



Quality medical data management within an open AI architecture – cancer patients case

Mirjana Ivanovic, Serge Autexier, Miltiadis Kokkonidis & Johannes Rust

To cite this article: Mirjana Ivanovic, Serge Autexier, Miltiadis Kokkonidis & Johannes Rust (2023) Quality medical data management within an open AI architecture – cancer patients case, Connection Science, 35:1, 2194581, DOI: [10.1080/09540091.2023.2194581](https://doi.org/10.1080/09540091.2023.2194581)

To link to this article: <https://doi.org/10.1080/09540091.2023.2194581>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 13 Apr 2023.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)



RESEARCH ARTICLE



Quality medical data management within an open AI architecture – cancer patients case

Mirjana Ivanovic ^a, Serge Autexier ^b, Miltiadis Kokkonidis^c and Johannes Rust^b

^aDepartment of Mathematics and Informatics, University of Novi Sad, Faculty of Sciences, Novi Sad, Serbia;

^bGerman Research Center for Artificial Intelligence (DFKI), Bremen, Germany; ^cNETCOMPANY-INTRASOFT S.A., Luxembourg, Luxembourg

ABSTRACT

In contemporary society people constantly are facing situations that influence appearance of serious diseases. For the development of intelligent decision support systems and services in medical and health domains, it is necessary to collect huge amount of patients' complex data. Patient's multimodal data must be properly prepared for intelligent processing and obtained results should be presented in a friendly way to the physicians/caregivers to recommend tailored actions that will improve patients' quality of life. Advanced artificial intelligence approaches like machine/deep learning, federated learning, explainable artificial intelligence open new paths for more quality use of medical and health data in future. In this paper, we will focus on presentation of a part of a novel Open AI Architecture for cancer patients that is devoted to intelligent medical data management. Essential activities are data collection, proper design and preparation of data to be used for training machine learning predictive models. Another key aspect is oriented towards intelligent interpretation and visualisation of results about patient's quality of life obtained from machine learning models. The Architecture has been developed as a part of complex project in which 15 institutions from 8 European countries have been participated.

ARTICLE HISTORY

Received 20 December 2022
Accepted 19 March 2023

KEYWORDS

Quality of life; cancer patients; data management in medical domains; Cloud/Edge distributed environment

Abbreviations

AI	Artificial Intelligence
DSS	Decision Support System
HL7 FHIR	Health Level 7, Fast Healthcare Interoperability Resources
EHR	Electronic Health Record
HIS	Hospital Information System
FL	Federated Learning
ICT	Information Communication Technologies
ML	Machine Learning

CONTACT Mirjana Ivanovic mira@dmf.uns.ac.rs Department of Mathematics and Informatics, University of Novi Sad, Faculty of Sciences, Trg Dositeja Obradovica 4, Novi Sad 21000, Serbia

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

QoL Quality of Life
XAI Explainable AI

1. Introduction

We are witnesses that contemporary society is constantly coping with serious stressful events and global health situations (like bird flu, dengue virus, the COVID-19 pandemic, etc.). Therefore, people everyday face situations that can negatively influence their health. Some studies (Cohen et al., 2019) present relation between stressful life events and health, and the role of stressors in increasing critical disease risks (Chida et al., 2008). Additionally, the population is getting old and more people suffer from different comorbidities and multi-morbidity (Fan et al., 2021; Kingston et al., 2018). Such circumstances raise awareness and importance of taking care of health status of population and necessity of the development of sophisticated services, devices, tools and general high-quality clinical/hospital information systems (He et al., 2019; Schulz et al., 2019). Improvements of health-related quality of life (QoL) aspects (Link 1) are important and widely recognised in therapies and follow-ups of serious diseases survivors. Apart from usual health problems like anxiety, sleep disorders, mental impairment, and so on (Norbye et al., 2022), cancer patients experience a serious disruption of important QoL aspects during active oncological treatment period or follow-up care like pain, appetite loss, psychological difficulties, sexual disfunction, and so on (Link 3).

For the development of sophisticated software services that can help patients to successfully cope with everyday activities, it is needed to collect and properly integrate wide spectra of complex data. As addition to traditional clinical data, it is necessary to collect data from other sources like smart wearable devices, environmental and nutritional data, data about patients' personal experiences, and so on. Collected data represent starting point for application of a range of Artificial Intelligence (AI) and Machine Learning (ML) methods with intention to find out and offer reliable and trustworthy solutions for patients' therapies, treatments, recommendations and interventions. Usually results achieved after AI/ML data processing are in format that is not acceptive and friendly for physicians and it is needed to process them further to reach more understandable forms (He et al., 2019). However, in any stage of managing, using, analysing sensitive patients' medical and health data it is necessary to take care of privacy protection (Kalloniatis et al., 2021). Modern approaches in medical and health data management support more appropriate interventions, and usually more personalised treatments (Claeys & Vialatte, 2014; Ivanović & Ninković, 2017; Ivanovic & Balaz, 2020). Considering several modern medical and health systems, it is possible to recognise crucial subsystems of a typical "Health AI system" (Ivanovic et al., 2022).

- 1) *Data management subsystem.* It is responsible for secured collection of patients' data from multiple sources (Siddique et al., 2018). This stage of data management encompasses the aggregation of heterogeneous data, curation and preparation in formats appropriate for application of AI/ML approaches. During all activities, privacy preservation of patients' data must be guaranteed.
- 2) *Subsystem for intelligent data processing.* It is responsible for comprehensive and intelligent data processing and analyses of computed results. Usually, characteristic patterns

of behaviours are recognised, and powerful predictive models are generated that support treatment recommendations, quality of life predictions and tailored interventions for patients.

- 3) *Intelligent/smart interface*. This subsystem usually represents connection between patients' data and results achieved by AI/ML subsystem. To obtain trustworthy insights of the patient's health the interface is usually implemented using powerful AI techniques (like explainable AI – XAI, data visualisation, virtual agents, etc.). Interface uses results achieved by AI/ML models and suggests adequate, personalised treatments, interventions, various activities, nutrition and other actions to increase patients' QoL. To make results of predictive models more understandable to end-users, (usually physicians) XAI methods like Shapley Additive exPlanations, LIME, Anchors, Textual Explanations of Visual Models, Integrated Gradients are used (Holzinger et al., 2022), and adjusted to different kinds of devices.

Under our ASCAPE project – Artificial intelligence supporting cancer patients across Europe (<https://ascap-project.eu/>), an innovative and unique AI/ML-based Open Architecture for personalised treatment of breast and prostate cancer patients has been developed. In ASCAPE Open AI Architecture, the three subsystems with similar functionalities are implemented. In this paper, we will attempt to highlight the importance of patients' data management and present essential functionalities of ASCAPE Architecture achieved within first and third subsystems. Also, we will briefly present crucial functionalities of Cloud/Edge approach which is adopted in our architecture for personalised medical decisions (Ilić et al., 2023; Tyler et al., 2020; Venne et al., 2020).

The rest of the paper is organised as follows. In Section 2, we will present several modern systems that support cancer patients and briefly compare them with our approach. Section 3 considers essential and modern aspects and complex management of multi-source and multimodal medical and health data. Section 4 is devoted to presentation of crucial functionalities of our approach with special attention on data management, i.e. collection and aggregation of cancer patients' data and presentation of results of predictive models in user-friendly form for physicians. Brief discussion on current implementation of ASCAPE architecture and possibilities for its future use is given in Section 5. Concluding remarks are pointed out in the last section.

2. Contemporary intelligent cancer-related projects

In recent years, the European Commission (https://ec.europa.eu/info/research-and-innovation_en) has been financing a number of large projects with focus on health, medical and wellbeing domains. Among them, a group of projects was oriented towards developing intelligent architectures and adequate services to improve support for cancer patients, addressing QoL aspects on the basis of analyses of complex patients' data. The main result of the ASCAPE project is implementation of an innovative Open AI architecture and its main functionalities to propose tailored actions for breast and prostate cancer survivors to improve their QoL aspects. However, in this paper, it is not possible to present all innovations and solutions achieved within three years of project duration. We decided to put main

focus on data management from two essential aspects. First, we will concentrate on preparation of multisource data for AI/ML processing. Second, we will focus on visual intelligent representation of results obtained from AI/ML predictive models after data processing.

In the rest of the section, our intention is to present current state of the art in cancer patients' treatment. We will briefly present several projects related to ASCAPE but also, we will mention results from several related papers.

ONCORELIEF (<https://oncorelief.eu/>) is a project focused on patients with two kinds of cancer Acute Myeloid Leukaemia and Colorectal cancer. The motivation of the project is based on the fact that the burden of cancer is rising globally but also that recent improvements in early detection and therapeutic treatment have improved cancer survival and follow-up care. The development of adequate user-centred AI System allows the utilisation of big datasets to facilitate the integration of QoL assessment instruments. The focus in the system is on the improvement of post-treatment health status, increase the overall wellbeing and follow-up care of cancer patients. The proposed framework includes rather standard components for data processing, using AI models and providing personalised support for patients in post-treatment activities and tasks. Primary it suggests actions regarding the patients' overall health through an intuitive smart digital assistant (Guardian Angel). In terms of data management, the ONCORELIEF system merges data from various sources but relies on collecting them on a central infrastructure for the purpose of training ML models. The distributed style of processing datasets from different health institutions and existence of local and global AI/ML models produced by contemporary Federated Learning paradigm is missed from this architecture. However, the ASCAPE architecture is highly based on Federated Deep Learning and distributed data processing using modern approaches for security and privacy preserving patients' data.

Another noteworthy project is BD2Decide (<http://www.bd2decide.eu/>) which aim is to develop a more precise prognostic prediction in cancers of the Head and Neck Region (HNC). The main intention of developed system is to support the first-line treatment, i.e. to maximise the therapeutic results and minimise the impacts of HNC therapy. BD2Decide has been developed a decision support system (DSS) that combines existing data and advances of existing personalised prognostic models with technologies for data visualisation to decide about HNC patients' treatment. The clinical DSS is meant to provide the facilities that allow physicians to manage the patients' data at the typical stages of diagnosis, treatment and follow-up. BD2Decide should help a physician to decide based on evidence medicine but also on scored prognostic factors derived by the system, specific for each individual and for the reference populations. Therefore, the goal is that the physician can better react having available visualisations showing how the patient is positioned with respect to the reference population. Several data management components and functionalities are developed (Deliverable D2.3; Deliverable D5.1): (a) Data collection from different sources (even images and some genomics data); (b) Preparation and training of several models (Imaging, Statistical, Genomics); (c) Development of dashboard and visual analytics tool for patient-physician and physician-physician communication. Within the project, data and patients from several hospitals were considered in implementation of the DSS. However again, comparing to ASCAPE, this system relies on a centralised infrastructure and collecting data from the various hospitals in central repository. In our approach and architecture, high priority is put on privacy data protection and security in distributed environment. Innovative aspect of our approach is based on use of advanced techniques like

Differential Privacy (DP) and Homomorphic Encryption (HE) within Federated Learning (FL) style.

The project that has several similar aspects with ASCAPE is FAITH (<https://www.h2020-faith.eu/>) – a Federated Artificial Intelligence solution for monitoring mental Health status after cancer treatment. The project aims to provide an AI application based on FL that remotely identifies depression marker in people that have undergone cancer treatment. Initial validation of the system is done on the data gathered from lung and breast cancer survivors in three hospitals. Within the system data on lifestyle, but also clinical and behavioural data are considered. Based on AI models, the aim is to identify parameters that reveal the transition from the healthy state to the initiation of depression. Similarly, as in ASCAPE, this approach is based on FL paradigm and training global AI models that will be used in hospitals on their local data. Within the project, data about patients are collected every 3 months for a 12-month period, concentrating on identifying and analysing depression markers in data pertaining to sleep, voice patterns, activity and nutrition. However, in our approach not only depression is considered but all together 15 QoL aspects for breast and 12 QoL aspects for prostate cancer are considered to be predicted using developed AI/ML models. Additionally, the ASCAPE predictive models were initially trained on huge dataset of already existing cancer patients' data concerning longer period and history of illness, starting from 3rd month up to 10 years.

ASCAPE project is one of the research efforts in that same space. Its' aim was to address the QoL aspects of breast and prostate cancer patients, providing a privacy preserving ML-based framework supporting both Federated Learning and Homomorphic Encryption for decision support to physicians providing personalised predictions and interventions for their patients on the basis of data coming both a multitude of sources (EHR, clinically validated Patient Reported Outcome Measures (PREM) questionnaires, wearables, open data, etc.). It serves as a prime example of the State-of-the-Art in research in this space as: (a) while its piloting stage was meant to be focused on breast and prostate cancer, its overall design and architecture produced an extensible, adaptable Open AI framework that can be usable in both future research and integrated into clinical practice, (b) even in the context of the two cancers it addresses in the context of its piloting phase, it builds up a particularly rich data model combining data from diverse sources as inputs and aspires to give to physicians outputs/results of direct relevance to the care of their patients including both predictions and intervention suggestions and (c) it goes beyond standard input data management as it aims to provide a scalable privacy preserving AI infrastructure that does not require that patient data are revealed to a centralised ML Cloud in order for models to be built from the data of multiple centres, i.e. health institutions (hospitals, primary care centres, etc.).

In this last respect, projects such as ASCAPE and FAITH, offer a technological answer to the multifaceted privacy concerns healthcare providers are faced with when considering contributing with their data to help improve ML models and make them more relevant to their patients, problems that projects such as ONCORELIEF and BD2DECIDE fail to address. At the same time, ASCAPE is quite ambitious with respect to the decision support it provides, and it can do so, because its suggestions and predictions are presented to physicians who may use their own expertise and understanding of their patients to filter the ASCAPE predictions and recommendations and not directly to patients. Due to its flexibility, the ASCAPE architecture can be adapted to different scenarios and applications (Lampropoulos et al., 2021). Its current application to breast and prostate cancer patient monitoring

includes highly rich and complex data about patients, treatments, interventions, follow-ups; it considers 27 QoL aspects, uses a wide range of predictive AI/ML models combining training and using three kinds of models, local and global via Federated Learning and Cloud-based homomorphically encrypted models; privacy preserving based on modern approaches as Differential Privacy and Homomorphic Encryption; more complex AI/ML architecture that offers high-quality and trustable data visualisation for physicians based on XAI methods and friendly dashboard user-interface. Thus, exploring the ASCAPE architecture and overall design specifically serves to highlight important aspects of a broad family of systems, especially because it includes an amalgamation of features that other projects may or may not have and packages them in one comprehensive Open AI framework (Deliverable D1.1; Deliverable D2.4; Deliverable D4.1).

QoL is important aspect studied in different cancer-related diseases (Sinha & Heuvel, 2011; Saadat et al., 2017; Link 1) as one of common chronic conditions nowadays. Some successful applications of QoL prediction in this area include: the comparison of the pre and the post-treatment QoL in cervical cancer patients (Kumar, 2014) as well as the prediction of five-year lung cancer survival on the basis of QoL aspects (Sim et al., 2020). However, to the best of our knowledge, ASCAPE project results represent the first efforts to develop a QoL prediction model for breast and prostate cancer patients as they are among the most widespread cancers of male and female population. For example, in the paper (Yang et al., 2021), authors focus on prostate cancer patients but consider only one type of intervention while within ASCAPE we consider wide range of them (see Figure 5).

3. Health and medical data management – several key aspects

Electronic Health Records (EHRs) are usually the core source of patient's data both in modernised clinical settings but also in intelligent health solutions explored in a research context. Such records keep data of various important aspects of a patient (like clinical information, diagnoses, medication, treatment, etc.) and provide a more comprehensive and longitudinal view of the patient's health. They also incorporate information sources outside traditional which may include Patient Reported Outcome Measure (PROM) questionnaires, nutritional, activity tracking and other data that help present a more complete view of a patient's health. Additionally, the collection of a patient data from other diverse and multiple sources is necessary prerequisite for obtaining richer datasets and reliable results of its processing (Autexier et al., 2022). From this point of view, interesting example is the CrowdHEALTH project (Gallos et al., 2019). It is focused on the combination of patient's data from various sources in order to form Holistic Health Record. Within the project, an integrated holistic platform is developed that incorporates big data management mechanisms: acquisition, cleaning, integration, modelling, analysis, information extraction and adequate powerful interpretation (Kyriazis et al., 2017). In ASCAPE project, patients' data records also contain data collected from multiple sources and similar big data management mechanisms are applied.

After collecting patients' data from multiple sources, the next logical step in medical systems is to select adequate AI/ML techniques for curation, cleaning, managing, modelling and processing the data to achieve high-quality results and outcomes. Obtained results are further exploited and usually presented to physicians and/or caregivers in understandable, trustworthy and user-friendly manner.

Contemporary development of reliable medical DSS is based on integration of “traditional” medical and health data sources with novel ones like data from smart and wearable devices, IoT and sensors generated data, open data, environmental data, etc. (Hiremath et al., 2014). Due to their capacity for continuous data recording, wearable devices (Wu & Luo, 2019) support the collection of huge amounts of patient’s data such as heart rate, number of steps, calories burned, sleep stages, and so on. Such additional sources of patient data can increase the reliability of developed predictive models. Within ASCAPE, AI/ML/FL-based architecture wearable devices are used to improve personalised treatments and interventions of cancer patients. Wearable devices also facilitate connection between physicians and patients and offer great potential to support and even increase medical decisions quality.

However, apart from care of developing services that can help cancer survivors in improving their QoL aspects, it is important to consider cancer survivors personal experiences. For this purpose, numerous questionnaires/instruments to measure cancer patient’s individual views of his/her health status are usually used. Two widely used measures are PROMs (Patient Reported Outcome Measures) and PREMs (Patient Reported Experience Measures). From that point of view, it is valuable to acquire early reliable predictors and QoL aspects over time to improve treatment decisions and especially follow-up strategies. Medical systems should support the utilisation of Holistic Health Records, powerful AI/ML approaches that facilitate the integration of QoL instruments (like PROMs and PREMs), implementation of a user-centred communication interface and personalised support. All these comprehensive instruments are also incorporated in ASCAPE data management subsystem using emergent technologies nowadays.

To process and use for different purposes rich patients’ health records, various health terminologies and ontologies have been used to achieve great level of standardisation and unification of data representations. A health terminology represents a kind of “language” used to code entries in EHRs and other medical and health data and there are a range of them: various International Classification of Diseases, ICD-9, ICD-10, IDC-11, LOINC, CPT, SNOMED CT, and so on (Link 4). Such terminologies are important for obtaining interoperability between systems and for integrating data; exchange of data between systems as codes from different systems have to be compatible; and are standards for obtaining mapping of various vocabularies and smooth “communication” between independent medical and health systems.

SNOMED CT – Systematised Nomenclature of Medicine Clinical Terms is a comprehensive, computerised healthcare terminology widely used in medical and health information systems in numerous countries (Lee, Cornet, et al., 2013). It obtains higher level of interoperability between different data sources, health information systems (HIS), and services. Medical and health data represented in such standard forms can be later easier transformed in some other formats more suitable for AI/ML data processing.

Finally, in medical systems, it is important to follow Health data models from international standards and among frequently used is HL7 FHIR – Health Level 7, Fast Healthcare Interoperability Resources. It supports six major categories of resources and possibility to present the content of the FHIR resources in widely used formats like XML, JSON, Turtle and other facilities to create a clinical data model that includes interoperability artefacts composed of a set of modular components.

The process of data aggregation is an important step in producing compact patients' health records combining information from multiple sources. Aggregate management offers important insight in patient's health status (Lahiri et al., 2019). ASCAPE data management and clinical data model are essentially based on SNOMED CT and HL7 FHIR standards (see simple example in Figure 2).

Another vital aspect in processing sensitive medical and health data is privacy preserving (Siddique et al., 2018). To prevent discovering personal patients' data at any stage of processing and using achieved results, numerous techniques for de-identification are developing and using in practical solutions: Character/record masking, Shuffling, Anonymisation, Collectively de-identification, Pseudonymisation, Generalisation. Some well-known and widely used privacy preserving techniques include K-anonymity, L-diversity and T-Closeness. However, recent, and reliable techniques as Differential Privacy and Homomorphic Encryption are getting to play crucial role in Cloud/Edge medical and health architectures for processing complex patients' data (Kaissis et al., 2020).

DP is the systematic randomised modification used for processing data in order to reduce information about the single patient (Ficek et al., 2021). The term represents the umbrella of possible techniques and implementations starting from simple random shuffling of the input data to the introduction of noise to the dataset (for example Gaussian, Laplace or Exponential mechanisms) (Ji et al., 2014). Local DP ensures privacy at the source of the data, which is well suited for medical applications, but also for Federated Learning applications that enhance EHR data with additionally collected data from smart-phones/wearable devices. Within ASCAPE, Laplace DP is used for introduction of noise to the dataset and experimental results showed its efficiency (Deliverable D2.4; Savić et al., 2023).

HE is security preserving technique that enables arithmetic operations on ciphertexts without need to decrypt original data. It is a popular way for protecting sensitive patients' data leakage within FL and in distributed environments. HE refers to homomorphism in algebra and includes multiple types of encryption schemes where the encryption and decryption functions are homomorphisms between plaintext and ciphertext spaces. HE approaches perform computations on encrypted data without decrypting it, resulting computations are left in an encrypted form and they are identical as those get from unencrypted data (Kaissis et al., 2020; Park & Lim, 2022).

Depending on general organisation of specific medical platform, it can consider and analyse patients' datasets from arbitrary number of health institutions. Such system can train and lately use AI/ML predictive models considering common knowledge gained from all available datasets. Such types of platforms adopt distributed data processing based on powerful AI/ML predictive models. Closely related to such approach is employment of Federated Learning which significantly reduces the risk of data privacy being compromised. FL is based on existence of multiple Edge nodes, i.e. clients (in ASCAPE case health institutions), that work together to train a single model organised and stored on single server, i.e. Cloud.

An FL system has two actors: multiple Edge nodes and the Server/Cloud. The Server coordinates the training process between all Edge nodes that participate in the construction of a global model. Each Edge node receives a copy of the global model to be trained and

updates it based on available local data in particular Edge node. When training phase is finished all Edge nodes are participating in the training send their updated model's weights back to the Server to synchronise them and produce unique global model.

In such an approach, sensitive patient's data remain decentralised and FL keeps the data at health institutions (locally) and transfers only models' updates to the main server. Predictive models created and trained on local nodes'/Edges' data are participating in creating global/centralised federated models. This is succeeded by distributing the model architecture and initial weights to all Edge nodes participating in producing a global/federated model. Furthermore, Edge nodes train their copy of the global model on local data. When training is finished achieving satisfactory results, updated weights are sent back to the FL Server (Cloud) to contribute to creating new or updating existing common global/central model (Ivanovic et al., 2022; Ilić et al., 2023). Presented approach is adopted in ASCAPE architecture as well.

All approaches, standards for data representation and technologies mentioned in this section are crucial in managing medical and health data in contemporary systems' development. They represent modern, efficient, reliable and promising approaches, for application in contemporary and future intelligent, personalised medical systems, that constantly evolving. Additional quality that they offer is based on high support of interoperability between different systems, high performances of execution in distributed environments, and reliable security and privacy data preserving.

ASCAPE AI Open Architecture at the moment encompasses four health institutions and all data management activities, training and evaluation of AI/ML/FL predictive models, visualisation and explainability of obtained results/outcomes were performed using datasets from some of these institutions. From the beginning, we faced the great differences between available datasets as they contained distinct features, i.e. predictive and target variables. So, to introduce unification and systematisation among different datasets the only logical solution was to use some of widely spread and accepted standards for medical and health data representation and in such a way support integration of data from different datasets and support common data processing. It was logical to select SNOMED CT and HL7 FHIR as open source and widely accepted standards in medical and health institutions and systems. Such decision guaranties greater possibilities for extending ASCAPE architecture and inclusion of arbitrary number of new health institutions that would like to use great functionalities that ASCAPE solution offers.

4. ASCAPE – an innovative approach for supporting cancer patients quality of life

With appropriate medical treatment and support, more and more people suffering from different critical diseases, including cancer, survive and go about their everyday routine activities. When receiving adequate treatment tailored to their needs, patients can cope with illness' negative effects and hopefully increase their positive experience and QoL aspects. Different personalised services that support individuals in their activities is a modern approach also crucial for medical and health domains (Burmester, 2018; He et al., 2019; Ivanovic et al., 2022; Sim et al., 2020).

QoL aspects are getting essential for cancer survivors (Sim et al., 2020; Tzelvels et al., 2022). Complex patient's data collected from multiple sources are essential for achieving

higher quality of supportive services. Patients' data represent starting point of complete data management, preparation of data for application of powerful AI/ML approaches that lead to better predictions, interventions, treatment recommendations to achieve good health status. However, before applying AI/ML techniques diverse data must be prepared in an adequate way as unclean and incorrect data can lead to wrong results and bad health decisions. Contemporary medical systems should utilise patients' big datasets integrated with QoL instruments, make more power and reliable AI/ML inference mechanisms, support user-centred communication, visualise results of AI predictive models employing effective XAI techniques, improve personalised support, and so on. Such holistic approach should: improve post-treatment patients' health status and QoL; follow-up the patients to meet their needs; but additionally, to support predicting the status of new patients.

In the rest of the section, we will present unique, innovative ASCAPE Open AI Architecture with its essential functionalities. Special attention will be paid to data management from two perspectives: (1) preparation of cancer patients' datasets in standardised form and its transformation in adequate form for application of AI/ML/FL approaches to build predictive models and (2) use of intelligent approaches to realise Dashboard with effective visualisation of predictive models' results tailored for particular patient.

4.1. The ASCAPE – architecture of intelligent system for supporting cancer patients

The ASCAPE architecture is organised in contemporary manner as Cloud/Edge computing framework that supports privacy-preserving AI/ML/FL approaches. It (see Figure 1) consists of ASCAPE Cloud and, generally speaking, arbitrary number of the ASCAPE Edge nodes. Current prototype that is implemented and evaluated includes four Edge nodes, i.e. project partners health institutions. The ASCAPE Cloud coordinates the training of Federated Deep Learning global models on Edge nodes and trains Homomorphic Deep Learning models based on homomorphically encrypted data at Edge nodes. The ASCAPE employs two contemporary privacy preserving ML technologies:

Federated Deep Learning that obtains global models. Such models are built on knowledge derived from data from all Edge nodes that participate in FL without need to transfer the data to the Cloud and training the global model on the local Edge node data.

Homomorphic Deep Learning supports training of Cloud-based models using homomorphically encrypted data from Edge nodes. When such models are used for inference based on homomorphically encrypted input data, they produce an output that can be homomorphically decrypted. Encryption and decryption use a secret key known to the collaborating Edge node but not to the Cloud. This means that encryption and decryption only happen on those Edge nodes sharing the same secret key. The main advantage of such approach is that ASCAPE Cloud may be used for both training and model inference without accessing unencrypted data.

Proposed the ASCAPE Open AI architecture provides the common building blocks for different Medical and Health Information Systems (HISs) and supports personalised medical decisions that physicians make for their patients.

The main building blocks contain a number of components for implementation of a range of functionalities.

HIS-ASCAPE Integration Components allow an existing HIS to send patients' data to ASCAPE and also integrate the ASCAPE widgets and supporting backend code that provides

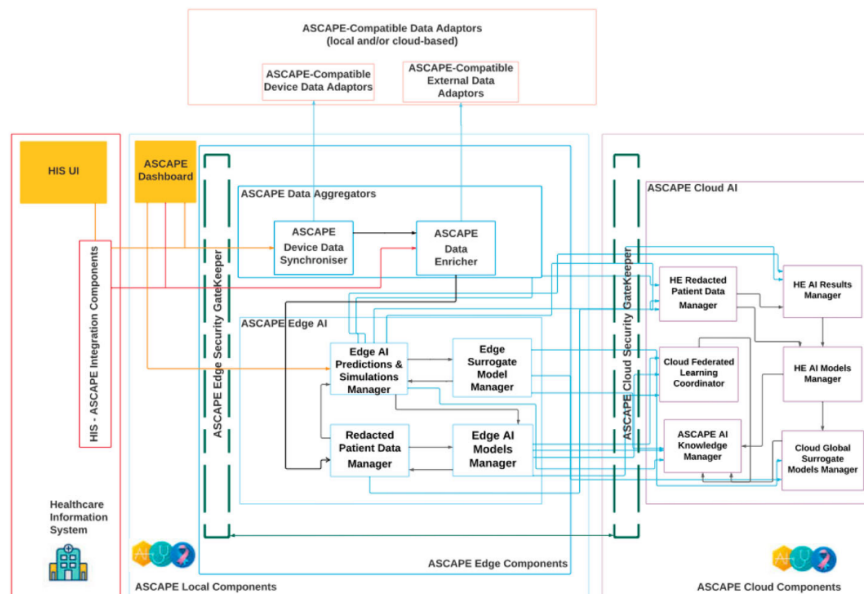


Figure 1. The ASCAPE AI architecture.

the HIS with ASCAPE functionality identical to the stand-alone ASCAPE Dashboard's and also offering physicians the additional ASCAPE benefits.

The ASCAPE Data Aggregators provide support to sending to the ASCAPE Edge node additional patient-related data not collected by the HIS (but rather by ASCAPE-compatible Data Adaptors deployed locally or remotely, like wearable, smart devices).

The ASCAPE Dashboard – if ASCAPE is not sufficiently integrated into the HIS physician may use it in order to access ASCAPE functionality including AI-assisted monitoring of patients' QoL status and recording information about proposed interventions.

The ASCAPE Edge Components – collaborate with the HIS and the Dashboard and with the ASCAPE Cloud which coordinates privacy-compliant collaborative model training with all participating health institutions, i.e. Edge nodes. In such a way, it provides collaboratively training predictive models to all participating health institutions. For Edge node components that interact with the ASCAPE Cloud all interactions are initiated from Edge nodes towards the Cloud in order to fit as best as possible to firewall settings in place at health institutions ICT environments.

The ASCAPE Cloud Components allow privacy-preserving ML technologies on the ASCAPE Cloud: (i) the coordination and storage components for FL, (ii) the training, storage components for model training on homomorphically encrypted data and encrypted predictions and (iii) the components used for collaborative surrogate model training (belonging to used XAI techniques within ASCAPE).

The ASCAPE Security GateKeeper – provides centralised authentication, authorisation and auditing so that all components can communicate through a secure environment. It is a flexible, modular solution to address user-level and component-level access. The main aim is to allow integration with HIS user identity that follows authentication standards.

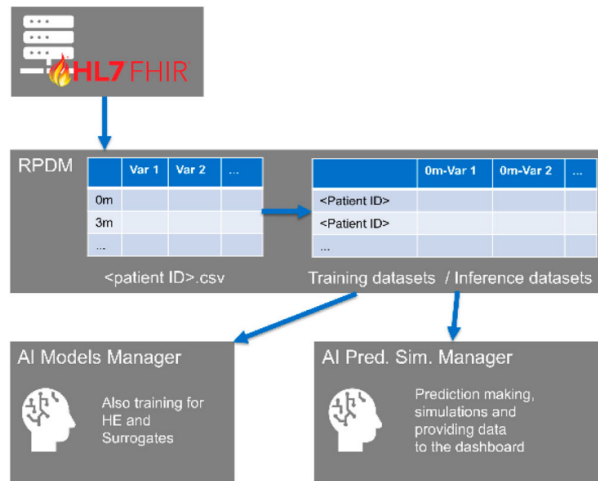


Figure 2. ASCAPE data model based on medical standards.

The *ASCAPE Edge Security GateKeeper* and the *ASCAPE Cloud Security GateKeeper* encompass components deployed at the Edge and at the Cloud respectively.

4.2. The ASCAPE – patients data pre-processing

The component responsible for the storage of all patients' data is *Redacted Patient Data Manager*. The data for a patient consist of data received from the HIS enhanced by additional data (environmental, from smart wearable devices, etc.) that are aggregated. Such data are represented in above mentioned and widely used health standard, i.e. FHIR HL7 compliant format, within the Edge node. The HL7 FHIR standard provides a unified way to store medical and health information in a database and acts as a bridge between the health institution's way to store medical data and the ASCAPE Edge node infrastructure. The HL7 FHIR ensures a consistent way to store data that at the same time is very flexible and can be support almost any type of data.

Data can be identified using different medical and health classification systems, such as International Classification of Diseases (ICD) or SNOMED CT (for details see again Section 3). The latter is used primarily in ASCAPE to identify variables i.e. features in a patient's record (see Figure 2).

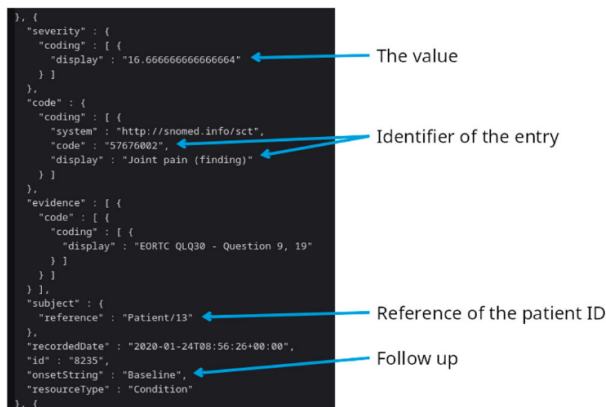
The main reason for the selection of these contemporary medical standards for representation of patients' data in ASCAPE lay in our intention to make the ASCAPE AI architecture open for all other health institutions interested to use it. The ASCAPE architecture offers numerous possibilities for new users like: to use it as is, i.e. as an additional Edge node with their own data; to extend different existing functionalities by introducing new AI/ML models, to use other XAI methods for presentation of obtained results, to implement new privacy preserving techniques, and so on; or extend architecture/components/functionalities for completely new diseases. To make such processes easier, it is extremely important to present data from different source in the same widely recognised interoperable format.

Additional advantage of adopted standards is that FHIR resources can be easily transformed in widely used formats like XML, JSON, Turtle and other facilities to create a clinical data model that includes interoperability artefacts. The importance of such approach is that other health institutions can rather simple use ASCAPE system and its functionalities by representing their datasets using this comprehensive medical and health data terminology standardisation. The ASCAPE architecture is rather easily extendable and can successfully manage several tens and hundreds of new health institutions as it is highly based on emergent distributed Federated Learning paradigm.

On the other hand, the data presented by the HL7 FHIR standard are not directly ready to be used for ML/FL processing and algorithms. The data for ML/FL training must be tabular with the naming and order of features consistently defined across all Edge nodes and common dataset contains all features for both cancer types. As we in the first implementation of ASCAPE Architecture considered four health institutions with different data features/variables for the same cancer type it was necessary to select crucial features/variables for each kind of cancer, target variables for each QoL aspect, variables for a range of possible interventions and produce the common data model. It was successfully done as result of intensive discussions of medical and IT experts. As different types of variables for both cancers finally were adopted additional data pre-processing step was necessary to produce unified record/entry for each patient. Since the data are pre-processed and all nominal variables are one-hot encoded, the possible values of nominal variables were defined consistently.

The *Redacted Patient Data Manager* has a dedicated API that provides endpoints to the FHIR database to ingest or extract data in JSON format. The functionality to convert the FHIR data into tables for training datasets and pre-process them is implemented in a dedicated service. That service uses a configuration, which indicates for each cancer type which list of variables shall be used. This allows flexibility and eases adaptation to new cancer types and possible use of ASCAPE architecture with other health institutions. The ASCAPE service for data export can extract the FHIR data over network via HTTP requests or read FHIR data in JSON files from local storage.

The whole process of training or re-training AI/ML/FL models and inference making is triggered by an export. Exports can be triggered manually via a REST API or scheduled for specific daytimes. The first step of an export is to request the JSON file of all data from the FHIR database. It is structured as a list with each entry corresponding to an entry in the database. The entries differ in their structure and some information might be implicit, for example, a specific diagnosis implying information about receptors in the tumour cell. Therefore, each variable that shall be used in the input data for the AI/ML/FL models is defined in the configuration file. Along each variable, a set of rules is defined that contains patterns that decide if an entry in the exported FHIR data includes information about the respective variable and if yes, what the respective variable is. Also, the patient ID and a date or timestamp (different durations for follow-up examinations: since baseline to M3, M6, M9, M12) are extracted that are needed to write the value into a table for the respective patient. These tables are then processed into datasets for different cancer types and timestamps. After that further pre-processing is applied to convert all data numeric values. Dates are converted to months relative to the 01/01/2000 and nominal categories are one-hot-encoded. The training datasets undergo privacy enhancing methods such as outlier



```

}, {
  "severity": {
    "coding": [ {
      "display": "16.666666666666664"
    } ]
  },
  "code": {
    "coding": [ {
      "system": "http://snomed.info/sct",
      "code": "57676002",
      "display": "Joint pain (finding)"
    } ]
  },
  "evidence": [ {
    "code": [ {
      "coding": [ {
        "display": "EORTC QLQ30 - Question 9, 19"
      } ]
    } ]
  } ],
  "subject": {
    "reference": "Patient/13"
  },
  "recordedDate": "2020-01-24T08:56:26+00:00",
  "id": "8235",
  "onsetTiming": "Baseline",
  "resourceType": "Condition"
}, {

```

Figure 3. Part of a patient's data in structured format.

detection and differential privacy and are then sent in CSV format to the *Edge AI Models Manager* and *Edge Surrogate Models Manager* as well as after being homomorphically encrypted to the ASCAPE Cloud. Part of above describer structure of the data is presented in Figure 3.

For training global models, *HE Redacted Patient Manager* component is responsible. This component receives and stores the homomorphically encrypted training datasets from all Edge nodes. The training datasets can be identified regarding cancer type and target variables and are combined to a single homomorphically encrypted dataset for each cancer type and target variable. These aggregated datasets are then forwarded to the *HE AI Models Manager* for training global homomorphically encrypted predictive models. The homomorphic encryption scheme used in ASCAPE is a variant of the MORE (Matrix Operation for Randomisation or Encryption) homomorphic encryption scheme, which relies on symmetric keys (a 2×2 matrix) and which is innovative research achievement of members of ASCAPE project (Vizitiu et al., 2020). The symmetric key is the same in all Edge nodes where it is used to encrypt the patient data before uploading it the ASCAPE Cloud.

The first prototype of ASCAPE Open AI Architecture has been implemented, and it was intensively trained and evaluated using already available datasets containing data for patients for both types of cancer (retrospective datasets) from three health institutions. In spite the fact that retrospective datasets for breast cancer did not include QoL aspects captured through questionnaires, from information of prescribed medications it was possible to derive the presence of certain QoL aspects (for Anxiety, Depression, Insomnia and Pain). For that purpose, four retrospective datasets (one for each QoL aspect) for breast cancer were produced for training and evaluating binary classifiers deciding whether a patient will suffer from anxiety (additional variable was used BcBase-Anxiety), depression (BcBase-Depression), insomnia (BcBase-Insomnia) and pain (BcBase-Pain) after breast cancer treatment.

For prostate cancer used dataset contained data on patient and cancer characteristics, treatment approaches, side effects based on direct questions or validated questionnaires and QoL aspects based on a validated questionnaire. Patients' health records also contained follow-up examinations repeated at six months intervals and QoL scores at the time

of the diagnosis and three different times relative to the date of diagnosis at months 36, 60 and 120. So, for continual observation of patients' health status and QoL aspects and training predictive models specific datasets were created with following naming scheme DATASET- n - m for training regression models predicting the questionnaire QoL score at month m considering all patient data collected up to month n . Predictive models were intensively trained and evaluated using following retrospective datasets: DATASET-30-36, DATASET-30-60, DATASET-30-120, DATASET-54-60, DATASET-54-120 and DATASET-108-120. For prospective datasets, i.e. data collected from new patients we adopted to examine the patients' health parameters in shorter periods (each three months), so the same dataset naming scheme is used but for period of three months (see more details in Section 5).

More details about used datasets, training predictive AI/ML models and ASCAPE Architecture evaluation can be found in (Savić et al., 2023).

4.3. Brief overview of AI/ML role in the ASCAPE architecture

ASCAPE AI architecture has been implemented, and AI/ML/FL models to support cancer patients' health status and QoL were intensively trained and evaluated using already existing retrospective datasets of two cancer for female and male: breast and prostate. The main intention of the architecture is to be used in real health institutions for new patients and their collected data (prospective datasets). For this purpose, apart from patients' data collected from different sources, several standard validated questionnaires are planned to be used to capture the particular patient's QoL aspect. Medical and health experts from project consortium proposed 15 QoL parameters for breast and 12 for prostate cancer that will be predicted using developed AI/ML/FL models. "To the best of our knowledge, ASCAPE is a unique research project that prospectively investigate an AI-based approach, towards a personalised follow-up strategy for cancer patients focusing on their QoL aspects" (Savić et al., 2023).

Generally speaking, two methodologies are prevalent for training predictive models: centralised and distributed that we also adopted in our architecture. Within centralised approach, patients' data are stored in health institutions locally and models' training is also performed locally. However, as intention of the project is to use the ASCAPE architecture by arbitrary number of health institutions it was necessary to adopt recently very popular decentralised ML technique, Federated learning. By applying FL patients' data collected by health institutions could be adequately processed as it enables the training of shared global models (in Cloud) while keeping all the sensitive data in local (Edge node). To improve quality of raw data, we curated and pre-processed data using two quality techniques: missing value imputation (MVI) and differential privacy (DP). The primary effect of FL is to support democratised access to ML models keeping data locally in health institutions.

The novelty in medical domain of the ASCAPE project is use of real datasets containing QoL aspects derived from prescribed medications and specific (LISAT-11) questionnaires. By our knowledge, we are the first who applied such analytical study of various ML models predicting such QoL aspects. Additional novelty is connected to AI/ML/FL approach, i.e. fact that the selected ML/FL models are examined in two different FL settings apart from centralised one (incremental and concurrent FL). The final contribution of our approach is consideration of impact of pre-processing techniques (MVI and DP) to the quality of explored AI/ML/FL models.

Within centralised approach, we used the following ML algorithms for training classification models predicting binary QoL aspects: NB (Naive Bayes), kNN (k nearest neighbours), SVM (support vector machines), DT (decision trees) and RF (random forests). For predicting numeric QoL aspects, the following algorithms for regression models are used: LINEAR (linear regression), RIDGE (ridge regression), LASSO (lasso regression), ELASTICN (elastic net regression), KRIDGE (kernel ridge regression), SVM (regression by support vector machines), RF (regression by random forests) and kNN (k -nearest neighbours regression). In application of FL, we used two scenarios: incremental and concurrent, based on neural networks as natural model choice (Ilić et al., 2023).

Detailed evaluation of AI/ML/FL approaches using retrospective datasets applied and adopted in the ASCAPE architecture is presented in Ilić et al. (2023), Savić et al. (2023).

4.4. The ASCAPE dashboard and predictive models' results visualisation

There are multiple challenges to be addressed by an AI/ML system that aims to enhance clinical practice. After collection of patient's data from multiple sources (EHR, wearable devices, nutrition and environmental data, etc.) and training and building high-quality predictive intelligent models, their use in real health institutions is essential. They should help physicians for assessment of patients with different stages of cancer (already under therapies, new ones, survivors) and to prediction their QoL after AI/ML models' recommended treatments/interventions/activities during different follow-up periods (three months, six months, etc.) (Deliverable D1.1). However, a good AI/ML/FL functionality with excellent analytical characteristics is not sufficient. Some other important aspects like user experience, integration, security, privacy, and so on also must be considered. The ASCAPE presents an interesting and innovative architecture covering all aforementioned aspects. Specifically, it aims to provide physicians with an AI-powered tool that monitors and predicts the progression of QoL metrics corresponding to overall QoL and specific aspects for a specific patient and offers suggestions for interventions that could improve outcomes. The ASCAPE Dashboard being the primary interface was conceived as a tool that helps physicians better support cancer patients after their treatment by means of effective visual presentation of recorded and predicted data values through efficient and meaningful interaction. In fact on physicians interaction with the ASCAPE system, the details of particular patient are retrieved and visualised, including predictions for their individual and overall QoL aspects and AI/ML/FL suggested interventions.

The comprehensive ASCAPE personalised visualisations widget produced by the ASCAPE system (Deliverable D4.1) is presented in Figure 4.

On the top of the widget, the patient's overall QoL timeline is presented (higher is better) and on the bottom specific psychological, physiological and other QoL aspects timelines (lower is better); just below the overall QoL or QoL aspects timeline the various interventions (non-pharmacological and pharmacological) relevant at each point of time are also visualised (as straight-line segments). In the middle, there is a spider chart (see details in Figure 5) depicting the latest recorded and the predicted values for the various QoL aspects ASCAPE deems relevant to display for the patient and a list of interventions ASCAPE deems relevant. Currently, both these lists depend on the broad cohort of the patient; it presents the ASCAPE personalised visualisations widgets inside the ASCAPE Dashboard page for a fictitious breast cancer patient.

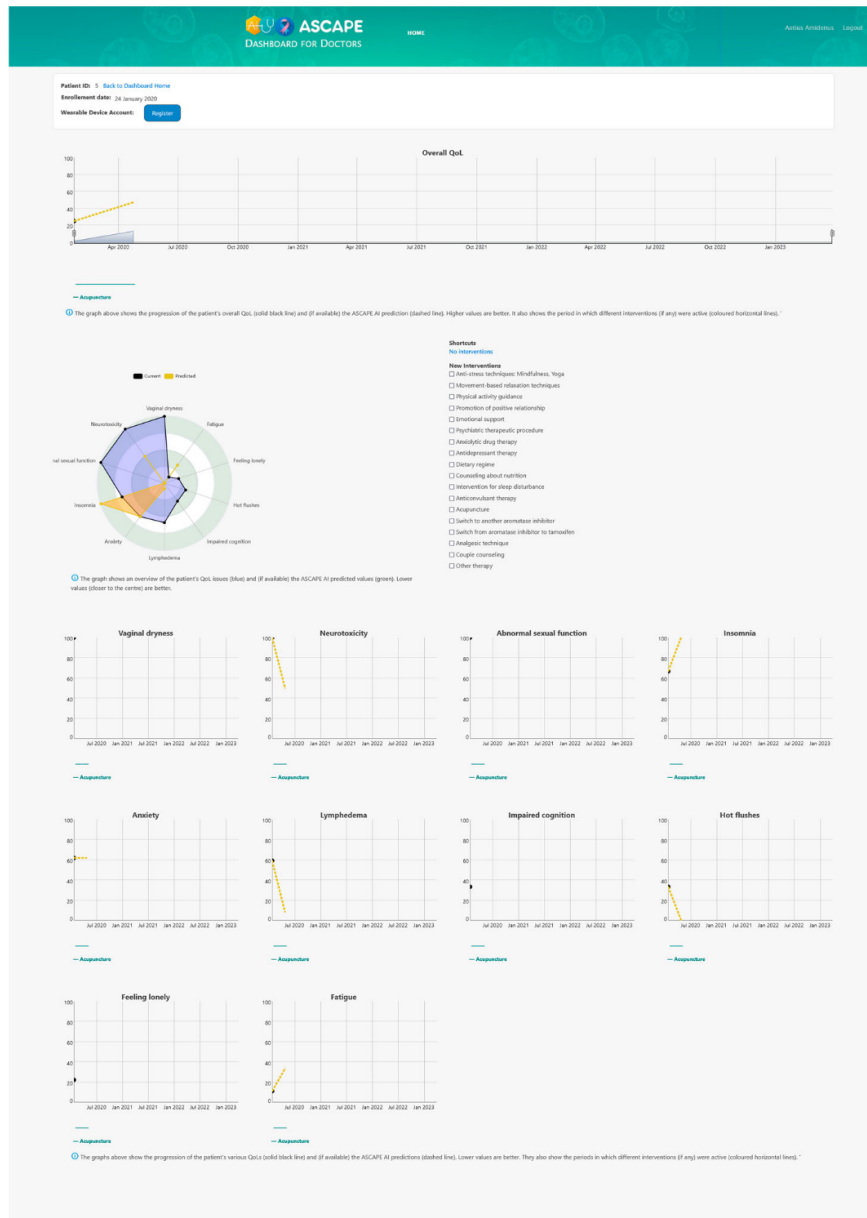


Figure 4. The patient visualisations widget of the ASCAPE system.

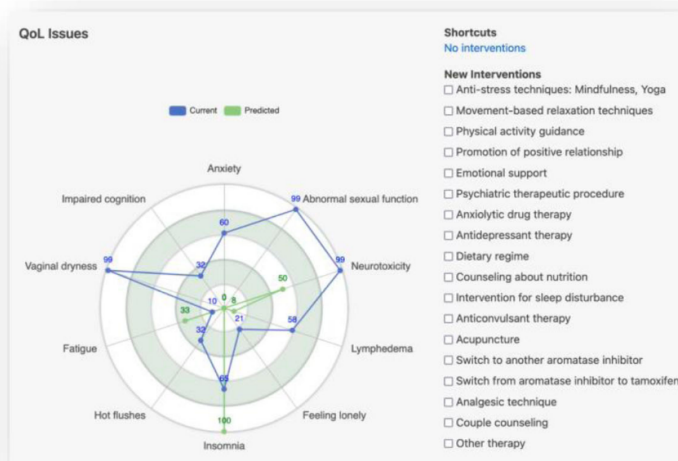


Figure 5. Current and the predicted values for the various QoL aspects with possible interventions.

The ASCAPE patient visualisations widget allows the physicians to get an overview of the patient QoL and the history of interventions without a litany of interactions. The default view provides both recorded data and predictions for the case that any currently active interventions remain so. Physicians can see how different choices of interventions from the previous period affect the predictions for the patient's overall QoL and all QoL aspects simply by clicking on it. This is a simple interaction producing a predictable response from the system. Possible shortcuts make the process even more efficient. These include: "No interventions", "Currently active interventions" and "ASCAPE-Proposed interventions". Selecting a shortcut is tantamount to selecting those interventions corresponding to the shortcut (and no others). The third shortcut selects the interventions the ASCAPE AI/ML models predict will have the greatest overall positive effect on the patient's QoL. One feature of this design is its combination of simplicity, usability and efficiency, both important factors for systems addressed to clinical personnel.

Another feature is that it allows physicians to easily experiment with different options but does not aim to direct them towards choosing the ASCAPE-recommended interventions.

ASCAPE, unlike the majority of similar clinically targeted AI-focused research projects, paid particular attention to providing an easy pathway for integration with existing systems. Part of this effort relates to the user interface (supported by XAI techniques) (Deliverable 2.4). The widget discussed and likewise the widget showing a summary of the current and predicted QoL aspects status can easily be embedded into existing HIS physicians are already using. Additionally, the physician may at any time allow the patient to see ASCAPE's visualisation and explain to him/her why he/she recommends a particular intervention. Also, the physician can discuss with the patient which interventions the patient most agrees with and is most likely to follow.

The ASCAPE priority is that health institutions on the one hand maintain control of their patient data and on the other are able to collaborate on building AI/ML models capturing knowledge from multiple hospitals' patients.

The key components of the ASCAPE architecture that highly support functioning of Dashboard are briefly presented below.

The first group are Edge components.

Edge AI Models Manager – This component is responsible for training local AI/ML/FL models in Edge nodes and for training global models with local data in collaboration with the ASCAPE Cloud, as well as for analytically evaluating models and choosing the ones that best fit local data. For each training dataset received from the *Redacted Patient Data Manager*, several types of models are trained both on local data only as well in federated manner orchestrated by the *Cloud Federated Learning Coordinator* (for details see Ilić et al., 2023; Savić et al., 2023). All locally trained models are stored in the component as well as any global model obtained from the ASCAPE Cloud. The quality of the models is evaluated over the locally available datasets evaluating widely used metrics.

Edge Surrogate Models Manager – This component is responsible for training local surrogate models (in the current implementation linear regression and decision trees as representative XAI techniques are implemented) and for training global surrogate models for global predictive models with using the local data in collaboration with the ASCAPE Cloud (Deliverable D2.4). These surrogate models are trained to make the same predictions as the primary models (of the Edge AI Models Manager) but due to their nature (e.g. decision tree models) lend themselves to being used for explaining these predictions.

Edge AI Predictions & Simulations Manager – This component uses the locally available models (local models or global models formed via federated learning) as well as the HE models at the ASCAPE Cloud to produce QoL-related predictions and intervention suggestions to the HIS and/or the Dashboard. The used models are those with the best evaluation over the local data and the predictions from the HE models are obtained by sending encrypted patient-specific inference requests to the ASCAPE Cloud and decrypting locally the received encrypted prediction. Furthermore, the component is responsible to compute feature attributions in form of Shapley Values to allow to visualise the impact of the different features on the predicted target values (Holzinger et al., 2022).

In addition to computing predictions and explanations, the component also pre-computes intervention suggestion: the goal is to use the predictive capabilities of trained models and interventions of any kind of the patient and selected by the health institutions partners to provide for each patient with suggestions of interventions that have a positive effect on the predicted value. This is performed by simulations estimating the treatment effect of interventions (possibly also combinations of different interventions) and provide that information for retrieval by the ASCAPE Dashboard (Deliverable 4.1) to show it to the physician treating the patients, which can then take a decision.

The second group are Cloud components.

Cloud Federated Learning Coordinator – This component coordinates the federated training of global predictive models based on the patient data available at each participating ASCAPE Edge node. The same type of models as locally is trained in federated manner for classification and regression tasks. Specific training schemes were designed that allow for flexible addition and removal of Edge node. The federated training is not initiated by the

Cloud Federated Learning Coordinator, but rather by the Edge nodes. If an Edge node needs a specific model and no global model is available in *Cloud Knowledge Manager*, it starts training locally and sends it as a first instance to the *Cloud Federated Learning Coordinator*. If a global model is available, the Edge node updates it with its local training data and submits it again to the Cloud (incremental training mode). If the *Cloud Federated Learning Coordinator* detects that more than one Edge node want to train a model, it switches to semi-concurrent mode, where training happens in several rounds by collecting the trained or updated model from each Edge node, creating an aggregated model by averaging and provide that model to all Edge nodes for the next training round. All final trained global models are then forwarded to the *Cloud Knowledge Manner*.

Cloud Knowledge Manager – This component stores all available final global models on the Cloud, from which they can be retrieved by the Edge nodes. This way new Edge nodes entering the federation can benefit from models previously trained on data from all other available ASCAPE Edge nodes.

HE AI Models Manager – This component stores all models trained on the aggregated homomorphically encrypted datasets. They can be retrieved by the *HE AI Results Manager* to provide encrypted predictions on encrypted inference requests submitted from the edge components.

HE AI Results Manager – The *HE AI Results* component receives all encrypted inference requests for predictions from the different Edge nodes. Based on the type, it retrieves the corresponding model from the *HE AI Models Manager*. If the model is not yet available, it waits until the model is available. The encrypted prediction is stored in the component in order to be retrieved by the Edge node that submitted the request. The inference requests can be of different kinds: of course, any inference request in the Edge node is also submitted to this component. However, during the computation of SHAPLEY values and the training of surrogate models, further requests are created by the edge components and submitted to this component in order to determine these for the HE models.

Cloud Global Surrogate Models Manager – This component coordinates all activities to train global surrogate models (belonging to XAI techniques), both for those obtained on plain text via FL as well as for global HE models. The training is initiated as soon as an Edge node requests a surrogate model which is not yet trained. The *Cloud Global Surrogate Models Manager* then initiates the training both for linear regression and decision tree models. Meanwhile, the *Edge Surrogate Model Manager* creates the local training for the surrogate models by taking the local training dataset used for the global model, but labelling it using the predictions of the global model; in case this is an HE model, the labelling consists of submitting appropriate encrypted inference requests to the *HE AI Results Manager* and labelling the local dataset by decrypting the encrypted answer.

The training of linear regression surrogate models essentially works like the FL of normal models, except that it uses the training datasets annotated with the predictions. Training of decision tree surrogate models is more involved, as separate training or update of models and aggregating via averaging is not possible. Hence, for this the computation consists of two iterations using a voting system. In the first iteration, each Edge node trains a decision tree from scratch and sends it back to the Cloud, where all decision trees are collected. In the second iteration, each Edge node evaluates each surrogate model from every Edge node using its own local evaluation dataset. The decision tree with the best overall score across all datasets is then used as the resulting surrogate model.

5. ASCAPE architecture – discussion on current state and future possible use

The ASCAPE architecture brings together Edge Computing with State-of-the-Art privacy-preserving AI techniques. It combines Federated Deep Learning, Homomorphic Encryption and Homomorphic Deep Learning, Differential Privacy, Outlier Detection and Explainable AI approaches, together with finding original and innovative solutions, producing a coherent AI framework (Lampropoulos et al., 2021; Ilić et al., 2023; Savić et al., 2023; Vizitiu et al., 2020). Additionally, it allows data to remain under the control of health institutions inside their ICT infrastructures, on their local ASCAPE Edge nodes. At the same time, the ASCAPE Cloud is central point for deriving knowledge from local data it does not have access to. Open AI network allows arbitrary number of local ASCAPE Edge nodes and the ASCAPE Cloud can lead to the creation of cross-AI models of medical knowledge trained on more data that a single health institution has, nevertheless how large and well-equipped, can obtain solely from its patients (Deliverable D1.1). This knowledge can be made available to other large or small health institutions and even individual physicians. Such rich knowledge can be accessed with the help of user interfaces designed to facilitate physicians' interaction with the ASCAPE AI functionality, taking into consideration their need to have quick access to relevant information (Deliverable D4.1).

The ASCAPE architecture covers all areas of functionality necessary to achieve the desired effect of integrating ASCAPE into clinical practice: data synchronisation (allowing the local ASCAPE Edge node to have access to health and medical relevant patient data), continuous learning (allowing local data to be transformed into cross-silo AI knowledge), obtain AI analytics on the basis of the created models and patient data (either locally or on the Cloud with patient data homomorphically encrypted) and presenting those results to physicians in a meaningful and helpful manner through powerful XAI techniques.

The architecture has been designed to work with a wide range of HIS, to allow different applications to be supported, beyond the breast and prostate cancer, to interoperate with different kinds of devices and open data sources, and even to support different AI techniques in addition to the ASCAPE implementations of Federated and Homomorphic Deep Learning.

Within three years of ASCAPE project duration, we faced different challenges and learned useful lessons while approaching to final implementation of first prototype. First version of the ASCAPE Open AI architecture that encompasses a number of predictive ML/FL models has been implemented and intensively evaluated and tested using retrospective datasets that were available from project's health institutions.

After defining common data model as consequence of selecting variables and characteristic features/attributes for both kinds of cancer, evaluation of the architecture has been done using anonymised retrospective datasets (Deliverable 2.4) while we are waiting for collecting prospective data for new patients. First part of data management was consisted of following steps: available retrospective datasets were transformed in HL7 FHIR format; then a CSV file for each patient was produced and cluster patients with the same type of cancer were created and organised in a dataset; datasets were pre-processed applying One-Hot-Encoding, Missing Value Imputation and Differential Privacy to prepare data for predictive ML/FL models training for all ASCAPE ML/FL models. Selected ML/FL models were trained, evaluated and used for predictions of QoL aspects with satisfactory level of accuracy and reliability (Savić et al., 2023).

Further development and evaluation of the architecture were continued by using prospectively collected datasets. As number of patients in these datasets was rather small and we had not enough data for different prediction periods and for different durations: since baseline to M3, M6, M9, M12; since M3–M6, M9, M12, and so on, the ASCAPE Architecture and their functionality evaluations were conducted on synthetically generated datasets derived from available prospective datasets. In spite the fact that synthetically generated datasets did not realistic and at satisfactory level of confidence this evaluation also showed that the ASCAPE predictive models produce good performance and achieve promising results.

Another main intention of ASCAPE data management was to present patient's data processing results in intelligent and friendly way (explained in Section 4.4). For explainability of model predictions, the SHAP framework was identified as the best basis for feature attribution. Surrogate models showed overall good to excellent approximation of the selected target model. To suggest the best treatments for particular patient, a simulation method based on the calculation of average treatment effects using cohort matching and exploiting trained QoL predictive models was adopted to be used for prospectively collected data (Deliverable D2.4).

At the moment, initial testing using real prospectively collected data is going on in one of the project's health institutions. Full use of the ASCAPE architecture is expecting to happen in near future using prospectively collected datasets from all four project's health institutions. These datasets consist of the collection of three main types of data. The first part includes patient's clinical data comprises demographics, cancer-related data, medications, QoL aspects, interventions and other medical conditions. The second type of data is Weather data, consisting of temperature and daylight duration information, which is linked to each patient based on their postal code or city name. The third type of data is Activity and related data is linked to each patient and is collected using FitBit wearable devices. It consists of daily steps, activity time, calories burned, heart rate, resting heart rate, sleep efficiency and sleep duration.

The ASCAPE Architecture was designed to be easily extensible and new health institutions can use its functionalities. The creation of an ASCAPE Edge node inside a health institution's ICT infrastructure is a prerequisite for the use of ASCAPE Architecture. Each health institution joining the ASCAPE ecosystem, needs to set up an ASCAPE Edge node and runs ASCAPE software that has access to (redacted and pseudonymised) patient data. An ASCAPE Edge node is responsible for:

1. Receiving from the health institution's systems, storing and making available in an interoperable standard format (FHIR) redacted and pseudonymised patient data.
2. Training local and global FL models using the health institution's patient data (if the Edge node has been configured to participate in FL).
3. Sending the health institution's redacted and pseudonymised patient data homomorphically encrypted to the ASCAPE Cloud (if the Edge node has been configured to participate in Homomorphic Learning).
4. Training surrogate models using local data and the predictions of the respective local or global FL models and Homomorphic Learning models.
5. Making predictions, running simulations and generating results to be presented to physicians via the ASCAPE Dashboard or other similar tools.

6. Ensuring only authorised users/systems have access to the above functionality, securing communications and other security and privacy enhancing functionalities.

In spite the fact that after three years of the project duration we achieved significant, innovative and valuable results we were constantly facing different challenges and needed to solve them for the benefit of all participants and future users of our system. The following are some of key lessons learned worth highlighting:

- The interdisciplinary nature of the work reveals a long and constructive evolutionary approach and discussions to share different understandings and views on medical and health data from physicians' point of view and ICT/technical point of view. Tight long-lasting cooperation and explanations of different understandings of the same aspects of medical and health data were challenging for both sides.
- In spite the fact that project participants came from different countries with different cultural and social background, all of them, i.e. medical and ICT experts finally recognised the necessity and challenges of introducing different, above mentioned and presented, complex technological innovations into the real medical environments and benefits that they can bring to physicians and patients in future exploitation.
- The diverse nature of each health institution, existence of different initial features/variables in patients' datasets, different forms of patients' records (paper documentation, simple xls records, full data representation in HL 7 FHIR format), different timelines for collecting data for new patients caused rather stressful working atmosphere and complicated a range of project activities. However, during constant meetings based on agile methodology and comprehensive explanations, we agreed about common data model and variables and characteristic features/attributes for both kinds of cancer.
- As patients' health and medical data are very sensitive and require privacy protection according to General Data Protection Regulation (EU GDPR), enabling permissions to use retrospective datasets from Ethical committees of project's health institutions was long-lasting and complex procedure that caused postponing initial activities at least three to four months.
- Different health institutions possessed completely incompatible hardware and software in their local HIS and settings. Such diversity caused a lot of problems to install necessary production environment in HIS Edge nodes to satisfy the ASCAPE Architecture requirements.
- In spite the fact that highly reliable privacy preserving methods (DP and HE) were adopted in ASCAPE some medical experts from health institutions did not allow sending neither retrospective nor prospective data to the technical partners to train and evaluate AI/ML/FL predictive models. So, specific workflows were created that caused extremely time-consuming process of training and evaluation of a range of models and accordingly preparation of results to be presented to physicians via Dashboard.
- We definitively faced a lot of complications and challenges that we should have solved during the last couple of months. However, the real-world prospective data collected about patients already provide some useful insights about their needs, which are rarely captured in the normal clinical setting. This experience is valuable for physicians in cancer patients' treatment. However, still there is some level of disinclination of some medical experts to adopt this new approach and suspicion in reliability of presented

results for the patient. Honestly, it is reasonable and expectable as final decision should be done by physician after careful analysis of results obtained by ASCAPE ML predictive models.

- The last but not the least is existence of rather standard problem in real medical and health studies that we also faced. Contrary to the initial expectations that we will be able to recruit more than 600 new patients and collect their data during the second and the third year of the project finally we must be satisfied with about 150 patients. This rather small number of patients with different stages of illness, therapies and follow ups, of course is not enough to achieve highly credible and trustworthy results of the ASCAPE ML predictive models at the moment. However, in near future, our projection is that more health institutions will be willing to use our architecture and accordingly the better performances and more promising results of the ASCAPE ML predictive models are expected.

6. Conclusion

The aim of the paper was to present innovative ASCAPE Open AI Architecture with its crucial functionality and especially to put focus on patients' data management, i.e. pre-processing data as input to AI/ML models training and evaluation, but also presenting achieved results in more user-friendly manner to physicians and caregivers. Now the whole system has been evaluated using already existing patient' data (retrospective datasets). As very promising results were obtained in such a way, we are continuing our research activities with evaluation of developed architecture on data collected from new patients (prospectively collected data) expecting accurate and reliable behaviour of developed architecture.

On the other hand, rapid technological development is fundamental for implementing sophisticated and personalised medical services especially based on advances and challenges that AI/ML methods constantly bring. Development of more and more refined AI/ML algorithms influence high-quality data management and processing in obtaining more robust and effective medical predictions. The numerous challenges in applying ML in medical and health domains are oriented towards data collection and management. Any ML model depends on high-quality data, so formulating effective data management, pipelines for ML data processing and user-friendly interfaces to present achieved results becomes essential requirements in modern HIS (Alanazi, 2022).

In line with data preparation for application of advanced AI/ML approaches is privacy preserving as very prominent issue. Apart from a range of already existing methods necessity to leverage privacy-enhancing technologies is essential to reap the benefits of AI/ML while minimising the risk of data violations (Link 5). Apart from traditional methods, emerging privacy-enhancing technologies that will significantly increase privacy and security of sensitive medical and health data in the future include: Differential privacy, Homomorphic encryption, Secure multi-party computation, Zero-knowledge proofs, and so on. Within ASCAPE architecture, we verified effectiveness of Differential privacy and Homomorphic encryption approaches as guarantee for privacy preserving of sensitive patients' data.

Another critical challenge in intelligent medical and health systems is that prediction based on ML usually does not provide clear reasons for that, especially for non-experts. At the moment, some promising techniques of XAI support better understanding of ML predictive models' results/outcomes and increase reliability in such system. Accordingly, in

ASCAPE architecture we experimented and obtained very encouraging results using XAI techniques for presentation of results at Dashboard. Positive ASCAPE results and achievements will further direct medical research and practice in very prominent directions (Lee, Braud, et al., 2021): powerful health analytics and predictive modelling (He et al., 2019), data visualisation techniques supported by advanced XAI approaches for communication (Holzinger et al., 2022) between different users (patients, physicians, medical stakeholders), personalised therapies, recommendations and interventions.

Newest concepts like avatars, metaverse (Lee, Braud, et al., 2021), holographic constructions (Kairouz et al., 2019; Lee, Braud, et al., 2021) possess a high potential and can significantly influence future development of holistic, sophisticated medical systems (Lloret et al., 2015; Salih & Abraham, 2016). So further development of ASCAPE could go into direction of extension of existing architecture by such novel approaches, especially in riching user-friendly interface and Dashboard.

However, the near future is not so optimistic (Link 2) concerning wide use and exploitation of intelligent medical and health systems in real environments. There are a lot of problems (like diverse, limited and distributed patients' data sources; satisfactory but not fully reliable AI/ML models; rather slow big data processing mechanisms; integration of wide variety of multiple AI services) and personalised medicine has limitations that make it not revolutionary but evolutionary research domain.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was supported by the ASCAPE project. The ASCAPE project has received funding from the European Union's Horizon 2020 Research and Innovation Framework Programme under grant agreement No. 875351.

ORCID

Mirjana Ivanovic  <http://orcid.org/0000-0003-1946-0384>

Serge Autexier  <http://orcid.org/0000-0002-0769-0732>

References

- Alanazi, A. (2022). Using machine learning for healthcare challenges and opportunities. *Informatics in Medicine Unlocked*, 30, paper 10092430. <https://doi.org/10.1016/j.imu.2022.100924>
- Autexier, S., Lüth, C., & Drechsler, R. (2022). Chapter the Bremen ambient assisted living Lab and beyond – smart environments, smart services and artificial intelligence in medicine for humans. In M. A. Pfannstiel (Ed.), *Künstliche Intelligenz im Gesundheitswesen* (German Title: Das Bremen Ambient Assisted Living Lab und darüber hinaus – Intelligente Umgebungen, smarte Services und Künstliche Intelligenz in der Medizin für den Menschen). Springer Fachmedien Verlag Wiesbaden, March. pp 835–850, https://doi.org/10.1007/978-3-658-33597-7_40
- Burmester, G. R. (2018). Rheumatology 4.0: Big data, wearables and diagnosis by computer. *Annals of the Rheumatic Diseases*, 77(7), 963–965. <https://doi.org/10.1136/annrheumdis-2017-212888>
- Chida, Y., Hamer, M., Wardle, J., & Steptoe, A. (2008). Do stress-related psychosocial factors contribute to cancer incidence and survival? *Nature Clinical Practice Oncology*, 5(8), 466–475. <https://doi.org/10.1038/ncponc1134>. Epub 2008 May 20. PMID: 18493231.

- Claeys, A., & Vialatte, J. S. (2014). Les progrès de la génétique: vers une médecine de précision? Les enjeux scientifiques, technologiques, sociaux et éthiques de la médecine personnalisée [Advances in genetics: Towards a precision medicine? Technological, social and ethical scientific issues of personalised medicine].
- Cohen, S., Murphy, M. L. M., & Prather, A. A. (2019). Ten surprising facts about stressful life events and disease risk. *Annual Review of Psychology*, 70(1), 577–597. <https://doi.org/10.1146/annurev-psych-010418-102857>. Epub 2018 Jun 27. PMID: 29949726; PMCID: PMC6996482.
- (Deliverable D1.1). ASCAPE Deliverable – D1.1 Positioning ASCAPE's open AI infrastructure in the after cancer-care Iron Triangle of Health. <https://ascapenproject.eu/node/57>
- (Deliverable D2.3) D2.3. System architecture. <https://www.bd2decide.eu/deliverables/d23-system-architecture>
- (Deliverable D2.4). ASCAPE Deliverable – D2.4 ML-DL Training and Evaluation Report. <https://ascapenproject.eu/node/118>
- (Deliverable D4.1). ASCAPE Deliverable – D4.1 Personalized interventions and user-centric visualizations. <https://ascapenproject.eu/node/120>
- (Deliverable D5.1) D5.1. Multilayer data acquisition and management services. <https://www.bd2decide.eu/deliverables/d51-multilayer-data-acquisition-and-management-services>
- Fan, Z. Y., Yang, Y., Zhang, C. H., Yin, R. Y., Tang, L., & Zhang, F. (2021). Prevalence and patterns of comorbidity among middle-aged and elderly people in China: A cross-sectional study based on CHARLS data. *International Journal of General Medicine*, 14, 1449–1455. <https://doi.org/10.2147/IJGM.S309783>
- Ficek, J., Wang, W., Chen, H., Dagne, G., & Daley, E. (2021). Differential privacy in health research: A scoping review. *Journal of the American Medical Informatics Association*, 28(10), 2269–2276. <https://doi.org/10.1093/jamia/ocab135>
- Gallo, P., Aso, S., Autexier, S., Brotons, A., De Nigro, A., Jurak, G., Kiourtis, A., Kranas, P., Kyriazis, D., Lustrek, M., & Magdalino, A. (2019). CrowdHEALTH: Big data analytics and holistic health records. *Studies in Health Technology and Informatics*, 258, 255–256. PMID: 30942764.
- He, J., Baxter, S. L., Xu, J., Xu, J., Zhou, X., & Zhang, K. (2019). The practical implementation of artificial intelligence technologies in medicine. *Nature Medicine*, 25(1), 30–36. <https://doi.org/10.1038/s41591-018-0307-0>
- Hiremath, S., Yang, G., & Mankodiya, K. (2014). Wearable internet of things: Concept, architectural components and promises for person-centered healthcare. In 2014 4th International Conference on Wireless Mobile Communication and Healthcare-Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH), pp. 304–307, IEEE.
- Holzinger, A., Saranti, A., Molnar, C., Biecek, P., & Samek, W. (2022). Explainable AI methods-a brief overview. In A. Holzinger, R. Goebel, R. Fong, T. Moon, K. R. Müller, W. Samek (Eds.), *xxAI - Beyond Explainable AI* (pp. 13–38, Vol. 13200). *xxAI 2020. Lecture Notes in Computer Science*. Springer. https://doi.org/10.1007/978-3-031-04083-2_2
- Ilić, M., Ivanović, M., Jakovetić, D., Kurbalija, V., Otlokan, M., Savić, M., & Vujanović-Sedlar, N. (2023). ASCAPE – An intelligent approach to support cancer patients. WorldCist'23 – 11st World Conference on Information Systems and Technologies, to be held in Pisa, Italy, 4–6 April 2023.
- Ivanovic, M., Autexier, S., & Kokkonidis, M. (2022). AI approaches in processing and using data in personalized medicine. In S. Chiusano, T. Cerquitelli, & R. Wrembel (Eds.), *Advances in databases and information systems. ADBIS 2022. Lecture notes in computer science* (Vol. 13389, pp. 11–24). Springer. https://doi.org/10.1007/978-3-031-15740-0_2
- Ivanovic, M., & Balaz, I. Influence of artificial intelligence on personalized medical predictions, interventions and quality of life issues. ICSTCC 2020 In 24th International Conference on System Theory, Control and Computing, ICSTCC 2020, Sinaia, Romania, IEEE 2020, ISBN 978-1-7281-9809-5: 445450.
- Ivanović, M., & Ninković, S. (2017). Personalized HealthCare and agent technologies. 11th KES International Symposium on Agent and Multi-Agent Systems, Technologies and Applications, Vilamoura, Portugal, 21–23 June, pp. 3–11, Springer.
- Ji, Z., Lipton, Z. C., & Elkan, C. (2014). Differential privacy and machine Learning: A survey and review. arXiv:1412.7584 [cs.LG]. <https://doi.org/10.48550/arXiv.1412.7584>

- Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, M. A. N., Bonawitz, K., Charles, G., Cormode, Z., Cummings, R., D'Oliveira, R. G. L., Eichner, H., Rouayheb, S. E., Evans, D., Gardner, J., Garrett, Z., Gascón, A., Ghazi, B., Gibbons, P. B., ... Jag, M. (2019). Advances and open problems in federated learning. arXiv preprint arXiv:1912.04977.
- Kaissis, G. A., Makowski, M. R., Rückert, D., Braren Rickmer F. (2020). Secure, privacy-preserving and federated machine learning in medical imaging. *Nature Machine Intelligence*, 2(6), 305–311. <https://doi.org/10.1038/s42256-020-0186-1>
- Kalloniatis, C., Lambrinouidakis, C., Musahl, M., Kanatas, A., & Gritzalis, S. (2021). Incorporating privacy by design in body sensor networks for medical applications: A privacy and data protection framework. *Computer Science and Information Systems*, 18(1), 323–350. <https://doi.org/10.2298/CSIS200922057K>
- Kingston, A., Robinson, L., Booth, H., Knapp, M., Jagger, C., & for the MODEM project (May 2018). Projections of multi-morbidity in the older population in England to 2035: Estimates from the population ageing and care simulation (PACSim) model. *Age and Ageing*, 47(3), 374–380. <https://doi.org/10.1093/ageing/afx201>
- Kumar, S., Rana, M., Verma, K., Singh, N., Sharma, A., Maria, A., Singh, G., Khaira, H., & Saini, S. (2014). Predigt-cx: Post treatment health related quality of life prediction model for cervical cancer patients. *PLoS One* 9(2), e89851. <https://doi.org/10.1371/journal.pone.0089851>
- Kyriazis, D., Autexier, S., Brondino, I., Boniface, M., Donat, L., Engen, V., Fernandez, R., Jimenez-Peris, R., Jordan, B. & Jurak, G. (2017). Crowdhealth: Holistic health records and big data analytics for health policy making and personalized health. *Informatics Empowers Healthcare Transformation*, 238, 19–23. <https://doi.org/10.3233/978-1-61499-781-8-19>
- Lahiri, C., Pawar, S., & Mishra, R. (2019). Precision medicine and future of cancer treatment. *Precision Cancer Medicine*, 2, 33, AME Publishing. <https://doi.org/10.21037/pcm.2019.09.01>
- Lampropoulos, K., Kosmidis, T., Autexier, S., Savic, M., Athanatos, M., Kokkonidis, M., Koutsouri, T., Vizitiu, A., Valachis, A., & Padron, M. Q. ASCAPE: An open AI ecosystem to support the quality of life of cancer patients. In Proceedings of ICHI 2021 – 9th IEEE International Conference on Healthcare Informatics, pp. 301–310.
- Lee, D., Cornet, R., Lau, F., & De Keizer, N. (2013). A survey of SNOMED CT implementations. *Journal of Biomedical Informatics*, 46(1), 87–96. <https://doi.org/10.1016/j.jbi.2012.09.006>
- Lee, L. H., Braud, T., Zhou, P., Wang, L., Xu, D., Lin, Z., Kumar A., Bermejo C., & Hui, P. (2021). All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda. arXiv preprint arXiv:2110.05352.
- (Link 1). H2020 project. <https://www.bd4qol.eu/wps/portal/site/big-data-for-quality-of-life>
- (Link 2). <https://www.england.nhs.uk/cancer/living/>
- (Link 3). Follow-up medical care. <https://www.cancer.gov/about-cancer/coping/survivorship/follow-up-care>
- (Link 4). Healthcare classifications and terminologies. <https://guides.library.kumc.edu/c.php?g=451743&p=3084407>
- (Link 5). Healthcare, artificial intelligence, top 6 challenges of AI in healthcare & how to overcome them. <https://research.aimultiple.com/challenges-of-ai-in-healthcare/>
- Lloret, J., Canovas, A., Sendra, S., & Parra, L. (2015). A smart communication architecture for ambient assisted living. *IEEE Communications Magazine*, 53(1), 26–33. <https://doi.org/10.1109/MCOM.2015.7010512>
- Norbye, A., Abelsen, B., Førde, O., & Ringberg, U. (2022). Distribution of health anxiety in a general adult population and associations with demographic and social network characteristics. *Psychological Medicine*, 52(12), 2255–2262. <https://doi.org/10.1017/S0033291720004122>
- Park, J., & Lim, H. (2022). Privacy-preserving federated learning using homomorphic encryption. *Applied Sciences*, 12(2), 734. <https://doi.org/10.3390/app12020734>
- Saadat, S., Aziz, A., Ahmad, H., Imtiaz, H., Sohail, Z., Kazmi, A., Aslam, S., Naqvi, N., & Saadat, S. (2017). Predicting quality of life changes in hemodialysis patients using machine learning: Generation of an early warning system. *Cureus*, 9. <https://doi.org/10.7759/cureus.1713>
- Salih, A., & Abraham, A. (2016). *Ambient intelligence assisted healthcare monitoring* (p. 192). LAP LAMBERT Academic Publishing.

- Savić, M., Kurbalija, V., Ilić, M., Ivanović, M., Jakovetić, D., Valachis, A., Autexier, S., Rust, J., & Kosmidis, T. (2023). The application of machine learning techniques in prediction of quality of life features for cancer patients. *Computer Science and Information Systems*, 20(1), 381–404. <https://doi.org/10.2298/CSIS220227061S>
- Schulz, S., Stegwee, R., & Chronaki, C. (2019). Standards in healthcare data. In P. Kubben, M. Dumontier, & A. Dekker (Eds.), *Fundamentals of clinical data science*. Springer. https://doi.org/10.1007/978-3-319-99713-1_3
- Siddique, M., Mirza, M. A., Ahmad, M., Chaudhry, J., & Islam, R. (2018). A survey of big data security solutions in healthcare. In *International Conference on Security and Privacy in Communication Systems*, pp. 391–406. Springer, Cham.
- Sim, J., Kim, Y., Kim, J., Lee, J., Kim, M. S., Shim, Y., Zo, J., & Yun, Y. H. (2020). The major effects of health-related quality of life on 5-year survival prediction among lung cancer survivors: Applications of machine learning. *Scientific Reports*, 10(1), Article 10693. <https://doi.org/10.1038/s41598-020-67604-3>
- Sinha, R., & Heuvel, W. (2011). A systematic literature review of quality of life in lower limb amputees. *Disability and Rehabilitation*, 33(11), 883–899. <https://doi.org/10.3109/09638288.2010.514646>
- Tyler, N. S., Mosquera-Lopez, C. M., Wilson, L. M., Dodier, R. H., Branigan, D. L., Gabo, V. B., Guillot, F. H., Hiltz, W. W., El Youssef, J., Castle, J. R., & Jacobs, P. G. (2020). An artificial intelligence decision support system for the management of type 1 diabetes. *Nature Metabolism*, 2(7), 612–619. <https://doi.org/10.1038/s42255-020-0212-y>
- Tzelves, L., Manolitsis, I., Varkarakis, I., Ivanovic, M., Kokkonidis, M., Useros, C. S., Kosmidis, T., Muñoz, M., Grau, I., Athanatos, M., Vizitiu, A., Lampropoulos, K., Koutsouri, T., Stefanatou, D., Perrakis, K., Stratigaki, C., Autexier, S., Kosmidis, P., & Valachis, A. (2022). Artificial intelligence supporting cancer patients across Europe – The ASCAPE project. *PLoS One*, 17(4), e0265127. <https://doi.org/10.1371/journal.pone.0265127>
- Venne, J., Busshoff, U., Poschadel, S., Menschel, R., Evangelatos, N., Vysyaraju, K., & Brand, A. (2020). International consortium for personalized medicine: An international survey about the future of personalized medicine. *Personalized Medicine*, 17(2), 89–100. <https://doi.org/10.2217/pme-2019-0093>
- Vizitiu, A., Nita, C. I., Puiu, A., Suciu, C., & Itu, L. M. (2020). Applying deep neural networks over homomorphic encrypted medical data. *Computational and Mathematical Methods in Medicine*, 2020, 3910250:1–3910250:26. <https://doi.org/10.1155/2020/3910250>
- Wu, M., & Luo, J. (2019). Wearable technology applications in healthcare: A literature review. *The Online Journal of Nursing Informatics*, 23(3). <https://www.himss.org/resources/wearable-technology-applications-healthcare-literature-review>
- Yang, Z., Olszewski, D., He, C., Pintea, G., Lian, J., Chou, T., Chen, R. C., & Shtylla, B. (2021). Machine learning and statistical prediction of patient quality-of-life after prostate radiation therapy. *Computers in Biology and Medicine*, 129, 104127. <https://doi.org/10.1016/j.combiomed.2020.104127>