Alliance for IoT and Edge Computing Innovation

# White Paper
# AI for better health

# Release 1.0

# AIOTI WG Health

# 12 December 2022

# Executive Summary

This white paper introduces the AIOTI WG Health vision and contribution to the European discussions on AI strategies applied to the health domain.

The main aim of this paper is to contribute to the identification and addressing of any issues that might be hindering the wider adoption of AI technologies in the healthcare sector based on AIOTI WG Health's members' best practices.

# Table of Content

Executive Summary ........ 2

Table of Figures ........ 4

1. Introduction ........ 5

2. Artificial Intelligence for health and care in the EU ........ 6

3. COVID-19's impact and the push towards AI implementation strategies ........ 7

4. AI, challenges and opportunities; the vision for a future better health ........ 9

5. AI Ethics and risk management ........ 12

6. AI Data analysis, privacy, data ownership and data access ........ 14

7. AI Standardization ........ 16

8. AI impact assurance for sustainability ........ 19

9. Best practices for better health deployed by AIOTI members ........ 21

10. Conclusions ........ 29

List of Contributors ........ 30

Acknowledgements ........ 31

About AIOTI ........ 32

# Table of Figures

Figure 1  Examples of AI applications at different stages of the COVID-19 crisis[1] .....................................................7

Figure 2 Impact assurance key steps ..........................................................................................................................20

 

# 1. Introduction

Nowadays computers are supporting human input, decision making and provision of data. In today's healthcare sector and medical profession, AI, algorithms, robotics and big data are used to derive inferences for monitoring long run medical trends, detecting and measuring individual associated risks and opportunities based on data-driven estimations. The healthcare sector highly depends on data and analytics to improve therapies, practices, services personalization and adherence. Due to the massive deployment of IoT (Internet of Things) and smart technologies (like smart wearable devices capable of regularly monitoring our health parameters), in recent years, there has been tremendous growth in the range of medical data collected, including clinical, genetic, behavioural and environmental data. In particular, the global healthcare IT market is projected to reach USD 821.1 billion by 2026 from USD 326.1 billion in 2021, at a CAGR of 20.3% during the forecast period. The growth in this market is mainly driven by government mandates & support for healthcare IT solutions; rising use of big data in healthcare; high returns on investment associated with healthcare IT solutions; the need to curtail escalating healthcare costs; the growing demand for and use of HCIT solutions due to COVID-19; and the growing mHealth, telehealth, and remote patient monitoring markets[1].

These include electronic health records (EHRs), genome sequencing machines, high-resolution medical imaging, smartphone applications and ubiquitous sensing. Through machine learning algorithms and unprecedented data storage and computational power, AI technologies have most advanced abilities to gain information, process it and give a well-defined output to the end-user. Daily monitoring thereby aids to create big data to recognize behavioural patterns' relation to health status in order to create predictions with highest mathematical precision based on big data capturing large-scale samples. AI thereby enlightens to analyse the relation between prevention and treatment and patient outcomes in all stages of diagnosis, treatment, drug development and monitoring, personalized medicine, patient control and care. Advanced hospitals are looking into AI solutions to support and perform operational initiatives that increase precision and cost effectiveness. Robotics have been used for disabled and patient care assistance. Medical decision making has been supported through predictive analytics and general healthcare management technology. Network connectivity allows access to affordable healthcare around the globe in a cost-effective way.

This white paper gets the goal to provide the main perspective, experiences and know-how from the AIOTI WG Health members on specific key issues, opportunities and open points because of the AI implementation in the healthcare domain.

[1] Healthcare IT Market by Products & Services (Healthcare Provider Solutions, Healthcare Payer Solutions, & HCIT Outsourcing Services), Components (Services, Software,Hardware), End-User, and Region (2022 - 2026)

## 2. Artificial Intelligence for health and care in the EU

Despite a number of initiatives undertaken by the EU in the last few years towards advancing the development and uptake of AI technologies to help EU citizens better monitor their health, receive better diagnoses and more personalized treatments, as well as live a healthier and more independent life, current situation in the EU indicates that healthcare organisations are slow in implementing AI technologies in healthcare and that the level of adoption is low overall.

Based on evidence gathered[1], while most EU MS that have developed AI strategies identify healthcare as a priority sector, there are no policies within those strategies targeting healthcare in particular. The lack of trust in AI-driven decision support is hindering the wider adoption, while issues around integrating new technologies into current practice are also prominent challenges identified by relevant stakeholders in EU Member States (MS).

AI-focused strategies and policy instruments include initiatives aiming to build the EU's adoption of AI across industries. As concern healthcare, most Member States have identified healthcare as one of the key sectors that needs to be prioritized in the context of AI.

In national AI strategies healthcare is included in various modalities with particular attention to topics such as data, ethics, research and innovation and public services.

The AI instruments development is strictly intertwined with the availability of a relevant amount of data as well as their quality.

On top of this, making the use of AI ethical and transparent is of great relevance to the healthcare sector, with the notion of trust emerging as a significant component of this. AI-powered methods and processes used within the sector should be trusted by society.

Steps towards the EU's vision of a more collaborative and unified Union that fully leverages the opportunities AI presents in tackling the big issues facing healthcare have also been undertaken and launched under European Commission umbrella projects, such as Horizon 2020 and include the 'AI-on-demand platform' (AI4EU[2]) and the network of European AI research excellence centres. The AI4EU platform, will facilitate the creation of a European AI ecosystem and will be an access point to all resources required to engage with AI. The network will focus on the mobilization of AI scientists to produce high-quality research and will promote cooperation between academics and industry.

A review(1) of the AI-promoting initiatives from EU Member States reveals that there are no specific initiatives within these strategies targeting the healthcare sector in particular thus revealing the need for the development of an effective regulatory framework.

[1] Study on eHealth, Interoperability of Health Data and Artificial Intelligence for Health and Care in the European Union, Lot 2: Artificial Intelligence for health and care in the EU Final Study Report, European PwC, Luxembourg 2021

[2] https://www.ai4europe.eu/

# 3. COVID-19's impact and the push towards AI implementation strategies

The use of advanced technologies, especially predictive computing in the health sector, is on the rise in this era, and they have successfully transformed the sector with quality insights, better decision-making, and quality policies. Even though notable benefits have been achieved through the uptake of the technologies, adoption is still slow, as most of them are still new, hence facing some hurdles in their applications especially in national and international policy levels. But the recent case of COVID-19 outbreak has given an opportunity to showcase that these technologies, especially artificial intelligence (AI), have the capacity to produce accurate, real-time, and reliable predictions on issues as serious as pandemic outbreak.

Due to the COVID-19 pandemic, AI tools and technologies are employed to support efforts of policy makers, the medical community, and society at large to manage every stage of the crisis and its aftermath: detection, prevention, response recovery and to accelerate research.

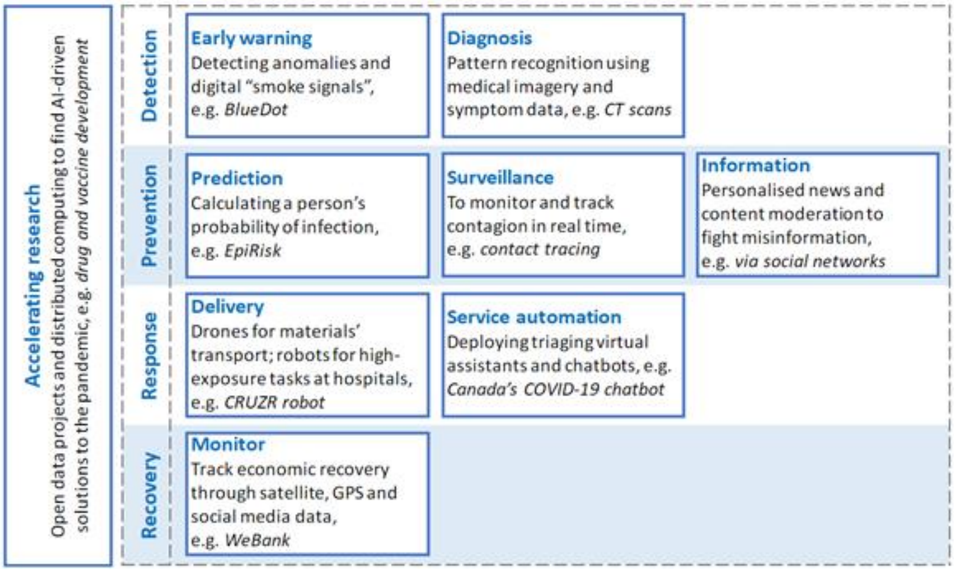


**Figure 1 Examples of AI applications at different stages of the COVID-19 crisis**[1]

AI tools and techniques can help policymakers and the medical community understand the COVID-19 virus and accelerate research on treatments by rapidly analysing large volumes of research data. AI text and data mining tools can uncover the virus' history, transmission, and diagnostics, management measures, and lessons from previous epidemics.

- Deep learning models have been implemented to streamline the process of discovering new medicines that might treat COVID-19. Several institutions are using AI to identify treatments and develop further updated vaccines.
- Access to datasets in epidemiology, bioinformatics and molecular modelling is being provided, e.g., through the European Centre for Disease Prevention and Control.
- Innovative approaches including prizes, open-source collaborations, and hackathons, are helping accelerate research on AI-driven solutions to the pandemic. For example, the United Kingdom's "CoronaHack – AI vs. Covid-19" seeks ideas from businesses, data scientists and biomedical researchers on using AI to control and manage the pandemic.

AI has been also employed to help detect, diagnose and prevent the spread of the virus. Algorithms that identify patterns and anomalies are working to detect and predict the spread of COVID-19, while image recognition systems are speeding up medical diagnosis. For example:

- AI-powered early warning system scan help detect epidemiological patterns by mining mainstream news, online content and other information channels in multiple languages to provide early warnings, which can complement syndromic surveillance and other healthcare networks and data flows (e.g. WHO Early Warning System, EWAS[2]).

- AI tools are helping in identifying virus transmission chains and monitor broader economic impacts. In several cases, AI technologies have demonstrated their potential to infer epidemiological data more rapidly than traditional reporting of health data. Institutions such as Johns Hopkins University and the OECD[3] (oecd.ai) have also made available interactive dashboards that track the virus' spread through live news and real-time data on confirmed coronavirus cases, recoveries, and deaths.

- Rapid diagnosis is key to limit contagion and understand the disease spread[4]. Applied to images and symptom data, AI could help to rapidly diagnose COVID-19 cases.

Conversational and interactive AI systems help respond to the health crisis through personalised information, advice and treatment, and learning.

- To fight misinformation – the COVID-19 "infodemic" – social networks and search engines are using personalised AI information and tools and relying on algorithms to find and remove problematic material on their platforms.

- Virtual assistants and chatbots have been deployed to support healthcare organisations, for example in Canada, France, Finland, Italy, the United States and by the American Red Cross. These tools help to triage people depending on the presence of symptoms.

- Identifying, finding and contacting vulnerable, high-risk, individuals.

---

[1] OECD, Using artificial intelligence to help combat COVID-19, April 2020

[2] WHO, WHO Early Warning System, https://www.who.int/emergencies/surveillance/early-warning-alert-and-response-system ewars#:~:text=WHO's%20Early%20Warning%2C%20Alert%20and,up%20a%20disease%20surveillance%20system. 14/07/2022

[3] https://oecd.ai/en/ 14/07/2022

[4] Shigao Huang, Jie Yang, Simon Fong, Qi Zhao, "Artificial intelligence in the diagnosis of COVID-19: challenges and perspectives", Int J Biol Sci. 2021; 17(6): 1581–1587. Published online 2021 Apr 10. doi: 10.7150/ijbs.58855, PMCID: PMC8071762

# 4. AI, challenges and opportunities; the vision for a future better health

Software has become significantly smarter in recent years.

The current expansion of AI is the result of advances in a field known as machine learning. Machine learning involves using algorithms that allow computers to learn on their own by looking through data and performing tasks based on examples, rather than by relying on explicit programming by a human[2].

A machine-learning technique called deep learning, inspired by biological neural networks, finds and remembers patterns in large volumes of data. Deep-learning systems perform tasks by considering examples, generally without being programmed, and out-perform traditional machine-learning algorithms.

Big Data, referring to extremely large data sets that can be analysed computationally to reveal patterns, trends and associations, together with the power of AI and high-performance computing, are generating new forms of information and insight with tremendous value for tackling humanity's greatest challenges.

Multiple are the possible applications, ranging from the agriculture to the healthcare domain. But due to the massive impact that these technologies can have in our society and daily life, it is so important to bear in mind which are the AI's associated challenges and opportunities in order to be ready to match and adequately address them. Below the main ones in terms of challenges and opportunities are reported.

**Challenges:**

1. **Explainable AI**: one of the main challenges of AI in health is to explain its decision to the physicians. This aspect is critical to increase the confidence in these technologies. Physicians make the final decision but their confidence on AI suggestions is crucial to combine Human and Artificial Intelligences. In this aspect, AI is going to have a relevant role in detecting cues or symptoms and in providing a support to physicians' proper treatment identification; AI is complementary to physicians' knowledge.

2. **Unbiased decisions**: AI bias is an important limitation in order to use AI systems in health. This bias limits the applicability of these systems in different contexts (different countries or social/economical scenarios, etc.).

3. **Frugality**: the performance of AI systems depends strongly on the quality and quantity of data used to train them. There are many health applications where it is not possible to work with all the necessary data. In these cases, it is necessary to consider other strategies like transfer learning, cross-corpus, cross-task, data augmentation, etc.

[2] https://www.wired.com/story/guide-artificial-intelligence/

 

**Opportunities:**

Thinking on future applications, we could consider:

**1. To study, analyse and model the use of health resources and processe**s in order to propose improvements that allow their efficient use.

- Use cases in Hospitals:
  - Prediction or prognosis of the evolution of a disease in hospitalized patients to manage short-term care needs.
  - Screening/Screening of patients in medical emergencies.
  - Management of blood donation campaigns.
  - Identification of high-efficiency surgical equipment for study of characteristics
  - Study of the effectiveness of radiological tests in the diagnosis (i.e. comparison of presumed diagnosis with final diagnosis).
- Use cases in Health Care Centers:
  - Study and modeling of the evolution and spread flows of contagious diseases to predict short-term resource needs.
  - Management of care for the elderly in primary care, segmentation, needs analysis.
  - Citation optimization for population segments based on their characteristics.
  - Prediction of need for hospitalization of patients with asthma, for prevention.
  - Management optimization of home care units
  - Analysis of nursing activity by centers, for comparison.

**2. To study of the incidence and the factors** that most influence a disease in order to prevent its appearance and improve early detection. Some examples would be:

- Segmentation for the design of prevention campaigns by area incidence of pathologies or type of population: alcohol, smoking, sleep disorders, weight, anxiety and depression, frailty, etc.). The development of very focused programs allows a great impact with a smaller investment.
- Detection of situations of vulnerability: suicide, gender violence, bullying, etc.
- Analysis of incidence of lung and colorectal cancer and cardiovascular disease by age ranges and health zones, combined with data from external context (e.g. environmental, towards epigenetics) and socioeconomic
- Integration of different tests to obtain biomarkers for the early diagnosis of diseases, according to the characteristics of the patient.
- Analysis of the evolution of mental health problems in the medical records of young people for the design of prevention strategies.

**3. To study the characteristics of chronic and elderly patients** to define patterns that allow improving care for these groups. Use cases:

- Analysis of comorbidities in chronic patients (temporary analysis, relationship with habits).
- Monitoring of therapeutic compliance of patients.
- Follow-up of the degree of independence to predict future care needs
- Predicting the need for hospitalization.

**4. Management and consumption of medical material** in order to propose optimization strategies for both the supply and management of this material. Use cases:

- Studies on pharmacy needs georeferenced and segmented by population characteristics to detect anomalies and the proposal of optimization algorithms
- Analysis of dispensations versus prescriptions, incidence of self-medication.

**5. To analyze the medical treatments that are being applied in the different health areas** and compare them with the evolution of the patients according to their characteristics. Use cases:

- Adaptation of treatments to the characteristics of the patient: à la carte care
- Search for more effective treatments.
- Search for interaction between drugs in specific pathologies.

# 5. AI Ethics and risk management

Artificial Intelligence and health can be really challenging depending on the application and the role of the AI in the healing process.

There is plenty of room for AI solutions oriented to clinical management, clinical costs improvements, processes optimization and it may not run into more challenges that AI applied to other sectors. Data Life Cycle, Origin of data, trust, explainability.

On the other hand, when AI is applied to the diagnostics, of decision-making processes, we run into more challenging scenarios.

In the clinical decision process, there are many information inputs that in other scenarios AI models will not work with as it can be Gender, Age, Ethics, and many other information that can be eliminated from not health purposes AI to avoid bias but that it has to be taken into account for medical purposes and opens the door to new challenges.

So far, we must consider two main health related domains were AI Ethics and risk management needs to have two different approaches that are going to be further treated below:

- Health environment related AI solutions.
- Medical decision-oriented solutions. (medical device)

**Health environment related solutions**

From the Ethic Perspective, the process to manage the Ethics of these AI models and solutions my not differ a lot from the other sector solutions. As the process may be alike, the development teams need to implement the sector requirements as it can be:

- data management for sensible data and other regulations that will affect the data usage of the health data (e.g., Data Act)
- Cybersecurity requirements,
- HIS systems integration requirements (See IEC 80001-1:2021 for further information)

From risk management point of view, the Impact of errors, (E.g., data corruption) within a health institution can be harmful to patients. Therefore, a risk management system needs to be set in place for any kind of development that is intended to be used in a health organisation or work with health organisation data.

**Medical decision-oriented solutions (Medical Device)**

In what refers to risk, Medical Device regulations define a strict Risk Management process that needs to be implemented in order to manage any potential hazard or risk that it may appear. This is properly handled and AI potential risk needs to be managed through this process.

As an important remark, currently there is a need from some AI systems to be trained from the data from the hospital to provide best results. That system does not perform equally well in other health institutions if it is not retrained with hospital data. This retraining and the different performance of some AI based systems depends on training and that can confront with medical device regulation.

From the Ethics perspective, there are mayor challenges that need to be addressed.

First, we can refer to patient decision of their own health data and how information can be shared. The personalised medicine is to come and in base for it to have warranties to succeed the patient need to understand the decisions made, as well as what is his data needed for. This scenario is quite challenging for elderly populations as well as less educated people. Those collectives that might have more cognitive challenges need to be able to understand the whole process and procedures.

Also, the physicians will have to understand the new tools they will have on their hands. As tool they are they need to provide appropriate use of them and be prepared (formation) to use them properly so they don't harm the patient, or the physician (reputation). The human decisions need to be protected and backed up for AI systems.

Also, there are some principles of medical device development that can be challenged by the AI models as it can be the auditory principle. All systems need to be auditable to the point of finding who has performed and specific action in a specific moment and the data the physician has used to take the decision. With the AI and the explainability "restrictions", a challenge may arise with this situation.

According to what just aforementioned, there is a need to define two different environments to AI in Health:

- Health institutions management
- Health management.

For Health institutions management:

- Ethics can me managed in the same approach as any other sector but taking into account the sensitive data requirements.
- Risk management is closer to medical device Risk management than to regular environments. Even though it works with data, alteration of that data because of AI failure can end in harm to patients. There are international standards that can be used as orientation

For health management:

- Risk management is handled within the regulations of Medical Devices Manufacturing. Software as a Medical Device Must follow the required processes.

Ethics are much more complicated in this case mainly because the still is an evident digital illiteracy in our society. Elderly person, which is to be called to be increasing in the future, is not completely capable of assuming the new technologies at the speed they are coming to them. This added to the need to co-creation of value between physicians and patients to involve them in their own health care challenges some AI models to be used in Health. Personalised medicine requires patients' involvement and understanding and for some collectives, that represents a challenge that can be generalized to a sector challenge itself.

# 6. AI Data analysis, privacy, data ownership and data access

Health services are evolving from being provided by specialists in several sectors (clinicians but also phycologists, nutritionists, physiotherapists, etc.) to adopt a patient-based approach where traditionally unconnected professionals and non-professional caregivers are called to collaborate. The use of data and artificial intelligent techniques could provide enormous advantage in view of coordinating efforts and enhancing the efficiency of health and caregiving processes in an integral perspective.

The use of personal data is especially relevant and tremendously sensitive when talking about health. Data privacy and data security are clue elements that interface with property rights (ownership of data) and ethical issues. The General Data Protection Regulation (GDPR) was put into effect on May 25, 2018, by the European Union. It imposes obligations onto any type of organisation that target, collect, or make use of personal data, and from different perspectives, it has been said to be the toughest privacy and security law in the world. The GDPR has derived into National and regional legislations all along the EU.

The GDPR defines 'personal data' as information relating to an identified or identifiable natural person, that is, information that can be identified, directly or indirectly, 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 (art. 2(a) GDPR). That opens a double option: a) any entity handling personal data, as for example, data from Electronic Health Records (EHR), will need to define special considerations for impact assessment; on the contrary, b) any data that cannot be related with a natural person is not considered personal data and therefore it is out of GDPR consideration.

The GDPR introduces the concept of Data Protection Impact Assessment (DPIA). The DPIA is a process designed to describe the processing, assess its necessity and proportionality, and help manage the risks to the rights and freedoms of natural persons resulting from the processing of personal data by assessing them and determining the measures to address them. DPIAs are important tools for accountability, as they help controllers not only to comply with the requirements of the GDPR, but also to demonstrate that appropriate measures have been taken to ensure compliance with regulation. Some important chapters for DPIA are:

- Definition of Data Governance model where controllers are responsible of making decisions about collecting, storing or processing personal data. National bodies regulate their role and enrolment as 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
- Data Security: Relevant EU standards, e.g., ISO/IEC 27001 and 27002 (Code of practice for information security management), to ensure confidentiality, integrity, and availability. It will additionally include the Directive on security of network and information systems ('Cybersecurity directive', NIS-Directive 2016/1148) on the security of critical infrastructures and the ePrivacy Directive 2002/58, as well as European Union Agency for Network and Information Security (~) guidance
- Data ownership: According to the Health Insurance Portability and Accountability Act of 1996 (HIPAA), the original physical medical record is the property of the physician's office that generated it. However, the data on the medical records are the property of the patients themselves. Therefore, any further processing of medical records requires consent forms

- Informed consent procedures should be facilitated before any processing of personal information. These forms should explain the purpose the data will be used for, the means and time horizon in which the information is going to be handled and (if any) the results that are expected to be obtained from the data processing

Any entity handling personal data and in particular health service providers will need well defined (and preferably open) Data Protection Impact Assessment plans to guarantee the necessary tools and mechanisms to promote suitable and secure data handling procedures. The use of Electronic Health Records (EHR), for example, as basis for any further data analysis would be susceptible of DPIA.

The secondary use of data for research can be facilitated by several anonymisation methods. Researchers should document data appropriately explaining the procedures and fieldwork methods, the objectives and methodology of the research, and explicitly describing the meanings of variables and codes used. Additionally, they should describe any derivation, transformations, de-identification or data cleaning carried out. The most commonly used de-identification methods are:

- Pseudo-anonymization of data: the entity monitoring the data codifies the original records before sending it to a third party. The results or conclusions obtained by the third party could be sent back to the former (entity monitoring the data), and this one would be able to relate the conclusions with the individual
- Anonymize the data: in this context that would imply removing direct identifiers (names, address, etc.) or aggregating the information providing measurements at building level. Both options imply losing information

Researchers should also ensure that data are held in an organised manner. Documentation is invaluable in enabling secondary users to contextualise data and conduct better, informed re-use of the material. Any consent and confidentiality concerns that may inhibit archiving data should be resolved before any secondary use of data.

European research programs funded by the European Commission (FP7, H2020, Horizon Europe, etc.) are evolving in terms of facilitating access to research data and research result for secondary access, and therefore applying more and more ambitious open science principles. The current Horizon Europe requires immediate open access to research data through trusted repositories and Data Management Plans that should be regularly updates according to FAIR principles (making data Findable-Accessible-Interoperable-Reusable). The ultimate objective is to:

- encourage collaboration and replication avoiding duplication of efforts
- involve citizens and society improving transparency and public participation
- ensure quality of data processing at the monitoring and evaluation as well as other analytical work performed
- build on previous results and experiences
- speed up the innovation uptake and therefore, facilitate faster and greater development of markets

This situation defines a fragile equilibrium, even more sensitive when it refers to health related data, where research (or research projects) should assess and provide the security measures to manage the risks to the rights and freedoms of natural persons resulting from any processing of personal data. The application of Open Data principles should be in line with an open and clear specification of DPIA, and both processes should be well aligned with corresponding Ethic Boards that will always ensure the necessity and proportionality to make use of personal data and to help manage any related risk.

# 7. AI Standardization

International standards – the technical specifications and requirements needed for AI and other technologies to perform well – can help address real and perceived risks by setting clear boundaries and making machine learning (ML) predictable, reliable and efficient.

**Explainable AI challenges** (mentioned in the section before) can be partially solve using AI technologies, more precisely, knowledge representation techniques (ontologies, knowledge graphs, etc.).

In particular, the focus has been put towards standards for AI and Health. Among them, specific attention has been devoted to ITU/WHO Focus Group on artificial intelligence for health (FG-AI4H), ISO (TC215 Health Informatics, and SC 42 Artificial Intelligence), ETSI SmartM2M SAREF4EHAW for eHealth/Ageing-well. Specifically, the following:

**ITU/WHO Focus Group on artificial intelligence for health (FG-AI4H)**

ITU/WHO Focus Group on Artificial Intelligence for Health (FG-AI4H) has a partnership with the World Health Organization (WHO) to standardize AI-based assessment framework and evaluation for health, diagnosis, triage or treatment decisions.

**ISO**

**ISO TC215 Health Informatics**

ISO TC215 Health Informatics (https://www.iso.org/committee/54960.html)'s scope: Standardization in the field of health informatics, to facilitate capture, interchange and use of health-related data, information, and knowledge to support and enable all aspects of the health system

**ISO SC42 Artificial Intelligence**

AI enabled Health Informatics: ISO SC42 AI created a joint working group "AI enabled Health Informatics" that will provide a landscape survey, and a set of recommendations for future work to the impact on ISO/TC 215 standards. The group comprises experts from ISO/IEC/JTC 1/SC 42, IEEE, and the ITU/WHO AI4Health focus group.

**JTC 1/SC 42 | ISO/IEC AWI 5392 Information technology** - Artificial intelligence - Reference architecture of knowledge engineering (under development) defines a reference architecture of Knowledge Engineering (KE) in Artificial Intelligence (AI). The reference architecture describes KE roles, activities, constructional layers, components and their relationships among themselves and other systems from systemic user and functional views. This document also provides a common KE vocabulary by defining KE terms.

**JTC 1/SC 42 | ISO/IEC AWI 5339 Information Technology** - Artificial Intelligence - Guidelines for AI applications provides a set of guidelines for identifying the context, opportunities, and processes for developing and applying AI applications. It can be used by ISO, IEC, and JTC1 Technical Committees and Sub-Committees to build on this work in developing standards for AI applications in their areas of interest. The guidelines provide a macro level view of the AI application context, the stakeholders and their roles, relationship to the life cycle of the system, and common AI application characteristics. The guidelines will reference but not duplicate or overlap other AI-related standards to build details.

### ETSI SmartM2M SAREF4EHAW for eHealth/Ageing-well

**ETSI SmartM2M SAREF4EHAW** for eHealth/Ageing-well aims to cover the following use cases:

1) Elderly at home monitoring and support, 2) Monitoring and support of healthy lifestyle for citizens, 3) Early Warning System (EWS) and Cardiovascular Accidents detection.

The use cases are classified into the following categories: 1) Daily Activity Monitoring, 2) Integrated care for older adults under chronic conditions, 3) Monitoring assisted persons outside home and controlling risky situations, 4) Emergency trigger, 5) Exercise promotion for fall prevention and physical activeness, 6) Cognitive simulation for mental decline prevention, 7) Prevention of social isolation, 8) Comfort and safety at home, and 9) Support for transportation and mobility.

SAREF deliverables reviewed standards (IEEE, ETSI, SNOMED International, OneM2M), Alliances (AIOTI), IoT Platforms, and European projects and initiatives, etc.

SAREF4EHAW investigated the following ontologies:

- WSNs/measurement ontologies: OGC (Open Geospatial Consortium) Observations & Measurements (O&M), Sensor Model Language (SensorML), Semantic Sensor Web (SWE): W3C & OGC SOSA (Sensing, Observation, Sampling and Actuation) and W3C SSN (Semantic Sensor Network).
- NASA QUDT (Quantities, Units, Dimensions and Types).
- eHealth/Ageing-well domain main ontologies:
  1) ISO/IEEE 11073 Personal Health Device (PHD) standards,
  2) ETSI SmartBAN Reference Data Model and associated modular ontologies,
  3) ETSI SmartM2M SAREF4EHAW,
  4) FHIR RDF (Resource Description Framework),
  5) FIESTA-IoT Ontology to support the federation of testbeds,
  6) Bluetooth® LE profiles for medical devices proposed by zontinua, MIMU-Wear (Magnetic and Inertial Measurement Units) ontology, and
  7) Active and Healthy Ageing (AHA) platform wearables' device ontology.
  8) SAREF has been mapped with oneM2M base ontology in 2017.

SAREF ETSI TR 103 509[3] defines 43 final ontological requirements, and 59 additional service level assumptions of the eHealth/Ageing-well domain (use cases included) are presented. For instance, the ontology will describe concepts to describe ECG devices.

ETSI TS 103 410-8 V1.1.1 (2020-07[4]) SmartM2M; Extension to SAREF; Part 8: eHealth/Ageing-well Domain.

[3] SAREF4EHAW2020
[4] SAREF4EHAW2020

**W3C**

**W3C Semantic Web in Health care and Life Sciences Community (HCLS)**

Develops, advocates for, and supports the use of Semantic Web technologies across health care, life sciences, clinical research and translational medicine. These domains stand to gain tremendous benefit from intra- and inter-domain application of Semantic Web technologies as they depend on the interoperability of information from many disciplines.

HCLS CG focuses on the use of Semantic Web technologies to realize specific use cases which themselves have a specific clinical, research of business values. The CG may also develop ongoing and mutually productive liaisons with relevant external organizations in healthcare, life sciences, and clinical research, including organizations that are actively working on relevant standards and/or implementations to which the HCLS's work might contribute.

## Recommendations

- Encourage more open standards and FAIR principles within standards
- Encouraging more synergies among standardization bodies such as ISO, IEC, W3C, NIST, ETSI, OneM2M, etc.
- Encouraging more experts to easily contribute to standards just by sharing their expertise (initiatives such as StandICT.eu 2023 and HSBooster are working towards this goal).

 

# 8. AI impact assurance for sustainability

Hippocrates stated that medicine has to be based on detailed observation, reason and experience in order to establish a diagnosis, prognosis and treatment. In particular, he developed the six sense methodology aimed to train doctors. Thanks to this new empirical method, physicians of ancient times began to use their intellect and five senses in order to gather information about their patients.

Nowadays, doctors keep using the same methodology described by Hippocrates over two thousand years ago, but they also have access to new, vast, and valuable sources of data not coming just from experience but from additional tools. In fact, laboratory tests, imaging, DNA sequencing, molecular pathology, and the technological advances of hyper-connectivity, allow for the analysis of new individual health features. It is calculated that these new technologies produce close to a zettabit (1 billion gigabytes) of data per year, and they will generate even more in the near future[1].

This huge data amount cannot be analysed just through human capacity. In this sense Artificial Intelligence (AI) has to be seen as a valuable support to use created knowledge in order to support the development of better health monitoring, better diagnoses and more personalized treatments, as well as live a healthier and more independent life in our societies. On top of this, AI can be the turning point instrument to address, at least, some of nowadays main challenges such as a) the ageing population and the spreading of chronic diseases, b) the lack of health personnel, c) the sustainability of the social-healthcare system, d) the healthcare system inequalities, e) the promotion of an advanced healthcare toward a personalised, predictive and participatory health.

As aforementioned, the wider adoption of AI is hindered by some associated risks preventing the full trust by its users. In particular, the main associated risks are variegated but can be summarised as follow.

1. **Lack of transparency and trust:** lack of understanding and trust in AI,
2. **AI algorithms errors due to for example:** data shift between AI training data and real-world data, unexpected variations in clinical contexts and environments,
3. **Privacy and security issue:** risk of data being exposed, shared without any consent, re-purposing, etc.,
4. **Misuse of medical AI tools:** lack of training, lack of digital literacy among patients, etc.,
5. **Gaps in AI accountability:** Legal gaps in current regulations, lack of ethical and legal governance for AI,
6. **Obstacles in AI's implementation into real-world healthcare:** limited data quality, lack of clinical & technical integration and interoperability of AI with existing clinical workflows

Various strategies could be implemented in order to smooth these risks and to favour the trust toward AI and to support its impact assurance.

In particular, what is relevant is to start from the needs and to analyse a possible solution in terms of: feasibility, available resources, acceptance from users' standpoint, management and sustainability elements.

 

In order to promote data quality and users' trust towards AI based solutions, different steps could be implemented (see Figure 2):

1. Needs and contexts analysis: in order to ensure the AI based solution alignment towards stakeholders' expectations and envisioned benefits as well as context's status in terms of services on stage, technological infrastructures, etc.

2. Co-design and stakeholders' training: co-creation of requirements and services' scenario based on identified needs and local ecosystem determinants. In this phase stakeholders training toward the innovative AI based tool/service is envisioned in order to share to them the values of innovative tool/service and smooth the issue of "misuse of AI tool.

3. Evaluation and monitoring: the AI based tool/service is implemented in a specific context. In this phase, a strong evaluation pathway with Key Performance Indicators (KPIs) has to be designed according to identified stakeholders' needs and expected benefits. A continues monitoring has to be executed in order to ensure the correct AI based innovation's implementation.

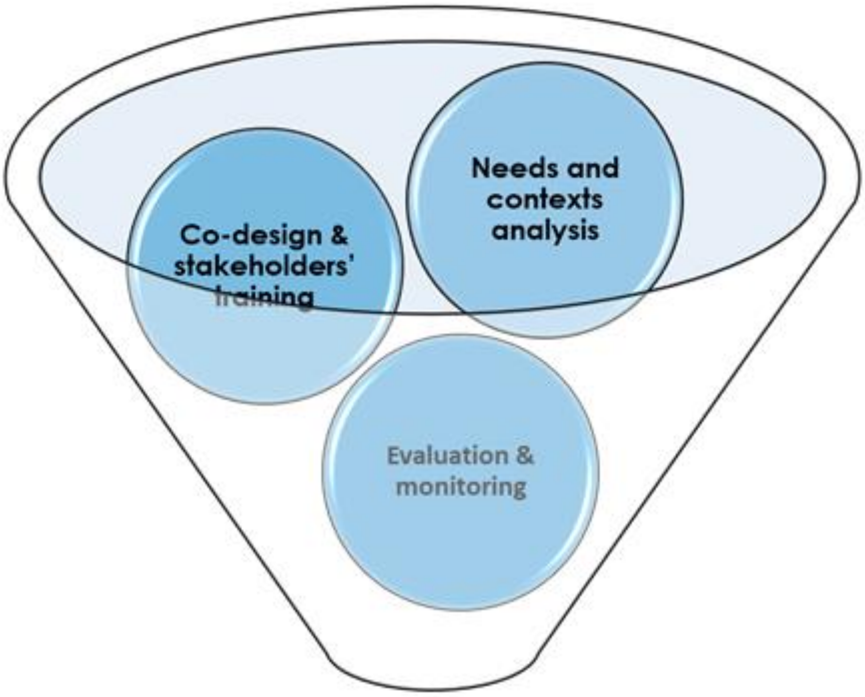


**Figure 2** ***Impact assurance key steps***

[1] Raghupathi W et al. Big data analytics in healthcare: promise and potential. Health Information Science and Systems 2:3-13, 2014.

# 9. Best practices for better health deployed by AIOTI members

<table>
<tr><td>Company/University/research centre name</td><td>Tyndall National Institute</td></tr>
<tr><td>Team's Name/Responsible</td><td>Salvatore Tedesco</td></tr>
<tr><td colspan="2">Reference Name of the key experience (deployed tools, algorithm, etc.)</td></tr>
<tr><td colspan="2">Wearable-AI</td></tr>
<tr><td colspan="2">Description of the reference domain and objectives of the deployed tools, algorithm, etc.)</td></tr>
<tr><td colspan="2">Research focuses on the adoption of novel wearable technology for healthcare and well-being applications integrating ML/AI, resulting in >60 peer-reviewed scientific publications and >20 collaborative industrial and research-oriented projects in this area. Those projects, covering different aspects of basic and applied research, lead to > € 1.1 million in grant funding as PI and co-PI. To achieve this, Salvatore has been responsible for leading a research team in the Wearable-AI for Health & Well-being area, involving postdoctoral researchers, junior researchers, research staff, and students.<br><br>Moreover, Salvatore has been also responsible for the creation of the Human Motion Lab at the Tyndall National Institute, with lab equipment worth > € 100 k including sport equipment, inertial-based motion capture systems, pressure insoles and mats, surface electromyography, force platform, video cameras, 10-camera motion capture system, acoustic emission sensors, and on-body physical activity monitors.<br><br>The ultimate goal is to develop AI-powered on-body wearable technology for healthcare & well-being applications, thus driving the vision for the ubiquitous adoption of wearables in healthcare and fitness to benefit society at-large.</td></tr>
<tr><td colspan="2">Experience innovative dimension and key strength points</td></tr>
<tr><td colspan="2">Research in the team can be mainly divided in five separate pillars:<br><br>• Human Motion Analysis, Biomechanics, Gait Analysis, and Lower-Limbs Rehabilitation<br><br>This pillar focuses on the investigation of the use of wearable sensors for gait analysis and lower-limb rehabilitation monitoring. The combination of wearable sensors and AI could represent a potential approach for developing an accurate functional assessment tool to curb the knee injury trend and identify factors that predispose to injury. For example, one study showed the feasibility of using body-worn motion sensors and machine learning approaches for the identification of post-ACL gait patterns even a number of years after the injury occurred. Moreover, the adoption of motion sensors for the assessment of adherence to home-based physical therapy programmes was investigated. Finally, a new wearable multi-sensing platform for knee remote rehabilitation was developed in another study.<br><br>• Healthy Ageing, Chronic Disease Management, and Parkinson's Disease<br><br>In this pillar, research is focused on the use of wearable sensors to collect data for monitoring elderlies with the aim to improve the quality of life of the elderly and support healthy aging, as well as monitoring falls in older adults. Moreover, the use of wearable sensors in hospital settings for the evaluation of specific therapeutic intervention was also investigated. Finally, the impact that wearable sensors have on patients with neurodegenerative diseases, such as Parkinson's, were considered in a number of studies. The analysis of physiological data for monitoring of subjects with Parkinson's disease is also currently ongoing as part of an industry-funded project.<br><br>• Health Markers and Physiological Monitoring</td></tr>
</table>

The aim of this pillar is to complete research in which physical and psychological data can be collected via wearable sensors and processed through ML models to evaluate the general health status of the users and potentially identify risk factors. Some examples involved:

- The adoption of acoustic emission sensors as a possible diagnostic tool for lower limbs pathologies,
- Knowledge-driven feature engineering for the detection of multiple symptoms on a blood pressure dataset,
- The development of novel ensemble techniques for all-cause and cancer-related mortality in an older adults population, showing the impact the wearable data has on those predictions.

An investigation on the adoption of wearable sensors and AI for the management of chronic migraine is also ongoing.

- Sports Analytics

This pillar mainly focused on investigating the use of body-mounted inertial sensors and neural networks for the estimation of Ground Reaction Forces (GRF) in runners to help prevent injury caused by running style. This research showed the potential application of ANN modelling for the estimation of all three GRF components (vertical, anteroposterior, and mediolateral) evident in running based on motion kinematic data. The adoption of AI for the prevention of hamstring injuries in athletes through EMG-based measurements is currently being investigated under the HOLISTIC project.

- Applied ML

This pillar focused on research in the field of Human Activity Recognition (HAR) showing that expert hierarchies constructed from domain knowledge are a suitable approach for multi-class decomposition method comparable in performance to standard methods, as well as the impact that subject-dependent and –independent models have on HAR datasets. Moreover as part of the edge AI Taskforce in EPoSS, we have also focused on defining the current opportunities, challenges, and future vision for edge AI.

**Experience encountered challenges**

Access to data is always a challenge in AI, especially in the healthcare domain, and that is why we have established our own Human Motion Lab in the institute. Otherwise, strong national/international collaborations were established to fill the gap in case of large multi-centre data which is not available in situ. Moreover, open access data and publicly available datasets have also been considered in a number of our publications.

Finally, given the lack of widespread computational resources in the institute, access to ICHEC (Irish Centre for High-End Computing) has been established for a number of projects.

The engagement with industry partners can be also challenging, as people without a technical AI background/knowledge may have very different expectations on what AI can achieve and the amount of data required for such purposes.

**Experience's implemented exploitation and eventual next steps**

A number of projects carried out by the team were in cooperation with national and US-based SMEs and MNCs. Licensing agreements are currently being discussed between the industry partners and UCC to license the developed AI technology.

**Experience dimension (number of users and/or stakeholders involved)**

Projects carried out by the team are mainly in the TRL range between 3-5. Developed technologies have been successfully tested on-the-field by the relevant stakeholders (SMEs and MNCs).

| Company/University/research centre name | **ITI – Instituto Tecnológico de Informática** |
|---|---|
| **Team's Name/Responsible** | F.Javier Pérez Benito |

**Reference Name of the key experience (deployed tools, algorithm, etc.)**

IA to optimze the process of breast density determination

**Description of the reference domain and objectives of the deployed tools, algorithm, etc.)**

Research focuses on helping to optimise the breast density assessment from digital mammograms. This biomarker is related to a higher risk of developing breast cancer. Supervised learning algorithms have been implemented to determine this. However, the performance of these algorithms depends on the quality of the ground-truth information, which expert readers usually provide. These expert labels are noisy approximations to the ground truth, as there is both intra- and inter-observer variability among them.

Thus, it is crucial to provide a reliable method to measure breast density from mammograms. The algorithms and tools provided by ITI present a fully automated method based on deep learning to estimate breast density, including breast detection, pectoral muscle exclusion, and dense tissue segmentation. A novel confusion matrix (CM)—YNet model for the segmentation step has been researched and implemented. This architecture includes networks to model each radiologist's noisy label and gives the estimated ground-truth segmentation as well as two parameters that allow interaction with a threshold-based labeling tool.

**Experience innovative dimension and key strength points**

This research task can be mainly divided in:

- Ethics in the AI development with sensible data

1.

2. This research was carried out with accordance with approval for usage of IMIM data was granted by the Ethics Committee at IMIM. Informed consent was obtained from all subjects involved in one of the datasets used in the study. For the IMIM dataset, patient consent was waived for the since anonymised retrospective data was used. The masks generated by our CM-YNet model for all INbreast images are described in a scientific paper and publicy available.

3.

- Collection of a dataset  from several sources and evaluation metrics

4.

5. A multi-center study covered women from 11 medical centers of the Generalitat Valenciana (GVA) as part of the Spanish breast cancer screening network. It included 1785 women with ages from 45 to 70. The cranio-caudal (CC) and medio lateral-oblique (MLO) views were available for 10 out of 11 of the centers, while one center only collected the CC view. This dataset was used for training, validation, and testing. The dataset was randomly partitioned into 75% (2496 mammograms) for training and validation (10%), and 25% for testing (844 mammograms). The mammograms of the same patient were always included in the same set.

6. Additionally, an independent dataset composed of 381 images obtained at the Institut Hospital del Mar d'Investigacions Mèdiques (IMIM) was included only for testing to obtain a better evaluation of the generalization performance of the models. Because the researchers at IMIM had a particular interest in testing the fully automated tool in various types of images, 283 out of the 381 images at IMIM were obtained from old acquisition devices with lower image quality, making the segmentation task more challenging. Only CC views were provided for this dataset.

7. The implemented CM-Ynet model achieved the highest DICE score averaged over both test datasets (0.82 ± 0.14) when compared to the closest dense-tissue segmentation assessment from both radiologists. The level of concordance between the two radiologists showed a DICE score of 0.76 ± 0.17. An automatic breast density estimator based on deep learning exhibited higher performance when compared with two experienced radiologists. This suggests that modeling each radiologist's label allows for better estimation of the

unknown ground-truth segmentation. The advantage of the proposed model is that it also provides the threshold parameters that enable user interaction with a threshold-based tool.

- Model selection and training for dense tissue estimation

8. Dense tissue segmentation can be obtained using two different approaches: (a) parametric approximation, in which threshold parameters are applied to the image to obtain the segmentation mask; and (b) mask approximation, which consists of directly assigning each pixel in the image as dense tissue or not. The advantage of using the mask approximation is that it can be used for any kind of segmentation as it is estimated at pixel level, but the expert labels should be defined by manually contouring the region of interest (ROI).
9. Therefore, the parametric approximation is more convenient for manual interaction with the expert as only a few (usually two or three) parameters need to be adjusted.
10. In a previous approach developed by ITI, the entirely convolutional neural network (ECNN) architecture was presented for dense-tissue segmentation. The DICE scores obtained with ECNN were close to the concordance achieved between the two radiologists. However, the performance on images with low gray-level resolution was not optimal.
11. Thus, four approaches were chosen for the study of the dense tissue estimation:
12. • ECNN: Parameter Estimation.
13. • U-Net: Mask Estimation.
14. • Y-Net: Hybrid Approach.
15. • CM-Segmentation: Learning Noisy Labels.
16. In the last study, ITI explored other architectures that directly estimate the segmentation mask by modeling each annotator's label independently. Finally, a new architecture named "confusion matrices-YNet" (CM-YNet) which aims to estimate the dense-tissue mask and the two segmentation parameters simultaneously was implemented and experimented. This architecture models each radiologist's label to estimate the dense-tissue mask and segmentation parameters for compatibility with threshold-based tools.

- Achieved results with the applied ML

17. Results were computed using the widely used Sørensen–Dice similarity coefficient which measures how much two masks M 1 and M 2 overlap. Our experiments were based on the assumption that the ground-truth label is not known. Therefore, we compared the results against each expert annotation independently and also obtained the mean DICE score between the estimated mask and the expert label which was closest to it.
18. The DICE scores for the different models are publicy available for the GVA and IMIM datasets. The mask approximation models (U-Net and YNET-mask) achieved better performance than their parametric counterparts (ECNN, YNet-param) for all training variations. Even though the expert labels used to train the model were obtained by setting the parameters predicted by the parametric models, these only have two degrees of freedom, and, therefore, the generalization performance could not outperform the mask approximation models, which have the freedom to adjust all pixels individually.
**19.** CM-YNet (mask) achieved a DICE score of 0.82, demonstrating a good level of concordance of CM-YNet with the segmentation provided by experienced radiologists.

**Experience encountered challenges**

Access to data is always a challenge in AI, especially in the healthcare domain. For this case, fortunately we had access to several datasets granted by GVA and IMIM institutions.

Also, some relevant technical challenges arouse. For example, while the CM-YNet hybrid approach improved the results of the fully parametric models, the results presented are based on comparing the DICE scores between the automatic model and the closest radiologist for each sample. This was performed due to the absence of a unique ground truth. Future work will involve including more than two expert labels in order to compare the results with fusion label methods such as majority voting or STAPLE . Another limitation is the pectoral muscle exclusion algorithm. The solution currently adopted, although robust, could be improved. The use of "for presentation" mammograms instead of "raw" images may be the reason for some of the differences among the acquisition devices. It would also be interesting to check if "raw" mammograms would avoid the preprocessing step. Finally, the estimation of the breast segmentation mask can be jointly estimated with a dense-tissue mask which would probably increase performance and decrease the segmentation time.

**Experience's implemented exploitation and eventual next steps**

| |
|---|
| During the last years and in an uninterrupted way, our AI-HEALTH research team has signed collaboration projects with the Department of Epidemiology and Evaluation, IMIM (Hospital del Mar Medical Research Institute). Actually the target is to deploy AI models on the hospital premises in order to run a pilot in a real environment. Subsequently and with the support of a good result, any kind of agreement could be proposed and discussed in order to transfer the developed AI technology. |
| **Experience dimension (number of users and/or stakeholders involved)** |
| The developed technology have been successfully tested in closed laboratory experiments, and in briefly should be going to be tested in a real environment by relevant stakeholders as IMIM Hospital. |

 

| Company/University/research centre name | ITI – Instituto Tecnológico de Informática |
|---|---|
| Team's Name/Responsible | Laura Arnal |

**Reference Name of the key experience (deployed tools, algorithm, etc.)**

IA to reduce hospital unplanned readmission

**Description of the reference domain and objectives of the deployed tools, algorithm, etc.)**

Research focuses on helping to optimise the discharge decision-making process by smartly ordering patients based on a severity score, thus helping to improve the usage of clinical resources.

Unplanned hospital readmissions mean a significant burden for health systems. A great number of heterogeneous factors can influence the readmission risk, which makes it highly difficult to be estimated by a human agent. However, this score could be achieved with the help of AI models, acting as aiding tools for decision support systems. In this paper, we propose a machine learning classification and risk stratification approach to assess the readmission problem and provide a decision support system based on estimated patient risk scores.

**Experience innovative dimension and key strength points**

This research task can be mainly divided in:

- Ethics in the AI development with sensible data

  This research was carried out with accordance with University and Polytechnic La Fe Hospital. Patient privacy was maintained by using data previously pseudo anonymised with non-traceable codes and only authorised people obtained data from electronic health records. The Hospital's Ethics Committee waived the requirement for written consent since data was pseudo anonymised and the study complies with national and European legal requirements regarding data protection.
- Collection of a dataset from unscheduled readmissions

  This task mainly focused on collecting and organizing data in collaboration with La Fe Health Department and the Health Research Institute of La Fe University Hospital. Therefore, the dataset used during the experiments was build from a subset of patient information included in their Electronic Health Record (EHR) system. La Fe Health Department has deployed an EHR at different care levels, including over 20 million records, effectively organized reaching stage 6 in the eight-stage (0–7) EMRAM maturity model. Currently, the data lake layer includes structured and semi-structured information, coming from several information systems involving clinical activity, such as emergency care settings, outpatient, hospitalization, clinical reports, surgical unit, intensive care unit, hospital at home care. The data feeding the platform is composed by the aggregation of 22 datamarts and comprises 750 million rows, 84 tables, 4064 columns resulting in a total size of 640Gb. Data updates are scheduled on a daily, weekly, or monthly basis, depending on the datamart. The subset of data used for our study contains information from 35034 episodes of 22370 patients. Data was gathered and merged into one table containing information from five categories: Consumptions, Laboratory, Treatments, Hospitalisation, Comorbidities.

  As a result, a table of 35034 rows and 962 columns was obtained, where each row accounts for an admission episode. For each one of these episodes we know the date of admission and discharge. The observation period for this dataset ranges between 1/1/2015 and 30/12/2018. In order to train a classification model, the target variable takes unit value if readmission has taken place within 30 days following previous discharge, and zero otherwise. Following this rule, only 9.85% of the episodes in the dataset are positive class (readmissions).
- Model selection and training

  Three ensemble methods were chosen for the binary classification task:

  • Random Forest (RF) is an ensemble supervised learning method which incorporates many weak decision tree models. During the training phase, each of these individual trees sees a random set of features, ensuring some level of individuality. To make a prediction, each of these trees makes its guess and then a voting phase takes place to select the ensemble's final decision.

 

• Gradient Boosting (GB) is also a decisiong tree ensemble technique, not bagged horizontally but in a stage-wise manner known as boosting. Arranged like a chain, each tree's training parameters depend on the result of the previous one, optimising the performance in the later tree.

• Extreme Gradient Boosting (XGBoost) is an enhanced, highly efficient and computationally effective implementation of the GB algorithm. Moreover, XGBoost is regularized in a way which usually lends to better results and prevents overfitting.

Additionally, decision tree models were trained in order to compare the performance of the previous with that of a simpler one. In this study, samples were distributed between train and test sets using nested 10-fold cross-validation. After the model was trained in each split, and performance metrics were calculated on the corresponding test set.

- Achieved results with the applied ML

Results were computed performing nested 10-fold cross-validation on the entire dataset and averaging the resulting metrics computed on each test set. The best ROC-AUC results were obtained with XGBoost models and no balancing method Moreover, the best calibration was achieved performing Beta Calibration, as it shows both the lowest BrierScore and LogLoss metrics. While the ROC-AUC varies slightly with different calibration options, improving the model performance is not a target of this task. The calibration step aims to scale the prediction results to a more understandable, truthful class probability range, but does not enhance class separability.

The best classification results are obtained with an XGBoost model, with a better match between predicted and real class probabilities after performing Beta Calibration (ROCAUC: 0.693, BrierScore: 0.0851, LogLoss: 0.3009). Besides, classification metrics had great room for improvement in terms of both recall and precision.

**Experience encountered challenges**

Access to data is always a challenge in AI, especially in the healthcare domain. For this case, the Valencia Health Care Department has imposed restrictions on the data used in this study due to its potential for use in identification of participants. The authors have been granted data access after signing specific agreements guaranteeing and demonstrating compliance with national and European legal requirements regarding data protection and after the approval of an Ethics Committee, and so are not publicly available.

Pseudo-anonymised data are however available upon reasonable request by contacting the Medical research Institute of Hospital La Fe to any researcher wishing to use them for non-commercial purposes and who could guarantee and demonstrate compliance with national and European legal requirements regarding data protection.

**Experience's implemented exploitation and eventual next steps**

During the last years and in an uninterrupted way, our AI-HEALTH research team has signed collaboration projects with the Valencia Health Care Department. Actually the target is to deploy AI models on the hospital premises in order to run a pilot in a real environment. Subsequently and with the support of a good result, any kind of agreement could be proposed and discussed in order to transfer the developed AI technology.

**Experience dimension (number of users and/or stakeholders involved)**

The developed technology has been successfully tested in closed laboratory experiments, and in briefly should be going to be tested in a real environment by relevant stakeholders as La Fe Health Department and the Health Research Institute of La Fe University Hospital.

| **Company/University/research centre name** | **Trialog** |
|---|---|

**Reference Name of the key experience (deployed tools, algorithm, etc.)**

- **SAREF4EHAW-Compliant Knowledge Discovery and Reasoning for IoT-based Preventive Healthcare and Well-Being**. Amelie Gyrard and Antonio Kung. Elsevier Book: Semantic Models in IoT and e-Health Applications 2022.
- **Reasoning Over Personalized Healthcare Knowledge Graph: A Case Study of Patients with Allergies and Symptoms**. Amelie Gyrard, Utkarshani Jaimini, Manas Gaur, Saeedeh Shekarpour, Krishnaprasad Thirunarayan, and Amit Sheth. Elsevier Book: Semantic Models in IoT and e-Health Applications 2022.
- **A Naturopathy Knowledge Graph and Recommendation System to Boost the Immune System**. Amelie Gyrard and Karima Boudadoud. Elsevier Book: Semantic Models in IoT and e-Health Applications 2022.
- **Knowledge Engineering Framework for IoT Robotics Applied to Smart Healthcare and Emotional Well-Being**. Amelie Gyrard, Kasia Tabeau, Laura Fiorini, Antonio Kung, Eloise Senges, Marleen De Mul, Francesco Giuliani, Delphine Lefebvre, Hiroshi Hoshino, Isabelle Fabbricotti, Daniele Sancarlo, Grazia D'Onofrio, Filippo Cavallo, Denis Guiot, Estibaliz Arzoz-Fernandez, Yasuo Okabe, Masahiko Tsukamoto. International Journal of Social Robotics 2021. Springer Nature.
- **IAMHAPPY: Towards An IoT Knowledge-Based Cross-Domain Well-Being Recommendation System for Everyday Happiness**. Amelie Gyrard, and Amit Sheth. IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE) Conference 2019. Published within the Elsevier Smart Health Journal.
- **Personalized Health Knowledge Graph.** Amelie Gyrard, Manas Gaur, Krishnaprasad Thirunarayan, Amit Sheth, and Saeedeh Shekarpour. Workshop on Contextualized Knowledge Graph (CKG) co-located with International Semantic Web Conference (ISWC 2018).

Other:

- **Drug Abuse Ontology to Harness Web-Based Data for Substance Use Epidemiology Research: Ontology Development Study**. Usha Lokala, Francois Lamy, Raminta Daniulaityte, Manas Gaur, Amelie Gyrard, Krishnaprasad Thirunarayan, Ugur Kursuncu, Amit Sheth. JMIR Public Health and Surveillance 2022.
- **Question Answering for Suicide Risk Assessment Using Reddit.** Amanuel Alambo, Manas Gaur, Usha Lokala, Ugur Kursuncu, Krishnaprasad Thirunarayan, Amelie Gyrard, Amit Sheth, Randon S. Welton, and Jyotishman Pathak. IEEE International Conference on Semantic Computing (ICSC) 2019.
- **Applying Internet of Things for personalized healthcare in smart homes.** Soumya Kanti Datta, Christian Bonnet, Amelie Gyrard, Rui Pedro Ferreira Da Costa, Karima Boudaoud. Wireless and Optical Communication Conference (WOCC) 2015

**Description of the reference domain and objectives of the deployed tools, algorithm, etc.)**

See section above

**Experience innovative dimension and key strength points**

Scientific papers are outcome or research and innovation projects

**Experience encountered challenges**

Linking health data not that easy
Too many standards data formats which one to choose

**Experience's implemented exploitation and eventual next steps**

Standardization:
ISO

 

# 10. Conclusions

In the post-covid-19 era, the healthcare domain represents a unique opportunity for favouring the AI wider adoption. In fact, the healthcare domain is suffering a huge pressure due to internal and external massive changes that are pushing a development and a transformation of its dynamics. In this transformation process, AI can play a major role for smoothing the human and economic resources' allocation and improving healthcare treatments management.

According to this white paper, we believe that AI has an important role to play in the healthcare offerings of the future. Despite this, relevant challenges that have to be addressed. We tried to provide some answers to this challenges highlighting which is the AIOTI WG Health vision according to its members' point of view.

As addressed several times throughout the document, the greatest challenge to AI in the healthcare domains is not whether the technology will be capable enough to be useful, but rather ensuring its adoption in daily clinical practice overcoming barriers such as mistrust and scepticism. For widespread adoption to take place, AI systems must be approved by regulators, integrated with healthcare systems practices, standardised to a sufficient degree that similar products work in a similar fashion, taught to clinicians, paid for by public or private payer organisations. These challenges will ultimately be overcome, but time is needed. In Europe a strong push is encouraged towards AI adoption in daily practices, as testified by the key experiences included in the paper (ref. par. 9) but stronger alignment among EU member states, as well as within each country among the various decision making and legislative institutions, could further facilitate the AI wider adoption. As a result, we expect to see an ever more increasing use of AI in health daily practice within the next future connected to an increasing mingling with other AI application area such as smart cities and green economy.

## List of Contributors

The document was written by several participants of the AIOTI WG Health.

**Editor:**

Pietro Dionisio, Medea Project

**Reviewer:**

Damir Filipovic, AIOTI

**Contributors:**

Pietro Dionisio (Medea)

Roumen Nikolov (Virtech)

Amelie Gyrard (Trialog)

Katerina Petelova (Quanterall)

Wolfgang Eberle (IMEC)

Lucas Sanjuan Vinas (ITI)

Ruben San Segundo (UPM)

Salvatore Tedesco (Tyndall National Institute)

Cristina Martín Andonegui (Vicomtech)

Cian O'Murchu (Tyndall National Institute)

## Acknowledgements

## About AIOTI

AIOTI is the multi-stakeholder platform for stimulating IoT and Edge Computing Innovation in Europe, bringing together small and large companies, academia, policy makers and end-users and representatives of society in an end-to-end approach. We work with partners in a global context. We strive to leverage, share and promote best practices in the IoT and Edge Computing ecosystems, be a one-stop point of information on all relevant aspects of IoT Innovation to its members while proactively addressing key issues and roadblocks for economic growth, acceptance and adoption of IoT and Edge Computing Innovation in society. AIOTI's contribution goes beyond technology and addresses horizontal elements across application domains, such as matchmaking and stimulating cooperation in IoT and Edge Computing ecosystems, creating joint research roadmaps, driving convergence of standards and interoperability and defining policies.