.

▪ **Architectures for Trustworthy AI**

Requirements for Trustworthy AI should be "translated" into procedures and/or constraints on procedures, which should be anchored in the AI system's architecture. This could be accomplished through a set of "white list" rules (behaviours or states) that the system should always follow, "black list" restrictions on behaviours or states that the system should never transgress, and mixtures of those or more complex provable guarantees regarding the system's behaviour. Monitoring of the system's compliance with these restrictions during operations may be achieved by a separate process.

AI systems with learning capabilities that can dynamically adapt their behaviour can be understood as non-deterministic systems possibly exhibiting unexpected behaviour. These are often considered through the theoretical lens of a "sense-plan-act" cycle. Adapting this architecture to ensure Trustworthy AI requires the requirements' integration at all three steps of the cycle: (i) at the "sense"-step, the system should be developed such that it recognises all environmental elements necessary to ensure adherence to the requirements; (ii) at the "plan"-step, the system should only consider plans that adhere to the requirements; (iii) at the "act"-step, the system's actions should be restricted to behaviours that realise the requirements.

The architecture as sketched above is generic and only provides an imperfect description for most AI systems. Nevertheless, it gives anchor points for constraints and policies that should be reflected in specific modules to result in an overall system that is trustworthy and perceived as such.

▪ **Ethics and rule of law by design (X-by-design)**

Methods to ensure values-by-design provide precise and explicit links between the abstract principles which the system is required to respect and the specific implementation decisions. The idea that compliance with norms can be implemented into the design of the AI system is key to this method. Companies are responsible for identifying the impact of their AI systems from the very start, as well as the norms their AI system ought to comply with to avert negative impacts. Different "by-design" concepts are already widely used, e.g. privacy-by-design and security-by-design. As indicated above, to earn trust AI needs to be secure in its processes, data and outcomes, and should be designed to be robust to adversarial data and attacks. It should implement a mechanism for fail-safe shutdown and enable resumed operation after a forced shut-down (such as an attack).

▪ **Explanation methods**

For a system to be trustworthy, we must be able to understand why it behaved a certain way and why it provided a given interpretation. A whole field of research, Explainable AI (XAI) tries to address this issue to better understand the system's underlying mechanisms and find solutions. Today, this is still an open challenge for AI systems based on neural networks. Training processes with neural nets can result in network parameters set to numerical values that are difficult to correlate with results. Moreover, sometimes small changes in data values might result in dramatic changes in interpretation, leading the system to e.g. confuse a school bus with an ostrich. This vulnerability can also be exploited during attacks on the system. Methods involving XAI research are vital not only to explain the system's

Chapter 3 provides a concrete and non-exhaustive Trustworthy AI assessment list aimed at operationalising the key requirements set out in Chapter II. This assessment list will need to be tailored to the specific use case of the AI system.3

behaviour to users, but also to deploy reliable technology.

- **Testing and validating**

Due to the non-deterministic and context-specific nature of AI systems, traditional testing is not enough. Failures of the concepts and representations used by the system may only manifest when a programme is applied to sufficiently realistic data. Consequently, to verify and validate processing of data, the underlying model must be carefully monitored during both training and deployment for its stability, robustness and operation within well-understood and predictable bounds. It must be ensured that the outcome of the planning process is consistent with the input, and that the decisions are made in a way allowing validation of the underlying process.

Testing and validation of the system should occur as early as possible, ensuring that the system behaves as intended throughout its entire life cycle and especially after deployment. It should include all components of an AI system, including data, pre-trained models, environments and the behaviour of the system as a whole. The testing processes should be designed and performed by an as diverse group of people as possible. Multiple metrics should be developed to cover the categories that are being tested for different perspectives. Adversarial testing by trusted and diverse "red teams" deliberately attempting to "break" the system to find vulnerabilities, and "bug bounties" that incentivise outsiders to detect and responsibly report system errors and weaknesses, can be considered. Finally, it must be ensured that the outputs or actions are consistent with the results of the preceding processes, comparing them to the previously defined policies to ensure that they are not violated.

- **Quality of Service Indicators**

Appropriate quality of service indicators can be defined for AI systems to ensure that there is a baseline understanding as to whether they have been tested and developed with security and safety considerations in mind. These indicators could include measures to evaluate the testing and training of algorithms as well as traditional software metrics of functionality, performance, usability, reliability, security and maintainability.

3.4.2 Non-technical methods

This section describes a variety of non-technical methods that can serve a valuable role in securing and maintaining Trustworthy AI. These too should be evaluated on an ongoing basis.

- **Regulation**

As mentioned above, regulation to support AI's trustworthiness already exists today – think of product safety legislation and liability frameworks. To the extent we consider that regulation may

need to be revised, adapted or introduced, both as a safeguard and as an enabler, this will be raised in our second deliverable, consisting of AI Policy and Investment Recommendations.

▪ **Codes of conduct**

Organisations and stakeholders can sign up to the Guidelines and adapt their charter of corporate responsibility, Key Performance Indicators ("KPIs"), their codes of conduct or internal policy documents to add the striving towards Trustworthy AI. An organisation working on or with AI systems can, more generally, document its intentions, as well as underwrite them with standards of certain desirable values such as fundamental rights, transparency and the avoidance of harm.

▪ **Standardisation**

Standards, for example for design, manufacturing and business practices, can function as a quality management system for AI users, consumers, organisations, research institutions and governments by offering the ability to recognise and encourage ethical conduct through their purchasing decisions. Beyond conventional standards, co-regulatory approaches exist: accreditation systems, professional codes of ethics or standards for fundamental rights compliant design. Current examples are e.g. ISO Standards or the IEEE P7000 standards series, but in the future a possible 'Trustworthy AI' label might be suitable, confirming by reference to specific technical standards that the system, for instance, adheres to safety, technical robustness and transparency.

▪ **Certification**

As it cannot be expected that everyone is able to fully understand the workings and effects of AI systems, consideration can be given to organisations that can attest to the broader public that an AI system is transparent, accountable and fair.53 These certifications would apply standards developed for different application domains and AI techniques, appropriately aligned with the industrial and societal standards of different contexts. Certification can however never replace responsibility. It should hence be complemented by accountability frameworks, including disclaimers as well as review and redress mechanisms.54

▪ **Accountability via governance frameworks**

Organisations should set up governance frameworks, both internal and external, ensuring accountability for the ethical dimensions of decisions associated with the development, deployment and use of AI systems. This can, for instance, include the appointment of a person in charge of ethics issues relating to AI systems, or an internal/external ethics panel or board. Amongst the possible roles of such a person, panel or board, is to provide oversight and advice. As set out above, certification specifications and bodies can also play a role to this end. Communication channels should be ensured with industry and/or public oversight groups, sharing best practices, discussing dilemmas or reporting emerging issues of ethical concerns. Such mechanisms can complement but cannot replace legal oversight (e.g. in the form of the appointment of a data protection officer or equivalent measures, legally required under data protection law).

▪ **Education and awareness to foster an ethical mind-set**

Trustworthy AI encourages the informed participation of all stakeholders. Communication, education and training play an important role, both to ensure that knowledge of the potential impact of AI systems is widespread, and to make people aware that they can participate in shaping the societal development. This includes all stakeholders, e.g. those involved in making the products (the designers and developers), the users (companies or individuals) and other impacted groups (those who may not purchase or use an AI system but for whom decisions are made by an AI system, and society at large). Basic AI literacy should be fostered across society. A prerequisite for educating the public is to ensure the proper skills and training of ethicists in this space.

▪ **Stakeholder participation and social dialogue**

The benefits of AI systems are many, and Europe needs to ensure that they are available to all. This requires an open discussion and the involvement of social partners and stakeholders, including the general public. Many organisations already rely on stakeholder panels to discuss the use of AI systems and data analytics. These panels include various members, such as legal experts, technical experts, ethicists, consumer representatives and workers. Actively seeking participation and dialogue on the use and impact of AI systems supports the evaluation of results and approaches, and can particularly be helpful in complex cases.

▪ **Diversity and inclusive design teams**

Diversity and inclusion play an essential role when developing AI systems that will be employed in the real world. It is critical that, as AI systems perform more tasks on their own, the teams that design, develop, test and maintain, deploy and procure these systems reflect the diversity of users and of society in general. This contributes to objectivity and consideration of different perspectives, needs and objectives. Ideally, teams are not only diverse in terms of gender, culture, age, but also in terms of professional backgrounds and skill sets.

The European Commission to support the implementation of this vision about Artificial Intelligence established the High-Level Expert Group on Artificial Intelligence (AI HLEG), an independent group whose objective was, , among other initiatives, the preparation of AI Ethics Guidelines.

The group identifies Trustworthy AI as foundational ambition, “since human beings and communities will only be able to have confidence in the technology’s development and its applications when a clear and comprehensive framework for achieving its trustworthiness is in place”.

Similarly to what happen in other technologically complex areas that focus on reliability Trustworthy AI hence concerns not only the trustworthiness of the AI system itself, but requires a holistic and systemic approach, encompassing the trustworthiness of all actors and processes that are part of the system’s socio-technical context throughout its entire life cycle.

Trustworthy AI has three components, which should be met throughout the system's entire life cycle:

1. lawful, complying with all applicable laws and regulations
2. ethical, ensuring adherence to ethical principles and values
3. robust, both from a technical and social perspective since, even with good intentions, AI systems can cause unintentional harm.

Each component in itself is necessary but not sufficient for the achievement of Trustworthy AI. Ideally, all three components work in harmony and overlap in their operation.

4. Legislation and general principles on data protection

4.1 EU regulation on data protection

The General Data Protection Regulation (EU) (GDPR) is a regulation in EU law on data protection and privacy in the European Union (EU) and the European Economic Area (EEA).

Definitions:

Personal data: means any information relating to an identified or identifiable natural person ('data subject').

Data subject: is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

Processing: means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.

Data controller: means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law.

Data processor: means a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller.

Pseudonymisation: means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.

4.2 Data protection principles

The GDPR in article 5.1-2 define six key principles which constitute the core of the rules related to processing personal information.

Lawfulness, fairness & transparency

Personal data must be "processed lawfully, fairly and in a transparent manner in relation to the data subject". This means that data controllers must define their lawful basis for collecting and using personal data. In addition, they must use personal data in a way that is fair towards data subjects. For the processing of data to be fair, data subjects must know of the existence of a processing

operation and must be given full and accurate information about it. Information that must be provided to data subjects is described in Articles 13 and 14 of the GDPR. In particular, according to Article 13 that refer to data collected from Data subject, this information has to include:

- identity and contact details of the controller,
- contact details of the data protection office,
- purposes of processing,
- recipients of the personal data (if any),
- information if the controller intends to transfer the data to a third country or international organisation,
- period for storage of personal data,
- right to withdraw consent,
- information about all other data subject' rights including the right to rectify their data, erase it or restrict its processing.

Transparency is intertwined with fairness. Transparent processing requires data controllers to be open, honest and clear with data subjects from the start about who they are, and how and why they use their personal data. Using clear and plain language is crucial in ensuring that people can understand the information given to them and make an informed choice.

Purpose limitation

Purpose limitation means that personal data must be "collected for specified, explicit, and legitimate purposes." Further or additional processing of data in a way that is incompatible with these initial purposes is prohibited. The exemptions are for archiving purposes in the public interest, scientific or historical research purposes, or statistical purposes. Purposes of the processing must be included in the privacy policy or privacy notice for individuals.

Data minimisation

Personal data must be "adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed". In general, this means that data controllers should collect only the minimum amount of personal data they need to fulfil their purposes. the direct consequence is that to assess whether the data controller holds the right amount of personal data, the scope of data collection should be clearly defined. In addition, it is necessary to periodically review the processing to check if the personal data collect are still relevant and adequate for the purposes.

Accuracy

Personal data must be "accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay".

That is to say that organisations should ensure that the personal data they hold is correct, not misleading and accurate. This might require updating it whenever necessary. If organisations discover that personal data is not correct or misleading, they should take steps to correct it or erase it as soon as possible.

Storage limitation

Personal data must be: "kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes". The exact data retention period is not specified in the GDPR. The key point about this principle is that an

organisation must not keep data for longer than it needs it. The only exemption is for public interest archiving, scientific or historical research, or statistical purposes (if the organisation has appropriate safeguards). Anonymised data can be kept for as long as the organisation wants i.e., indefinitely.

Integrity and confidentiality

Personal data shall be ‘‘processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures”.
More specific indications on the security of data processing are defined in Article 32 of the GDPR which says: “Taking into account the state of the art, the costs of implementation and the nature, scope, context and purposes of processing as well as the risk of varying likelihood and severity for the rights and freedoms of natural persons, the controller and the processor shall implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk”.
In general, this means that both data controllers and processors must ensure a level of security that is appropriate to the risks that may be presented by their processing.

In addition, there is one more general principle in article 5.2 of the GDP which stands for “**Accountability**” defined as the responsibility of the data controller to demonstrate compliance with the other six principles.
Accountability principle include two key elements. The first is that organisations are responsible for compliance with the GDPR. The second element is that they must be able to demonstrate their compliance. Demonstration of compliance may include adoption of certain security measures within the organisation, ensuring a good level of understanding and awareness of data protection amongst staff, designation of a data protection officer, performance of impact assessments etc.

The GDPR grants individuals several rights that must be guarded by any actor who processes personal information. These individual rights include the following:
The right to be informed – Data subjects must be informed by the data controller about all personal data that is collected from them, purposes of the processing, legal basis for the processing, as well as the period for which the personal information will be stored.
The right to access – The data controller must supply data subjects with a copy of all the data they have collected from them.
The right to rectification – Data controllers must correct any data that a data subject feels are incorrect, inaccurate or complete data that an individual feels is incomplete.
The right to erasure – Data subjects has the right to erasure i.e. the ‘right to be forgotten’. Article 17 of the GDPR states that personal data must be erased by the data controller when consent has been withdrawn, there is unlawful processing, personal data is no longer necessary in relation to the purposes for which it was collected.
The right to object – Data subjects have the right to object to processing of their data for direct marketing purposes, profiling or on personal grounds at any time.
Rights in relation to automated decision making and profiling - Data subjects ‘have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her’

4.3 National legislation on data protection

Research and development in the Emilio project will be conducted in Italy, Belgium Switzerland and Romania. The results are not planned to be transferred in the framework of the Emilio project to other countries so the national legislations of the four involved countries will be considered.

4.3.1. Italy

Italy implemented the General Data Protection Regulation (Regulation (EU) 2016/679) ('GDPR') by amending the Personal Data Protection Code, with Decree 101/2018 which integrates the old 196/2003 with the new European legislation and repealing those sections directly conflicting with the GDPR. Supervision over the Personal Data Protection Code is conducted by the Italian data protection authority ('Garante'), which, among other things, acts upon data subjects' complaints, provides specific data protection measures for data controllers and processors, and adopts guidelines to assist organizations' compliance with the GDPR.

4.3.2. Belgium

Belgian data protection law is registered in the three national languages: Dutch, French and German languages.The special categories of the personal collection of data in research projects are described in Article 9.1 of the AVG (Algemene Verordening Gegevensbescherming) and are subject to a ban in principleon processing these data. Nevertheless, Article 9.2 of the AVG provides for a number of exceptions to this ban in principle, i.e. well-defined situations in which the processing of these sensitive data is nevertheless permitted:

- the data subject has expressly consented to it;
- the processing is necessary for the performance of obligations under labor law or social security and social protection law;
- the processing is necessary for the protection of vital interests; processing by associations and bodies -working in the political; philosophical, religious, or trade union fields- in relation to their members; the data may not be communicated to third parties without the consent of the data subjects;
- the processing is necessary for reasons of substantial public interest by or pursuant to law;
- the processing is necessary for the provision of health care; for archiving in the public interest or scientific, historical, or statistical research.

Important is to conduct a letter of consent and the possibility to give easy access to the data used in the project, only gather data that is necessary and delete the data after the project is finished, and inform the participant that the stored data is not available and being used anymore. There are 6 defined rights for this letter of consent:

- Right of access: The participant has the right to ask about the personal data that the project holds about him/her.
- Right of rectification: The participant has the right to ask for a copy to correct it.
- Right to object: The participant has the right to ask to stop processing their personal data.
- Right to restrict processing: The participant has the right to restrict the processing of their personal data.
- Right to oblivion or deletion: The participant has the right to ask to destroy personal data. However, if the project is legally obliged to keep certain personal data, the project may not comply with this

request.
- Right to withdraw the consent: for processing personal data based on the participant's consent, they may withdraw that consent.

4.3.3. Switzerland

Swiss data protection law is rooted in the civil law protection of personality rights. The Federal Constitution of the Swiss Confederation ('the Constitution') provides a constitutional right to privacy. Article 28 of the Swiss Civil Code ('the Civil Code') and the Federal Act on Data Protection 1992 ('FADP') put this fundamental right to privacy into concrete terms at a statutory level. On 25 September, 2020, the Federal Parliament enacted a revised FADP (the final text of which is accessible in German here, French here, and Italian here) ('the Revised FADP'). The Revised FADP will enter into force on 1 September 2023. It implements the requirements of the Council of Europe's Modernised Convention for the Protection of Individuals with Regard to the Processing of Personal Data ('Convention 108+'), and it aligns the FADP with the requirements of the European Union's General Data Protection Regulation (Regulation (EU) 2016/679 ('GDPR') with the aim of retaining the European Commission's adequacy finding.

4.3.4. Romania

The legal rules in Romania are mainly set in the Law No. 190/2018 Implementing the General Data Protection Regulation (Regulation (EU) 2016/679) which reiterates the General Data Protection Regulation (Regulation (EU) 2016/679) ('GDPR') rules and in specific decisions issued by the National Supervisory Authority for Personal Data Processing ('ANSPDCP'), regulates main areas of the GDPR such as when Data Privacy Impact Assessments ('DPIA') will be mandatory, the accreditation of certification bodies, the conducting of investigations and managing complaints, and notifying security breaches. The Law no. 190/2018 reffers to the measures for the application of Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).
This law establishes the measures necessary for the implementation at national level, mainly, of the provisions of Article 6 (2), Article 9 (4), Articles 37-39, 42, 43, Article 83 (7), Article 85 and Articles 87-89 of the General Data Protection Regulation.

The ANSPDCP's guidelines[5] are quite scarce and generic, only reiterating the main GDPR principles and standards.

4.4 EU Digital Services Package

Tuesday 5th July European Parliament approved new rules to codify open markets and consumer rights in the realm of digital technology, including voice assistants and smart homes.
The Digital Services Package (DSP) combines Digital Markets Act (DMA) and Digital Services Act (DSA) built on initial reports of problems in fair competition in tech released last year.
The DSP sets out rules for gatekeeper companies, meaning those with a certain level of revenue,

[5] Reference: https://www.dataprotection.ro/index.jsp?page=home&lang=en

platform userbase, and what the EU calls an “entrenched and durable position.”
The DMA asserts that the biggest names in tech will need to offer a third-party option for voice assistants on their devices, potentially an issue for the respective smart speakers built by Amazon, Google, and Apple. In late 2022, the Council of the European Union is expected to formally adopt the DSA & DMA.

4.5 Guidelines on virtual voice assistants

In recent years the popularity of virtual voice assistant (VVA)greatly increases. Whether it is interfaces integrated in smartphones or connected vehicles and smart TVs or autonomous smart speakers these devices are now quite commons in the everyday life of a lot of people.
There are several advantages to using speech-based interactions such as: the naturalness of the interaction which does not involve specific learning from the users, the speed of execution of the command and the extension of the field of action which can allow faster access to information.
For their ability to facilitate human computer interaction, the main beneficiaries of the voice interface could be people with disabilities or dependency for whom the use of traditional interfaces is problematic. Virtual voice assistance can provide easier access to information and computer resources and thus promote a more inclusive approach.
Health care application based on conventional are becoming diffuse not only to support well-being and prevention, but also for treatment and support.
The technical development and the spread diffusion has given to VVAs access to information of an intimate nature that could, if not properly managed, harm the individuals’ rights to data protection and privacy. Consequently, VVAs and the devices integrating them have been under the scrutiny of different data protection authorities.
The European Data Protection Board ('EDPB') published on 14 July 2021, the final version of its Guidelines 02/2021 on virtual voice assistance following their adoption on 7 July 2021. In particular, the guidelines note that data controllers providing virtual voice assistant services and their processors have to consider both the General Data Protection Regulation (Regulation (EU) 2016/679) ('GDPR') and the Directive on Privacy and Electronic Communications (2002/58/EC) (as amended).
The guidelines consider four of the most common purposes for which VVAs are used and consequently process personal data: executing requests, improving the machine learning model, biometric identification and profiling for personalized content or advertising.
As such, the guidelines identify the most relevant compliance challenges and provide recommendations to relevant stakeholders on how to address them.
Firstly, Guidelines recommend the respect of general principles relating to the processing of personal data, with particular reference to special categories (e.g. in the health industry) as well as: lawfulness, fairness and transparency of the processing, purposes limitation, data minimization and storage limitation.
Regarding the purposes there are further aspects that VVAs have to respect: information about the processing of personal data, specific storage period for each processing of personal data, the identification of the legal basis for each specific processing of personal data, as well as the implementation of appropriate technical and organisational measures to ensure a level of security appropriate to the risk.
More specifically, the guidelines highlight the increased complexity for data controllers who provide

such services in meeting GDPR transparency requirements, as a result of the multiple users of virtual voice assistants, the ecosystem complexities, and the specificities of the vocal interface. Moreover, the guidelines recommend that users should be informed at the earliest time possible and, at the latest, at the time of processing. Furthermore, the guidelines also recommend that users would be informed of the purposes of processing personal data which should accord with their expectations of the device they purchase and, that when a controller seeks consent for various different purposes, they must provide a separate opt-in for each purpose to allow users to give specific consent for specific purposes.

More specifically, if voice messages are to be used to inform users according to Article 13 GDPR, data controllers should publish such messages on their website so they are accessible to all the parties involved.

Regarding the retention period, it depends on the specific purpose of processing. Generally, after a query has been answered or a command executed, the personal data should be deleted unless the VVA designer or developer has a valid legal basis to retain longer these data. Before considering anonymization as means for fulfilling the data storage limitation principle, VVA providers and developers should check the anonymization process renders the voice unidentifiable. When during the review process the VVA provider or developer detects a recording originated on a mistaken activation, the recording and all the associated data should be immediately deleted.

Finally, the guidelines recommend state-of-the-art authentication procedures to ensure the security of personal information processed, as well as ensuring that data subject rights can be properly facilitated.

The Guidelines emphasize the obligations of data controllers, mainly:

1. providers or designers should consider that when collecting user's voice, the recording might contain other individuals' voice or data such as background noise that is not necessary for the service. Whenever possible, voice assistants' designers should therefore consider technologies filtering the unnecessary data and ensuring that only the user voice is recorded.
2. Data controllers providing voice assistant services should ensure users can exercise their data subject rights using easy-to-follow voice commands.
3. Data controllers providing voice assistant services through screenless terminal devices must still inform users according to the GDPR when setting up the voice assistant or installing or using a voice assistant app for the first time.

5. Application of ethical and data management principles to Emilio project

Based on the above-mentioned principles and legislation, Emilio will follow the above listed general ethical principles:

- Respect for the integrity and dignity of persons: protecting them from being used for any other purpose than stipulated.
- Follow the “do no harm” principle: clearly communicating any potential risks to the elderly person involved.
- Acknowledge the rights of individuals to privacy, personal data protection and the freedom of movement.
- Honour the requirement of informed consent and continuous dialog with elderly constructively and transparently.
- Respect the principle of proportionality: not imposing more than is necessary on the subjects, nor going beyond stated objectives.

All subjects participating to the different phases of the project (codesign, rapid test, filed trial) will participate in the Emilio project on a fully voluntary basis, and they must not be misled in any situation.
Participants should be informed clearly that they at any time have the right to withdraw from their participation, and that any data that they have provided will be destroyed if they so request and that there will be no resultant adverse consequences.
All project partners and eventually external persons involved in the Emilio project must at all times take all the appropriate safety measures to ensure that the voluntary participants are not subject to any potential danger, physical harm or any wrong-doing as a result of their participation in the project.
The scope of the research and the organization of the interview/test/experimental condition should be explained at the participant at the beginning of each situation.

5.1.1. Informed consent

Informed consent is one of the founding principles of research ethics. It is one of the elements of the application Its intent is that human participants can enter research freely (voluntarily) with full information about what it means for them to take part, and that they give consent before they enter the research.
Consent should be obtained before the participant enters the research (prospectively), and there must be no undue influence on participants to consent. The minimum requirements for consent to be informed are that the participant understands what the research is and what they are consenting to.
The information given to the participant will be in understandable language to the participant and will include appropriate and adequate information (e.g. the nature, duration, and purpose of the experiment; the method and means by which it will be conducted; any potential inconveniences and hazards reasonably to be expected; the effects upon his/her health, and that he/she may quit the testing at any point) shall be given in order to ensure informed consent.
A model consent form in English has been attached to this document. This consent has been used

in Task 2.1 for the participants to interviews and co-design workshop where no personale and/or health related data are collected. In the following tasks that will involve participation the document will be adapted if necessary.

5.1.2. Data management

In the Emilio project we can identify two main kind of data:

1) Data collected directly form the participant to the different project activities
2) Data collected through the Emilio ICT solution

Table 4 contain the information on how data collected from users are managed in the project. The next paragraphs will describe the Data Management plan for data automatically collected.

DM Component	Description how main points are addressed
1. Data Summary	1) Semi structured interview with professionals: transcription of the interview 2) Semi structured interviews with resident: transcription of the interview 3) Workshops with future resident and relatives: aggregated report 4) Workshops with primary end users to evaluate services: aggregated report 5) Workshops with primary end users for the codesign of the interfaces: aggregated report 6) Interview with end users about the usability of the developed prototypes: transcription of the interviews and usability questionnaires with no personal data 7) Recruitment and evaluation questionnaires for the field trial in that include data on health state, well-being and quality of life.
2. Privacy	Data collected in activities from 1to 6 are completely anonymous and no personal data are collected Data collected in activity 7 will be pseudonymized assigning to each participant an internal code like the following EM-IT001.
3. Generation	Data will be generated through interview and questionnaires directly administered to the selected participants
4. Storing	The transcription of the interviews and of the questionnaires (anonymized or pseudonymized according to the source) will be stored in the Basecamp repository (secured with access restriction). The paper copies of the questionnaires (when applicable) will be conserved in the end user partners promises in a close storage according to good clinical practice.
5. Access	The access to anonymized or pseudonymized data in Basecamp is reserved to consortium partners. Access to the paper questionnaires is allowed to the partner that collected them and the

	authority for inspection as required according to national laws.
6. Processing	Transcription of the interviews could be analyzed using tools for qualitative analysis like MaxQda. Data from questionnaires will be analyzed with statistical software like SPSS or Stata.

Table 4: Data management plan for data directly collected from participants to activities

5.1.3. Data collection

Data collected within the Emilio project can be defined in three categories:

- Personal Data:
- Medical Data
- Smart home sensors

Personal attributes are those that can explicitly identify a person such as name, address and phone number and attributes that could potentially identify a person such as gender and birth date. Medical attributes includes sensitive medical information and sensor data, like blood pressure, medical history and presence sensors.

Real-time data: The collected data is used for the monitoring of the vitality, comfortability and safety of the customer. Real time data is collected via a virtual representation (digital twin) of the resident and the environment he/she is living in.

Time-series data: Machine learning and artificial intelligence based models are used for prediction and decision support systems to alert or propose adequate services to the resident, depending on his actual mood and health status within a specific context.
The model will be continuously improved by ongoing collection of monitoring data (time-series). This data is collected with the corresponding time line as batch data and will be stored in a "delta lake" for the training and testing of the different models.

5.1.4. Data storage and handling

The system architecture foresees the implementation of a multi-layered system.

Edge Layer: Close to the resident (on-prem installation) the Edge layer will collect the data from the different on-prem sensors (home automation, wearables, Magicview). And perform the services which must be prodided even without internet connectivity, such as fall detection and alerting. A local prediction system with limited AI-capability is required.
Sensor data will collected for monitoring and will be stored on the edge device for time based decision makings.

IoT Cloud Layer: The IoT cloud layer collects the realtime monitoring data and provides the accessibility to the different attributes and features of the virtual representation of the resident via secured (authenticated, authorized and encrypted) API access. Data forewarding to deiffernt applications can be granted with access control of very deep granularity (down to the attribute and feture level).

Machine Learning Lifecycle: Large amout of data sets are collected for the purpose of training and testing of the different models used for the decision support system. The execution of the models will be continuously monitored and improved. The models will be distributed to the cloud based decision support system (more abstracted decision making) and to the edge computing devices (using the software deployment capabilities of the device management system) for the execution of the vital decision support makings and alertings.

Busniness Support System: In the Business Support System (BSS) Customer data, subscription data, contractual data and billing data will be stored. The personal data attributes will be stored in the billing support system. This forms another abstraction of the personal data to the device data. In the Iot System and Machine Leraning lifecycle no personal data shall be stored. The digital twin is based on the "thing ID". All the model training and testing can be performed based on the thing-ID. Only in the BSS System a link to the real person with the address will be possible. The access the the BSS capability is controled by a rolebased Access Control System.

5.1.5. Security measures

The objective of the measures are basing on state.of-the art technologies. Wherever possible, the maximum level of data protection measures shall be used to ensure the best privacy and security for the customer. The following principles form the base for the implementation of the measures:

- Authentication using Standard Methods such as OAuth.
- Token based API requests to the collected customer data
- Role Based Access Control
- Anonymization of personal data (disclosure of personal data only where explicitly allowed)
- Encryption for all data in motion, data in use and data at rest

Device Connectivity

Device Authentication

Devices connect to protocol adapters in order to publish telemetry data or events. Downstream applications consuming this data often take particular actions based on the content of the messages. The device connectivity layer relies on protocol adapters to establish a device's identity before it is allowed to publish telemetry data or send events. Conceptually, the device connectivity layer distinguishes between two identities:

- an identity associated with the authentication credentials (termed the authentication identity or auth-id),
- an identity to act as the device identity or device-id.

A device therefore presents an auth-id as part of its credentials during the authentication process which is then resolved to a device identity by the protocol adapter on successful verification of the credentials.

In order to support the protocol adapters in the process of verifying credentials presented by a

device, the Credentials API provides means to look up secrets on record for the device and use this information to verify the credentials.

The Credentials API supports registration of multiple sets of credentials for each device. A set of credentials consists of an auth-id and some sort of secret information. The particular type of secret determines the kind of information kept. Based on this approach, a device may be authenticated using different types of secrets, e.g. a hashed password or certificates, depending on the capabilities of the device.
Once the protocol adapter has resolved the device-id for a device, it uses this identity when referring to the device in all subsequent API invocations, e.g. when forwarding telemetry messages downstream to the Messaging component of the device connectivity layer.
Every device connecting to the device connectivity layer needs to be registered in the scope of a single tenant. The Credentials APIs therefore requires a tenant identifier to be passed in to their operations. Consequently, the first step a protocol adapter needs to take when authenticating a device is determining the tenant that the device belongs to.
The means used by a device to indicate the tenant that it belongs to vary according to the type of credentials and authentication mechanism being used.

Transport Layer Encryption
The device connectivity layer uses encryption for all connectivity. This is accomplished by using industry standard TLS protocol for all TCP-based endpoints and DTLS for UDP-based endpoints.
Due to security best practices, the device connectivity layer does not allow potentially insecure protocols like SSL or TLS lower than version 1.2.

Certificates for the device connectivity layer
To secure the endpoints of the device connectivity layer, we use X.509 certificate Public Key Infrastructure.

The kind of certificate used differs by endpoint.

- Application and management endpoints
  For all application and management endpoints world-trusted certificates, issued by well-known certificate authorities, are used. This brings the advantage that most systems will be able to validate those certificates by default.

- Device endpoints
  - For the device endpoints and protocol adapters we use a more narrowed down certificate approach. Devices often do not have the resources to maintain and validate many multiple root CAs.
  - Our device certificates are therefore all issued by the Let's Encrypt Certificate Authority.

Digital Twin

A policy describes who has permission to work with a specific resource. The digital twin layer provides three types of policy:

- Policy of a thing

A specific policy provides someone (called **subject**), permission to READ and/or WRITE at a given **resource**. The resource can be defined as rough or as fine-grained as necessary for the respective use case: e.g. “thing:/” as top-level resource, or on sub-resources such as “thing:/features”. At runtime, the permissions are propagated down to all thing sub-entities.

- Solution policy
  The solution policy defines the access rules for your digital twin instance. The solution policy ID is defined automatically by our service at the time of subscribing the service, thus you will not be able to set a different ID (e.g. like you could do it for a thing entity).
- Namespace policy
  The namespace policy defines who is allowed to create things or policies in the corresponding namespace. With namespace policy you can restrict permissions to create things and policies specifically for each namespace.

Analytics / Machine Learning[6]

The new EU Data Protection regulations applying from 2018 onwards will make privacy aware machine learning necessary. Consequently, issues of privacy, security, safety and data protection move more and more into the focus of AI and ML, thereby fostering an integrated ML approach, which emphasizes the importance of the human-in-the-loop.

The new EU General Data Protection Regulation (GDPR) defines personal data as “any information relating to an identified or identifiable natural person” and specifically acknowledges that this includes both ‘direct’ and ‘indirect’ identification. While so far there is not one privacy definition yet that is able to encompass all the different aspects of privacy, there are guidelines that list the possible identifiers that could be used to identify a person from a group:

- Names, Geographical subdivisions smaller than a state, Dates (other than year)
- Phone & Fax Numbers
- Electronic mail addresses
- Social Security, Medical Record & Health plan beneficiary numbers
- Account & Certificate/license numbers
- Vehicle identifiers and serial numbers (including licenseplate numbers)
- Device identifiers and serial numbers
- Web Uniform Resource Locators (URLs) & Internet Protocol (IP) address numbers
- Biometric identifiers, including finger, retinal and voiceprints
- Full face photographic images and any comparable images
- Any other unique identifying number, characteristic or code

The following methodes are possibilities for the preservation of privacy in machine-learning. The implementation of the methods will be developed along the Emilio Project and will depend on the

[6] Source: A Deep Learning Approach for Privacy Preservation in Assisted Living~~ing~~ I. Psychoula et al., "A Deep Learning Approach for Privacy Preservation in Assisted Living," 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), 2018, pp. 710-715, doi:

models which will be used for the decision support system.

Privacy Preservation with Anonymization Methods
Methods most commonly used for privacy preservation:

- k-Anonymity: achieved by suppressing (deleting an attribute value from the data and replacing it with a random value that matches any possible attribute value) or generalizing the attributes in the data
- l-diversity: anonymization conditions are satisfied if, for each group of records sharing a combination of key attributes, there are at least l -"well-represented" values for each confidential attribute
- t-closeness: requires the distribution of the sensitive attributes in an equivalent class to be close to the distribution of the attribute in the overall table, which in turn means that the distance between the two distributions should be no more than a specified threshold t

Privacy Preservation with Deep Learning
Usually, deep learning architectures are constructed as multi-layer neural networks.The existing literature on privacy protection mostly focuses on traditional privacy preserving methods, as described in the previous section, and not on deep learning.

Differential privacy is one of the few approaches of privacy protection that makes use of machine learning methods. Applications of Differential Privacy include boosting, principal component analysis, linear and logistic regression support vector machines, risk minimization and continuous data processing.

LSTM Encoder-Decoder Model: A proposed method for the preservation of the privacy with deep learning is based on the Encoder-Decoder system. In this case the model uses he multi- layered Long Short-Term Memory (LSTM) encoder to map the input sequence to a vector of a fixed dimensionality, and then another LSTM is used to decode the target sequence from the vector.
So the data from each entry of a collected dataset can be fully disclosed to the receiver, generalized or deleted. Each value for a given attribute has different range aligned with real life values. Separate views are created for different receivers, such as family member, doctor, caregiver, and researcher. Each receiver has a different decoder output due to their privacy clearances on resident information. The datasets can only be displayed with the use of the corresponding decoder.

5.1.6. Integration of 3rd party Services

We can only take over the chain responsibility if the 3rd party prodvider accepts our privacy terms and conditions in an auditable way.
If this is not possible, the integration has to be handled case by case, e.g. separate service contracts have to be agreed directly between the 3rd party providers and the Emilio customers.

6. Bibliography

Gillon, R (1994). "Medical ethics: four principles plus attention to scope". British Medical Journal. 309 (184): 184–188. doi:10.1136/bmj.309.6948.184. PMC 2540719. PMID 8044100.

Beauchamp TL, Childress JF. Principles of bioethics. 7th ed. Oxford University Press; 2013.

World Medical Association (2013). "Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects". JAMA. 310 (20): 2191–2194. doi:10.1001/jama.2013.281053. PMID 24141714.

Porcari, A., Borsella, E., Mantovani, E., Contributors, Stahl, B., Flick, C., Ladikas, M., Hahn, J., Brem, A., Yaghmaei, E., Søraker, J. H., Barnett, S. J., Schroeder, D., Chatfield, K., Ikonen, V., Leikas, J., & Obach, M. (2016). Responsible-Industry: A Framework for implementing responsible research and innovation in ICT for an ageing society.

Guidelines on virtual voice assistants Version 2.0 Adopted on 7 July 2021

WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust, Brussels, 19.2.2020 COM(2020) 65 final

Ethics and governance of artificial intelligence for health: WHO guidance
ISBN 978-92-4-002920-0 (electronic version) ISBN 978-92-4-002921-7 (print version)
© World Health Organization 2021

Psychoula et al., "A Deep Learning Approach for Privacy Preservation in Assisted Living," 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), 2018, pp. 710-715, doi: 10.1109/PERCOMW.2018.8480247.

7. Annex

7.1 Informed consensus model

INFORMED CONSENT FORM

TITLE OF STUDY: Emilio
NAME OF PROJECT COORDINATOR:
NAME OF INTERVIEWER: _____

I I, the undersigned (name and surname) _____
Age _____ gender M ☐ F ☐ date of birth ___/___/_____
Adrress _____n. _____ Post codeP _____
City _____ tel. _____

I declare

- To participate voluntarily in the interview aimed at defining the functional requirements of the system developed within the Emilio project.
  The Emilio project
  The EMILIO project intends to develop a series of services aimed at strengthening the self-sufficiency and counteracting the social isolation of self-sufficient elderly persons living in an assisted living facility. Through a web platform and a voice interface it will offer services to improve residents' comfort, safety and autonomy such as: access to telemedicine service, medication adherence, automatic health assessment and control of home systems. Web services are provided in an intuitive way, suitable for use by the target audience. To this end, an Internet of Things-based technological infrastructure will be implemented in the facility that observes the daily activities of the client and interacts vocally with the client when necessary.
- Have received from the interviewer/researcher mentioned above all clear and comprehensive information about the purpose and procedures of the interview in which I have been asked to take part.
- Having had the opportunity to ask clarifying questions and having had satisfactory answers, as well as having had the opportunity to enquire about the details of the study with a person I trust.
- To be aware
  - that my data may be examined or used for research purposes, but will remain strictly confidential in accordance with current legislation and subsequent amendments and additions;
  - that my data will be used in aggregate form, for the preparation of a final report for the Health Authorities or for a publication, whatever the outcome of the study, always respecting the confidentiality and anonymity of my identity (Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016, art. 13 of Legislative Decree no. 196/03 in force since 1 January 2004);
  - that I must sign two identical forms of this informed consent: one original will be retained by the interviewer/researcher (and kept for at least 15 years) and the second will be given to me;
  - that in case of any problems or for any further information I should contact:

Name and Surname of person in charge
Address
Telephone number

I therefore freely agree to take part in the interview.

The signature on this form will not affect my legal rights.

Read and approved (handwritten) _____