NOVA SCHOOL OF SCIENCE & TECHNOLOGY

DEPARTMENT OF ELECTRICAL
AND COMPUTER ENGINEERING

NUNO MIGUEL DO CARMO NUNES

BSc in Electrical and Computer Engineering

KNOWLEDGE BASE
FOR ONCOLOGY PATIENT CARE

MASTER IN ELECTRICAL AND COMPUTER ENGINEERING

NOVA University Lisbon
April, 2024

DEPARTMENT OF ELECTRICAL
AND COMPUTER ENGINEERING

KNOWLEDGE BASE
FOR ONCOLOGY PATIENT CARE

NUNO MIGUEL DO CARMO NUNES

BSc in Electrical and Computer Engineering

Adviser: Pedro Alexandre da Costa Sousa
Associate Professor, NOVA University Lisbon

Co-adviser: João Paulo Branquinho Pimentão
Assistant Professor, NOVA University Lisbon

Examination Committee

Chair: Luís Filipe Lourenço Bernardo
Associate Professor, NOVA University Lisbon

Rapporteur: José António Barata de Oliveira
Full Professor, NOVA University Lisbon

Adviser: Pedro Alexandre da Costa Sousa
Associate Professor, NOVA University Lisbon

MASTER IN ELECTRICAL AND COMPUTER ENGINEERING

NOVA University Lisbon
April, 2024

Knowledge Base for Oncology Patient Care

This document was created with the (pdf/Xe/Lua)LaTeX processor and the NOVAthesis template (v7.0.3) [**novathesis-manual**].

Para a minha família.

Acknowledgements

Firstly I would like to extend my deepest appreciation to my advisors, Pedro Sousa and João Paulo Pimentão, for accepting me into this project, and for providing their guideance and mentorship both within and beyond its scope, their enthusiasm and passion for the field is inspiring.

Special thanks to Dra. Maria Torrente for her expertise and feedback throughout the project.

I am grateful to all my colleagues at Holos for their warm welcome, ongoing support, and the invaluable knowledge they've shared.

To my group of friends, thank you for all your support, this journey wouldn't have been nearly as fulfilling without you by my side.

I also want honor the memory of my grandparents. Although they could not see me finish this chapter, their values and guidance have shaped the person I am today.

And above all, I would like to extend my most heartfelt gratitude to my parents, for always believing in me, and to my brother, who sparked my passion for technology. I would not be where I am without them and I am eternaly grateful. I owe my success to them.

"All's well that ends well."

— John Heywood

Abstract

Within the current healthcare environment, patient-centered care stands as a core principle, emphasizing the active involvement of patients in their care journey to enhance treatment outcomes and overall patient satisfaction. Fulfilling this premise requires robust communication mechanisms that foster information exchange between healthcare providers and patients and reciprocally, between patients and healthcare providers, as they navigate their healthcare journey.

This accentuates the need for intelligent platforms to share pertinent information. Despite advancements in digital technologies and the increasing availability of healthcare information, challenges persist in accessing and disseminating crucial data across healthcare systems. This dissertation aims to address these challenges by developing knowledge base modules within the CORE Cardio-Oncology web application to promote effective information exchange between patients and healthcare providers.

These modules serve as an intermediary between patients and healthcare professionals, facilitating the exchange of information across various types of data such as questions and answers, medication management and tracking of quality of life indicators like sleep patterns, blood pressure and exercise routines. Each module possesses distinct functionalities that enhance the overall interactivity. Together, these modules form an integrated system that is designed to support patients by providing comprehensive resources and tools to enhance their well-being and quality of life.

By working closely with the Oncology Department of Puerta de Hierro Majadahonda University Hospital, it was assured that the developed solutions were aligned with the specific needs and challenges faced by patients. This collaboration ensured that the interventions were not only effective but also tailored to enhance patient outcomes and elevate the overall quality of care of users.

Keywords:

Association Rules, Cloud Computing, Knowledge Base, Knowledge Discovery, User Interfaces, Web Development

Resumo

No ambiente atual dos cuidados de saúde, os cuidados centrados no paciente destacam-se como um princípio fundamental, realçando o envolvimento ativo dos pacientes na sua jornada de cuidados para melhorar os resultados dos tratamentos e a satisfação geral dos pacientes. Para cumprir esta premissa, são necessários mecanismos robustos de comunicação que promovam a troca de informações entre os profissionais de saúde e os pacientes e reciprocamente, entre os pacientes e os profissionais de saúde, à medida que seguem com os seus tratamentos. Isto sublinha a necessidade de plataformas inteligentes para compartilhar informações pertinentes. Apesar dos avanços nas tecnologias digitais e da crescente disponibilidade de informações de saúde, persistem desafios no acesso e na disseminação de dados cruciais nos sistemas de saúde. Esta dissertação tem como objetivo abordar estes desafios ao desenvolver módulos para a knowledge base da aplicação web CORE Cardio-Oncology para promover a troca eficaz de informações entre pacientes e profissionais de saúde.

Estes módulos servem como intermediários entre pacientes e profissionais de saúde, para a troca de informações em diversos tipos de dados, como perguntas e respostas, gestão de medicação e acompanhamento de indicadores de qualidade de vida, como padrões de sono, pressão arterial e rotinas de exercícios. Cada módulo possui funcionalidades distintas que fornecem uma maior interatividade geral. Juntos, estes módulos formam um sistema integrado projetado para apoiar os pacientes, ao fornecer recursos e ferramentas abrangentes para melhorar seu bem-estar e qualidade de vida. Ao estabelecer uma colaboração com o Departamento de Oncologia do Hospital Universitário Puerta de Hierro Majadahonda, foi assegurado que as soluções desenvolvidas estavam alinhadas com as necessidades e desafios específicos enfrentados pelos pacientes. Esta colaboração garantiu que as intervenções fossem não só eficazes, mas também adaptadas para melhorar os resultados dos pacientes e elevar a qualidade geral do cuidado prestado aos utilizadores.

Palavras-chave:

Regras de Associação, Computação Cloud, Knowledge Base, Knowledge Discovery, Interfaces de Utilizador, Desenvolvimento Web

Contents

List of Figures

List of Tables

Glossary

association rules Techniques that discover relationships and patterns in datasets. *(pp. 3, 24, 52)*

backpropagation A key algorithm used in training neural networks that involves calculating and adjusting the weights and biases of the network based on the difference between the predicted output and the actual output, propagating this error backward through the network to update the parameters in a way that minimizes the error. *(p. 17)*

biomarkers Measurable indicators or characteristics that are used to evaluate biological processes, disease presence, or response to treatment. *(pp. 11, 12)*

data mining The process of extracting valuable insights or knowledge from large datasets using various techniques such as machine learning, statistical analysis, and pattern recognition. *(p. 24)*

knowledge base A centralized repository of structured information and knowledge within a specific domain or subject area, used for efficient storage, retrieval, and management of information. *(pp. v, vii, 3, 6, 41, 69)*

knowledge discovery The process of identifying, extracting, and interpreting useful knowledge or patterns from large volumes of data. It encompasses various techniques and methodologies from fields such as data mining, machine learning, statistics, and information retrieval. Knowledge discovery aims to uncover hidden insights, trends, and relationships within data that can inform decision-making, improve understanding, or support prediction and inference tasks. *(pp. 3, 43)*

neural networks Computational models inspired by the structure and function of the human brain which consist of interconnected nodes, or artificial neurons, organized in layers, that process and transform input data to produce desired outputs, enabling the network to learn and make predictions or decisions without being explicitly programmed. *(p. 17)*

pipeline A series of connected processes or stages through which data or information flows. Each stage in the pipeline performs a specific task or transformation, and the output of one stage becomes the input for the next, allowing for efficient and modular data processing. *(pp. xi, 16, 17)*

web sockets A communication protocol that provides full-duplex communication channels over a single TCP connection, allowing for interactive communication between a client (such as a web browser) and a server. *(p. 48)*

Acronyms

AI Artificial Inteligence *(pp. ix, xi, 3, 6, 9–15, 17–20, 43)*
APIs Application Programming Interfaces *(p. 22)*
AR Augmented Reality *(pp. 14, 19)*
ARM Association Rule Mining *(pp. 24, 25)*

BaaS Backend as a Service *(p. 40)*

CT scans Computerized Tomography Scans *(p. 15)*

DOM Document Object Model *(p. 42)*

ePHR Electronic Personal Health Record *(p. 5)*

FaaS Function as a Service *(p. 40)*

HTTP Hypertext Transfer Protocol *(pp. 40, 46, 52, 63)*

IaaS Infrastructure as a Service *(pp. 39, 40)*
IAM Identity and Access Management *(p. 48)*

JSON JavaScript Object Notation *(pp. 48, 49, 52)*

mHealth apps Mobile Health Applications *(p. 8)*
MRI Magnetic Resonance Imaging *(p. 15)*

NLP Natural Language Processing *(pp. xi, 6, 10, 15–17)*

PaaS Platform as a Service *(pp. 39, 40, 42)*

Q&A Questions and Answers *(pp. 3, 48, 52, 55, 60, 63, 70)*

SaaS Software as a Service *(pp. 39, 40)*

UI User Interface *(p. 42)*
UMLS Unified Medical Language System *(p. 10)*

1

Introduction

1.1 Background

In the realm of modern healthcare, patient-centered care has emerged as a fundamental principle, emphasizing the importance of involving patients as active participants in their own care journey [56]. With the advent of digital technologies and the increasing availability of healthcare information, patients are now empowered to take charge of their health and make informed decisions [19]. However, despite these advancements, the process of accessing and sharing crucial information within the healthcare ecosystem remains a significant challenge.

Hospitals play a pivotal role in delivering comprehensive healthcare services. However, according to [32], the current state of information sharing in hospitals often falls short of meeting patients' expectations and needs. Fragmented communication channels, lack of interoperability between systems and limited patient engagement create barriers that impede the seamless flow of information among healthcare providers, patients and caregivers. Effective information sharing is crucial for public health initiatives as well [37], as public health heavily relies on data reported by healthcare stakeholders, which enables efficient reporting, investigation, emergency response and communication within the clinical community [60]. Simplifying and streamlining information exchange processes can greatly enhance public health outcomes and foster collaboration among stakeholders involved in patient care. By addressing these challenges, information flow can be optimized and healthcare services can be improved, leading to better patient outcomes.

Patients consistently report enhanced engagement and improved retention and comprehension when exposed to shared information, as revealed by multiple studies conducted across various organizations, resulting in a greater understanding of their care plans, diagnostic tests and referrals [43].

Moreover, patients often find themselves encountering questions and uncertainties regarding their healthcare journey that they desire to have addressed [22]. The availability of comprehensive and easily accessible information plays a crucial role in addressing these concerns and fostering patient satisfaction. When patients have access to accurate and timely information, they can seek clarification on treatment options, potential risks and benefits and other aspects of their care plan. By actively involving patients in the decision-making process and providing them with the necessary information, healthcare providers can empower patients to make well-informed choices and actively participate in their own healthcare management.

Furthermore, medication nonadherence among patients remains a significant issue, with studies indicating that patients are nonadherent to their medication approximately 50% of the time. Physicians often attribute nonadherence primarily to factors such as lack of access or forgetfulness [11]. Non-adherence to medication regimens can result in increased morbidity and higher healthcare costs [57]. Addressing medication nonadherence through interventions aimed at enhancing patient education, and implementing reminders or monitoring systems can mitigate the adverse effects associated with nonadherence and improve patient outcomes.

In addition, routine assessment of patients' quality of life has an impact on physician-patient communication and results in emotional end health benefits for patients [70] as they often experience various aspects of their well-being that warrant attention and understanding throughout their healthcare journey. This regular assessment enables healthcare providers to gain deeper insights into the multifaceted aspects of patients' well-being, ranging from physical discomfort to lack of sleep and exercise, thereby facilitating more empathetic and patient-centered care approaches, fostering a more collaborative relationship with their healthcare providers. Consequently, patients are more likely to articulate their concerns and preferences, leading to more tailored and effective treatment plans that address their individual needs and priorities.

1.2 Problems

Promoting patient-centered care is a challenge. Patients are inundated with vast amounts of information from various channels, which frequently results in feelings of confusion and anxiety [74]. This issue underscores the need for patient-centered care, necessitating a solution that addresses communication barriers and empowers individuals to seek personalized guidance and clarity at their own pace. Additionally, the complexity of treatment regimens poses another significant obstacle, with intricate medication schedules coupled to the fact that nonadherence to medication plans can often be an intentional choice made by the patient [11]. This nonadherence also extends to the registration of quality of life indicators. The necessity for ongoing assessments and interventions throughout the care journey further underscores the importance of addressing barriers to patient engagement and adherence.

1.3 Contributions

In order to promote patient-centered care, key modules contributing to the knowledge base of a cardio oncology web application were developed. The first module is a designated Questions and Answers (Q&A) page that facilitates the exchange of information between healthcare professionals and patients, serving as a platform where patients can pose their questions.

To further enhance the information exchange process, the Q&A page within the proposed software solution is designed to leverage knowledge discovery AI capabilities. By utilizing association rules, the Q&A page aims to recommend relevant questions to users, streamlining the information-seeking process for patients. The system can analyze previous user interactions to generate tailored question suggestions that align with the specific needs and interests of individual patients. This AI-driven recommendation feature serves as a valuable tool for facilitating efficient and effective information exchange between healthcare workers and patients, enabling users to obtain targeted and personalized insights while minimizing potential barriers associated with formulating accurate and relevant inquiries. By leveraging AI in this manner, the software strives to optimize the patient-centered information exchange experience, promoting enhanced engagement, knowledge acquisition and overall satisfaction.

The second module consists of a medication management page, providing users with the capability to input and track their medication intake. This information is seamlessly made available to healthcare professionals, offering a comprehensive overview of patients' treatment adherence and progress. Users can also get medication intake reminders through Google Calendar, enhancing their ability to adhere to prescribed medication regimens. This integration not only adds a practical dimension to the application but also contributes to a holistic approach to patient care by addressing medication adherence, a crucial aspect of treatment success.

The third essential module entails a quality of life indicators registration and monitoring page, allowing users to input data associated with various aspects of their daily life such as sleep patterns, cardiovascular data, exercise routines, pain levels, and overall well-being. Through intuitive interfaces, users can effortlessly record their data, which is then visualized through interactive graphs that depict the evolution of these indicators over time. These graphical representations provide users and physicians with valuable insights into health status indicators and allow for easy interpretation of trends and patterns.

These components serve as an integral part of the web application known as CORE[1]. This project's front page is shown in figure 1.1.

CORE is a versatile web application designed to support patients across various medical conditions, emphasizing its utility in Oncology and Cardiology. It offers a comprehensive array of resources, tools, and support to assist individuals in managing their health and improving their overall well-being.

[1]https://core.holos.pt/

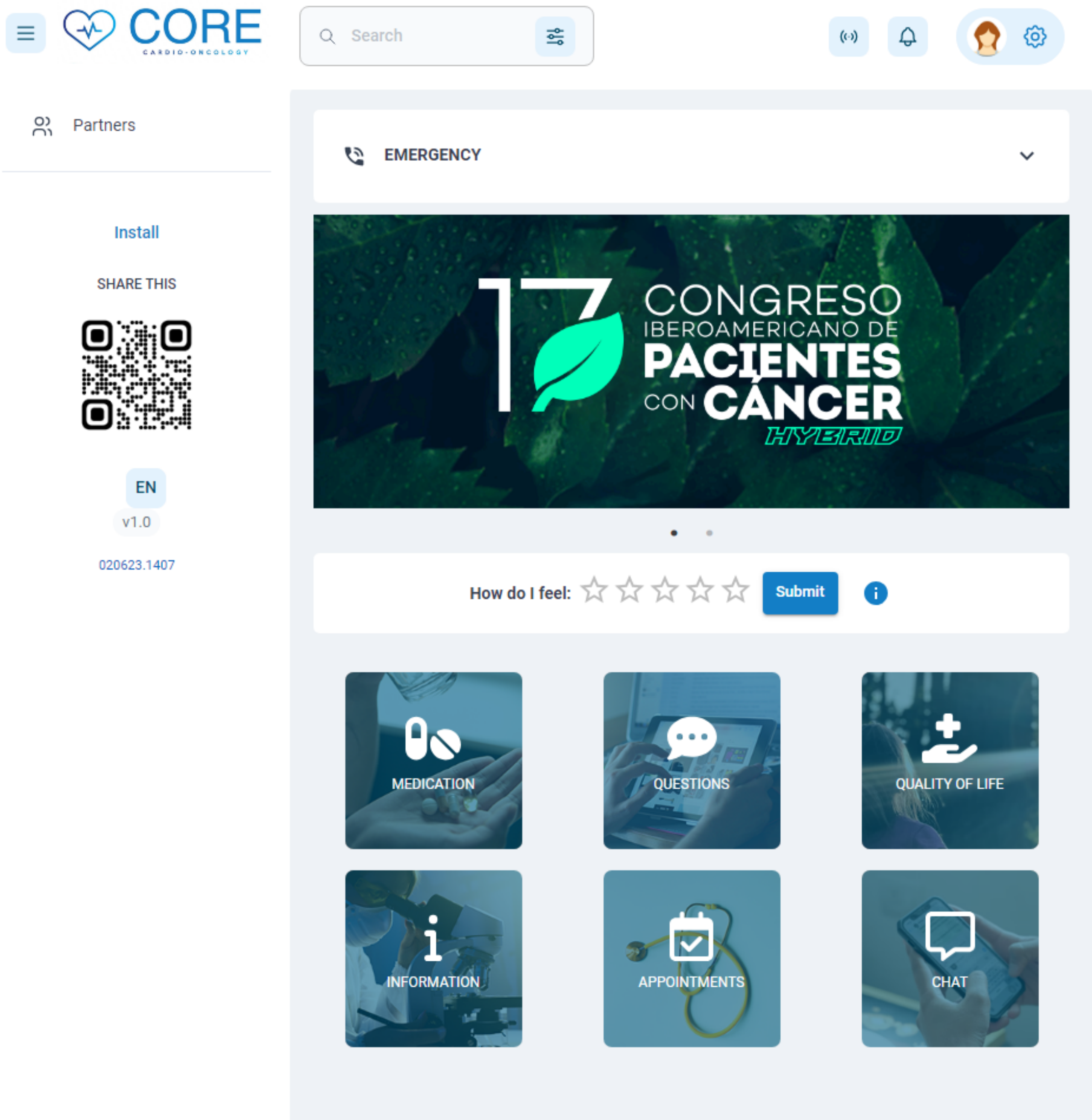

Figure 1.1: Project CORE - Front page.

CORE provides a diverse range of features and functionalities aimed at empowering patients to navigate their healthcare journey effectively. It offers access to educational materials covering different medical conditions, treatment options, and potential side effects, ensuring patients have access to reliable information to make informed decisions about their health.

The primary objective of this initiative is to enhance the well-being and quality of life for individuals affected by various health conditions. By developing intelligent information-sharing platforms, CORE aims to provide user-friendly and accessible resources, including essential medication management features and a wealth of educational materials tailored to the specific needs of patients across different medical domains.

2

State of the Art

2.1 Patient-Centered Information Sharing

The need for platforms that facilitate information sharing and communication with healthcare patients arises from the increasing emphasis on patient engagement and empowerment in healthcare. These platforms serve as valuable tools for promoting patient-centered care and enabling patients to actively participate in their healthcare decisions.

One important feature of patient engagement platforms is interactive, near real-time communication. This allows patients to directly connect with their healthcare organization and engage in two-way communication [54], enabling them to ask questions, seek clarifications and receive timely responses from healthcare providers, enhancing their overall engagement and satisfaction.

2.1.1 Patient-Centered Information Sharing Technologies

Individual patients have distinct healthcare needs, prompting healthcare providers to offer platforms that cater to personalized and tailored information. These platforms enable patients to inquire, share pertinent information about their conditions and receive customized responses. The most relevant examples of these platforms are:

- **Patient Portals**: web-based platforms that provide individuals with controlled access to their personal health information and various healthcare services. These portals are typically provided by healthcare organizations, such as hospitals, clinics, or health systems, to enhance patient engagement, facilitate communication and empower individuals to actively participate in their healthcare management. Given that patient portals did not gain widespread use until 2006 when several initiatives coincided, including the launch of the Electronic Personal Health Record (ePHR) by Microsoft and Google [33], these initiatives aligned with the prevailing social movement of embracing advanced information and communication sharing tools, such as smartphones and social media, in daily routines. This movement reflects the general population's willingness to adopt technology as a means of interactive social engagement.

- **Chatbots and Generative AI**: Chatbots and generative Artificial Inteligence (AI) solutions, such as ChatGPT and Google Gemini are systems designed to simulate human conversation. They interact with users through text-based or voice-based interfaces, providing automated responses and performing various tasks. They are commonly used in customer service, virtual assistants and other applications where human-like interaction is required [51]. These systems' key components are:

  – Natural Language Processing (NLP): They utilize NLP techniques to understand and interpret user inputs. NLP enables the tool to extract meaning from text or speech, identify the user's intent and generate appropriate responses. It involves tasks such as language understanding, sentiment analysis and entity recognition.

  – Dialog Management: They employ dialog management techniques to maintain a coherent conversation flow. They can handle multi-turn interactions, remember previous context and respond appropriately to user queries or prompts. Dialog management ensures a seamless and engaging user experience.

  – Knowledge base: They often rely on a knowledge base or database of information to provide accurate and relevant responses. This knowledge base may include structured data, pre-defined responses, or access to external sources of information. They can retrieve and present information from the knowledge base in a user-friendly manner.

  They offer numerous advantages that contribute to enhanced user experiences and organizational efficiency. Some key benefits of these systems include:

  – Accessibility: They are available 24/7, allowing users to access information and services at their convenience, without the need to wait for human assistance.

  – Scalability: They can handle multiple conversations simultaneously, making them scalable for businesses or organizations with high customer or user interaction volumes.

  – Efficiency: They automate routine tasks and inquiries, reducing the workload on human agents. They can provide an almost instant responses, saving time for both users and businesses.

  – Consistency: They deliver consistent and standardized information and responses, ensuring a uniform user experience across different interactions.

  However, it's important to note that chatbots and generative AI solutions have limitations. They may struggle with understanding complex or ambiguous queries, lack empathy and emotional intelligence and may not be suitable for highly sensitive or critical situations that require human intervention. According to [51], they prove most useful in less significant decisions, such as scheduling doctors appointments and locating health clinics, where more than 70% of interviewed physicians stated that these systems cannot provide detailed diagnosis and treatment because of not knowing all of the personal factors associated with the patient, also stated in [10] and that they could be a risk to patients if they self-diagnose too often.

- **Telemedicine**: the use of telecommunications technology to provide healthcare services remotely. It allows healthcare professionals to evaluate, diagnose and treat patients without the need for in-person visits [72]. Through telemedicine, patients can connect with healthcare providers using video conferencing, telephone calls, or other digital communication platforms. Telemedicine encompasses a wide range of medical services, including consultations, follow-up appointments, prescription refills and monitoring of chronic conditions. It enables patients to access healthcare from the comfort of their homes, eliminating the need for travel and reducing waiting times. Other significant benefits include:

  – Increased accessibility to healthcare for individuals in rural or underserved areas [73].

  – Minimizes the potential for patients to encounter pathogens in a clinical or hospital setting.

  – Potentially helps in reducing healthcare costs, as it minimizes travel expenses and can potentially reduce the number of hospital admissions or emergency room visits.

  – Facilitates better management of chronic conditions through remote monitoring and regular communication with healthcare providers. It enables timely intervention, early detection of complications and continuous support for patients in managing their health.

  While telemedicine offers numerous benefits, it also suffers from significant downfalls, from technical challenges, given that it relies on technology (internet connectivity, device malfunctioning, audio or video lag) to the lack of critical physical examinations, where some medical conditions may require hands-on assessment, palpation or diagnostic tests that cannot be performed remotely, limiting the accuracy of diagnoses and potentially resulting in missed or delayed diagnoses.

- **Mobile Health Applications**: software applications designed to run on smartphones, tablets, or other mobile devices to support health-related activities and facilitate healthcare information sharing. These apps are specifically developed to provide users with tools and resources to manage their health and well-being conveniently through their mobile devices [46]. They offer a wide range of features and functionalities that can assist users in various aspects of health management.

They can include features such as fitness tracking, allowing users to monitor their physical activity levels, track steps, calories burned and set exercise goals. Some Mobile Health Applications (mHealth apps) also incorporate features for monitoring vital signs, such as heart rate, blood pressure and sleep patterns, using sensors or wearable devices [53], enabling them to detect any changes or abnormalities and gain insights into their overall well-being. Furthermore, mHealth apps enable users to manage their medications effectively, providing reminders for medication intake, dosage instructions and alerts for potential drug interactions or refills, helping users stay organized with their medication schedules and reduces the likelihood of missed doses or incorrect usage. Despite the numerous advantages offered by mHealth apps, it is essential to acknowledge and address the limitations and challenges associated with their use. By understanding these factors, one can work towards optimizing their effectiveness and ensuring their integration into healthcare systems. Some of these challenges include:

– The dependence on the Internet connection to provide updated information [53].

– Their effectiveness relies on user engagement and adherence. Users must actively utilize the features, input accurate data and follow recommended guidelines for optimal outcomes.

– Ensuring the privacy and security of user data is of utmost importance, where developers should implement robust measures to protect sensitive health information.

Table 2.1 compares the aforementioned technologies.

Table 2.1: Comparison of Patient-Centered Healthcare Technologies

Technology	**Main Features**	**Appropriate use Cases**	**Implementation Challenges**
Patient Portals	Access to personal health records	Appointment scheduling, prescription refills	Technical barriers, limited functionality
Chatbots	24/7 availability for patient inquiries with real-time responses	Basic medical questions, appointment reminders	Limited ability for complex cases
Telemedicine	Remote healthcare interaction with a specialist	Specialist consultations, follow-up care	Provider training and adoption
Mobile Health Applications	Health information access, monitoring	Self-management, health tracking	Dependence on user engagement and adherence challenges

2.2 Practical implementations of AI in the healthcare industry

AI has emerged as a disruptive and transformative technology that is reshaping multiple industries, including healthcare [75]. The healthcare sector, with its vast and complex data landscape, stands to benefit significantly from AI applications. AI has the potential to revolutionize patient care, research and administration [15, 68] by leveraging advanced algorithms to analyze large datasets, identify patterns and generate valuable insights. The integration of AI in healthcare holds promise for improving diagnostic accuracy, optimizing treatment plans, accelerating drug discovery, enhancing patient engagement and streamlining healthcare operations. This sub-chapter will delve into the diverse applications of AI in the healthcare industry, exploring how it is driving innovation, transforming healthcare delivery and advancing the field towards more personalized and efficient care. By harnessing the power of AI, healthcare organizations can unlock new opportunities to improve patient outcomes, enhance clinical decision-making and ultimately, advance the overall quality of healthcare services.

2.2.1 Healthcare domains impacted by AI advancements

AI has found diverse and wide-ranging applications in the healthcare industry, leading to a paradigm shift in multiple facets of patient care, research and administration. The integration of AI technologies has ushered in a new era of possibilities, enabling significant advancements and improvements across various areas. The following paragraphs outline some key AI technologies making substantial contributions in the domains of healthcare. Figure 2.1 shows different branches of AI and the areas of healthcare they have had an impact on.

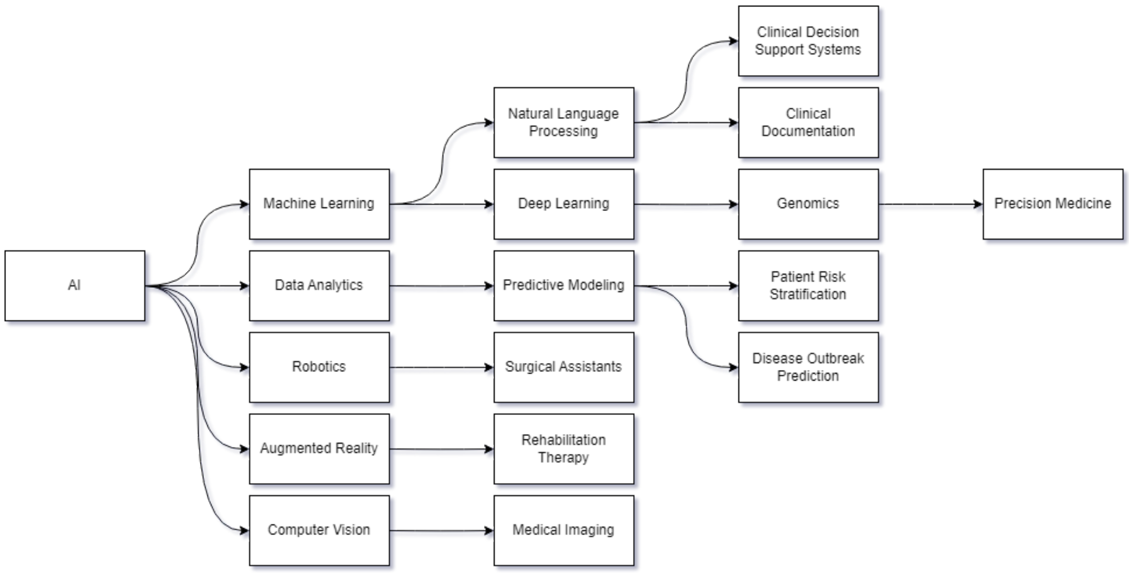

Figure 2.1: An overview of the impact of AI on different branches of healthcare, synthesized from [16, 24, 20, 17, 36, 9, 29, 71].

2.2.1.1 Clinical Decision Support Systems

Clinical Decision Support Systems leverage the power of NLP and AI to enhance healthcare decision-making processes, they consist of computer-based tools that assist healthcare professionals in analyzing patient data, medical knowledge and clinical guidelines to provide evidence-based recommendations in order to improve patient care. NLP plays a crucial role in clinical decision support systems by enabling the system to understand and process human language in various forms, including clinical notes, medical literature and patient records. NLP techniques extract relevant information from unstructured text, such as symptoms, diagnoses, medications and treatment plans. By transforming unstructured data into structured, machine-readable formats, NLP enables the interpretation and analysis of patient information effectively. Using machine learning and deep learning algorithms it is possible for the these tools to learn from datasets and identify patterns and correlations in patient data. AI models can recognize complex relationships between symptoms, diagnoses, treatments and outcomes, allowing clinical decision support systems to provide personalized recommendations and predictions, all while reducing errors and providing alerts for potential medication interactions or allergies [16]. Projects such as the EU-funded CLARIFY[1] healthcare software solution enhance patient medication management by providing advanced features that include warnings for possible medication conflicts.

2.2.1.2 Clinical Documentation

Clinical documentation is a critical aspect of healthcare, encompassing the recording and communication of patient information throughout the care continuum. NLP enables the analysis and understanding of unstructured clinical text, such as physician notes, diagnostic reports and discharge summaries. The algorithms extract relevant information from these documents, including patient demographics, medical conditions, medications, procedures and treatment plans. By automatically structuring and codifying this information, NLP facilitates the creation of comprehensive and standardized clinical documentation. These algorithms can learn from annotated data to classify and extract specific information from clinical text, as an example, AI models can automatically identify medical codes, such as Unified Medical Language System (UMLS) codes from clinical documents [24], reducing the burden of manual coding. AI can also assist in automating documentation tasks, such as generating progress notes or extracting pertinent information for research and quality improvement initiatives.

These tools enable healthcare professionals to focus more on direct patient care rather than spending excessive time on documentation tasks. Additionally, they improve the accuracy and completeness of clinical documentation by reducing human errors and ensuring consistent and standardized information capture, which, in turn, enhances the

[1] https://www.clarify2020.eu/

quality of healthcare data, which is essential for clinical research, decision-making and reporting.

2.2.1.3 Precision Medicine - Genomics

Precision medicine, a novel approach to healthcare, encompasses the consideration of individual variability in genes, environment and lifestyle when formulating tailored treatment strategies and disease prevention protocols. Genomics, the field dedicated to studying an individual's genetic information, assumes a pivotal role in precision medicine by unraveling the intricate complexities of the genome. Deep learning, a branch of AI, has emerged as a potent computational tool within genomics research, enabling the extraction of profound insights from vast and complex genomic datasets [20].

Genomics involves analyzing an individual's DNA sequence and identifying genetic variations that may influence disease susceptibility, drug response and overall health. Precision medicine leverages this genetic information to provide personalized treatment options and preventive measures. By considering a person's unique genetic profile, healthcare professionals can make more informed decisions about the most effective interventions, medications and dosages.

Deep learning, a machine learning technique inspired by the structure and function of the human brain, has revolutionized genomics research. Deep learning algorithms can analyze vast amounts of genomic data, identifying patterns and relationships that may not be readily apparent to human researchers. These algorithms can automatically extract features, such as genetic markers, from raw genomic data, allowing for more accurate disease risk prediction, prognosis assessment and treatment response prediction. Deep learning models have been employed in various genomics applications, including variant calling, gene expression analysis and genomic sequence classification [20]. These models can uncover complex associations between genetic variations and disease phenotypes, leading to the discovery of novel biomarkers and therapeutic targets. Additionally, deep learning can integrate genomic data with other clinical, environmental and lifestyle factors to provide a more comprehensive understanding of individual health and disease.

The combination of precision medicine, genomics and deep learning has the potential to transform healthcare, enabling the identification of subpopulations with similar genetic profiles, allowing for targeted interventions and personalized therapies. Deep learning models can assist in the interpretation of genomic data, aiding in the prediction of disease risk and treatment response for individual patients, holding promise for improving patient outcomes, reducing adverse drug reactions and optimizing healthcare resource allocation.

2.2.1.4 Patient Risk Stratification

Patient risk stratification is a process that involves categorizing individuals into different risk groups based on their likelihood of developing certain health conditions or experiencing specific outcomes, such as disease progression or treatment response. Predictive modeling, based on data analytics and AI plays integral roles in risk stratification by leveraging patient data to generate models and algorithms that can accurately predict and quantify individual risk profiles. It uses statistical and machine learning techniques to develop models that can forecast future events or outcomes based on historical data. In the context of patient risk stratification, predictive modeling utilizes a wide range of patient variables, including demographic information, medical history, genetic factors and clinical biomarkers to create models that estimate the likelihood of specific health outcomes [44]. These models enable healthcare providers to identify patients at high risk and implement preventive measures or interventions to mitigate adverse outcomes.

2.2.1.5 Disease Outbreak Prediction

Disease outbreak prediction involves the use of predictive modeling, data analytics and AI techniques to forecast and anticipate the occurrence, spread and severity of infectious diseases within populations [34]. These powerful tools enable proactive measures to be taken, including resource allocation, public health interventions and preparedness strategies.

Predictive modeling forms the foundation of disease outbreak prediction by utilizing historical data, epidemiological factors and environmental variables to build mathematical models [34]. These models estimate the likelihood of disease outbreaks, identify vulnerable populations and forecast the potential trajectory of an outbreak. By incorporating factors such as population density, travel patterns, climate conditions and social behaviors, predictive models can assess the risk of disease transmission and the effectiveness of intervention strategies. Data analytics plays a critical role by extracting insights from diverse data sources, including surveillance data, demographic information, clinical records, social media data and environmental data and advanced analytics techniques identify patterns, anomalies and early warning signals that indicate outbreak emergence or escalation. Leveraging big data analytics, public health agencies and researchers can rapidly detect outbreaks, monitor progression and allocate resources efficiently.

AI, particularly machine learning algorithms, complements predictive modeling and data analytics by uncovering complex relationships and patterns within large datasets. These algorithms process vast amounts of data, learn from historical outbreak patterns and generate reasonably accurate predictions about disease dynamics [36]. They identify risk factors, predict disease spread and inform decision-making for more effective outbreak control. AI-powered algorithms integrate real-time data streams, such as social media and sensor data, enhancing situational awareness and providing timely insights. By combining predictive modeling, data analytics and AI, disease outbreak prediction enhances public health preparedness and response, as early detection and accurate forecasting enable targeted interventions, resource allocation and timely communication to healthcare providers and the public. Outbreak prediction systems are able to optimize surveillance, inform vaccination strategies and identify areas requiring enhanced public health measures.

However, successful implementation of disease outbreak prediction necessitates addressing challenges such as ensuring data quality, integrating diverse data sources, maintaining data privacy and security, addressing algorithm biases and promoting collaboration among stakeholders. Ethical considerations and transparent decision-making processes build trust and support public health interventions. In summary, disease outbreak prediction utilizes predictive modeling, data analytics and AI to forecast and anticipate infectious disease occurrence and spread. Analyzing diverse datasets and uncovering hidden patterns provide valuable insights for public health agencies, enabling timely and targeted interventions to mitigate outbreak impact. Advancements in predictive modeling, data analytics and AI hold great potential for improving our ability to predict, prevent and respond to disease outbreaks, safeguarding public health and saving lives.

2.2.1.6 Surgical Assistants

With advancements in robotics and AI, surgical assistants have witnessed significant transformations, leading to improved surgical outcomes and patient care. Robotics in surgery involves the use of robotic systems that are controlled by surgeons to perform precise and complex tasks. These systems, equipped with robotic arms and surgical instruments, offer enhanced dexterity, precision and stability compared to traditional surgical techniques. Surgical robots can be guided by AI algorithms that enable autonomous or semi-autonomous actions, such as tissue manipulation, suturing and incision closure. This collaboration between robotics and AI empowers surgical assistants by augmenting their capabilities and enabling them to perform tasks with enhanced accuracy and control. For instance, robotic systems, guided by AI algorithms, can suture bowel better than a human hand without human interference, ensuring optimal precision and efficacy in the surgical process [9].

Moreover, AI-powered robotics can enable remote surgical assistance, allowing surgical experts to provide guidance and expertise from a remote location. This is particularly valuable in scenarios where access to specialized surgical expertise is limited or when rapid response is required in emergency situations. Through teleoperated systems, surgical assistants can benefit from real-time feedback and guidance from experienced surgeons, ultimately improving patient outcomes.

The integration of robotics and AI in surgical assistance offers several benefits, it reduces the physical strain on surgical assistants, as robotic systems can perform repetitive and physically demanding tasks, this, in turn, decreases the risk of fatigue and potential errors during surgery. Additionally, robotic-assisted procedures can result in smaller incisions, reduced blood loss, shorter hospital stays and faster recovery times for patients. The precise movements and enhanced visualization provided by robotic systems contribute to improved surgical accuracy, leading to better patient outcomes and reduced post-operative complications [49].

Despite these advantages, the adoption of surgical assistants using robotics and AI also presents challenges. Technical limitations, high costs and the need for specialized training and maintenance of robotic systems are among the key considerations. Ethical and legal aspects, including liability and patient safety, require careful attention as well. Additionally, ensuring proper integration and collaboration between surgical assistants and robotic systems is crucial to maximize the benefits and minimize any potential risks associated with these technologies [39, 35].

2.2.1.7 Rehabilitation Therapy

Rehabilitation therapy has undergone significant advancements with the integration of Augmented Reality (AR) and AI, leading to more effective and engaging treatment approaches [26]. AR, which overlays virtual elements onto the real world, enables the creation of interactive simulations that mimic real-life scenarios for patients to practice and enhance their motor skills, coordination and balance. By utilizing AR devices or applications, individuals can participate in immersive exercises and games that provide visual and auditory feedback, fostering active engagement and motivation throughout the rehabilitation process.

The synergy between AI and AR further enhances rehabilitation therapy by personalizing treatment plans and offering intelligent feedback. AI algorithms analyze patient data, such as movement patterns, muscle strength and progress, to develop tailored rehabilitation programs. These programs adapt based on individual needs and progress, ensuring optimal therapy outcomes. Additionally, AI-powered systems provide real-time guidance and feedback during exercises, assisting patients in improving their techniques and adjusting movements for better results [47].

Moreover, the integration of AI and AR technologies enables remote rehabilitation therapy, breaking down barriers of distance and accessibility. Tele-rehabilitation platforms, supported by AI algorithms, enable healthcare professionals to remotely supervise and guide patients. These platforms track and assess patients' movements, offering personalized feedback and guidance. By allowing individuals to access therapy services from their homes, remote rehabilitation enhances convenience and reduces the need for frequent hospital visits.

The combination of AR and AI in rehabilitation therapy provides numerous advantages, including increased patient engagement, precise monitoring of progress and expanded accessibility to therapy services [14]. Overcoming technical challenges related to device compatibility, system complexity and data privacy is crucial for seamless integration. Ensuring proper training and support for healthcare professionals and addressing ethical considerations related to patient consent and privacy are also essential. With continued advancements, AR, AI and related technologies hold immense potential for transforming rehabilitation therapy, improving patient outcomes and enhancing the overall rehabilitation experience.

2.2.1.8 Medical Imaging

In the realm of medical imaging, AI has achieved notable advancements and has emerged as a powerful tool for diagnostics. Through the utilization of advanced algorithms and machine learning techniques, AI systems have exhibited remarkable capabilities in the analysis and interpretation of various types of medical images, including X-rays, Computerized Tomography Scans (CT scans) and Magnetic Resonance Imaging (MRI). This integration of AI-powered solutions has significantly enhanced the accuracy and efficiency of diagnostic processes [61].

One of the key advantages of AI in medical imaging lies in its ability to assist radiologists in identifying subtle patterns, anomalies and potential disease indicators that might go unnoticed to the human eye. By leveraging its computational power and sophisticated algorithms, AI can perform intricate image analysis, highlighting areas of concern and providing additional insights to radiologists. This collaborative approach between AI and radiologists leads to more accurate and comprehensive assessments, enabling early detection of diseases and timely intervention [25].

AI-powered systems can detect and classify abnormalities in medical images with impressive precision. Through the analysis of vast datasets and training on annotated images, AI algorithms learn to recognize specific patterns associated with different diseases or conditions. This enables them to identify and flag abnormalities, such as tumors, fractures, or lesions, accurately and swiftly, while sometimes outperforming radiologists [15]. Consequently, AI helps radiologists in prioritizing and focusing their attention on areas that require further investigation, potentially leading to faster and more accurate diagnoses.

Moreover, AI in medical imaging has the potential to streamline workflows and enhance efficiency [61]. By automating certain tasks, such as image segmentation, registration and quantification, AI algorithms can significantly reduce the time and effort required for image analysis. This allows radiologists to allocate their expertise and attention to more complex and critical cases, thereby optimizing their productivity and improving overall patient care.

2.2.2 AI Technologies Deployed in Healthcare Domains

In the previous chapter, chapter 2.2.1, different branches of healthcare where AI has had a significant impact were introduced. This new chapter aims to explain in some detail the technologies behind these AI tools.

- **Natural Language Processing**: a subfield of AI that focuses on the interaction between computers and human language. It involves developing algorithms and models that enable computers to understand, interpret and generate human language in a meaningful way. At its core, NLP teaches machines to understand and process human language by breaking it down into its constituent parts, such as words, sentences and paragraphs [40].

NLP encompasses a range of tasks, including language understanding, language generation, sentiment analysis, information extraction and machine translation. One of the key challenges in NLP is dealing with the inherent ambiguity and complexity of natural language [50]. Human language is rich with nuances, context-dependent meanings and linguistic variations, making it a challenging domain for computational analysis. NLP algorithms strive to overcome these challenges by leveraging large-scale language models, semantic analysis, syntactic parsing and context-aware processing.

Tokenization is a fundamental algorithm used in NLP, breaking down text into smaller units called tokens. Part-of-speech tagging assigns grammatical labels to each word in a sentence, aiding in understanding the role and function of each word. Sentiment analysis algorithms analyze words and phrases to classify text as positive, negative, or neutral. Named entity recognition algorithms identify and classify named entities, such as names of people, organizations, or locations. Machine translation algorithms analyze sentence structure and meaning to facilitate accurate translations between languages [38].

Machine learning algorithms, including deep learning, are also utilized in NLP. These algorithms learn patterns and relationships from large datasets to make predictions and generate natural-sounding responses. NLP has a wide range of applications across industries, such as healthcare, finance, customer service and information retrieval. In healthcare, it plays a crucial role in clinical documentation, chatbots (2.1.1), sentiment analysis, machine translation and information extraction from medical records (2.2.1.2) [58].

Overall, NLP algorithms enable computers to understand and process human language, transforming unstructured text into meaningful information. They analyze patterns, structures and meanings in text to extract insights, facilitate communication and provide valuable tools for various applications. NLP is a rapidly evolving field that holds great potential for enhancing human-computer interactions, improving information retrieval and advancing language-related tasks across different domains.

Figure 2.2 shows a common NLP pipeline that involves a series of steps to process and analyze text. First, spelling correction is performed to fix any spelling errors. Then, tokenization breaks the text into individual units called tokens. Sentence splitting identifies and separates sentences within the text. Part-of-speech tagging assigns grammatical labels to each token, indicating their syntactic role. Lemmatization reduces words to their base form. Syntactic analysis helps understand the grammatical structure and word sense disambiguation resolves multiple meanings of words.

In summary, this pipeline helps in analyzing text by correcting spelling, breaking it into tokens and sentences, assigning grammatical labels, reducing words to their base form, understanding the syntactic structure and resolving word meanings.

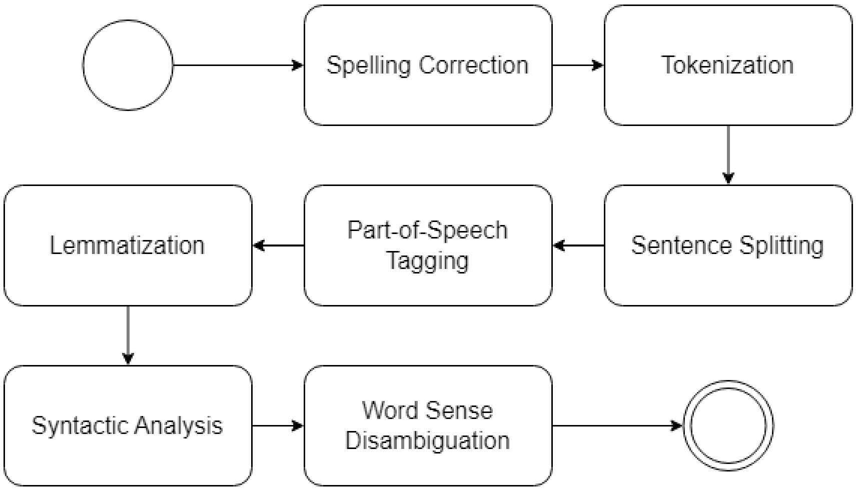

Figure 2.2: A common NLP Pipeline, adapted from [59].

- **Deep Learning**: a subset of AI that focuses on training artificial neural networks to learn and make predictions or decisions without being explicitly programmed. At the core of deep learning are neural networks, which are inspired by the structure and function of the human brain. These networks consist of interconnected layers of artificial neurons, called nodes, that process and transform input data to produce desired outputs.

  What sets deep learning apart from traditional machine learning approaches is its ability to automatically learn hierarchical representations of data. In deep learning, multiple hidden layers are stacked together, allowing the neural network to learn increasingly complex and abstract features from the input data [12]. Each layer extracts higher-level representations by combining and transforming the features learned in the previous layers. This hierarchical representation learning enables deep learning models to capture intricate patterns and relationships in the data, making them highly effective in solving complex tasks.

  To train a deep learning model, a large labeled dataset is used and the neural network adjusts its internal parameters through a process called backpropagation [8]. During backpropagation, the model calculates the difference between its predicted output and the actual output and this error is used to update the weights and biases of the neural network in a way that minimizes the error. This iterative learning process allows the deep learning model to continually improve its performance by fine-tuning its internal representations and making more accurate predictions.

  Deep learning has demonstrated remarkable success in various domains, including image and speech recognition, natural language processing and even medical diagnosis. The ability of deep learning models to automatically learn and extract meaningful features from raw data has led to breakthroughs in tasks that were previously considered challenging or unsolvable. However, deep learning models often require a large amount of labeled data and computational resources for training and their black-box nature can make it difficult to interpret their decision-making process [62].

- **Predictive Modeling**: a technique in AI that aims to make predictions or forecasts based on historical data and patterns. It involves the use of statistical and machine learning algorithms to analyze past observations and uncover relationships between variables, enabling the creation of models that can predict future outcomes [5].

  At the core of predictive modeling is the concept of training a model on existing data to learn patterns and correlations. The model captures the underlying structure and trends within the data, allowing it to generalize and make predictions on new, unseen data. This process involves splitting the available data into a training set and a test set, where the training set is used to build the model and the test set is used to evaluate its performance.

  Predictive modeling algorithms vary depending on the problem at hand, but they typically involve the use of regression, classification, or time series analysis techniques. Regression models aim to predict continuous numerical values, such as sales figures or temperature. Classification models, on the other hand, assign data points to predefined categories or classes, such as predicting whether an email is spam or not. Time series analysis models are specifically designed to predict future values based on patterns observed in sequential data, such as stock prices or weather patterns [13].

  To build a predictive model, features or variables that are relevant to the prediction task are selected and transformed into a format that can be understood by the algorithm. This often involves preprocessing steps such as data cleaning, normalization and feature engineering. The model is then trained using the training data, where it adjusts its internal parameters to minimize the difference between the predicted outcomes and the actual outcomes.

  Once the model is trained and evaluated on the test set, it can be applied to new data to make predictions. The accuracy and reliability of the predictions depend on the quality of the data, the selection of appropriate features, the choice of algorithm and the model's ability to generalize to unseen data. Predictive modeling is widely used in various industries and domains, including finance, healthcare, marketing and manufacturing, to make informed decisions, optimize processes and improve outcomes [27].

- **Robotics**: By integrating artificial intelligence techniques, such as machine learning and computer vision, robots have become more autonomous, adaptable and intelligent. AI-powered robots are equipped with sophisticated algorithms and sensors that enable them to perceive and interact with their environment, make decisions and perform complex tasks with precision and efficiency [7].

  Moreover, AI has revolutionized the field of robotic manipulation and dexterity. Through advanced computer vision techniques, robots can perceive and understand the objects and tasks they are assigned to handle. Machine learning algorithms enable robots to learn from examples and past experiences, improving their ability to manipulate objects with precision and dexterity. This has opened up new possibilities in areas such as manufacturing, healthcare (2.2.1.6) and logistics [63], where robots can perform intricate tasks that previously required human intervention.

  In summary, the integration of AI into robotics has ushered in a new era of intelligent and autonomous machines. AI-powered robots are capable of perceiving their environment, making decisions and performing complex tasks with accuracy and adaptability. As AI continues to advance, one can expect even more sophisticated and capable robots that will further transform industries, enhance productivity and improve our daily lives.

- **Augmented Reality**: The convergence of AI and AR has brought about groundbreaking advancements and transformative experiences. AI plays a vital role in augmenting reality by providing intelligent and context-aware capabilities to AR systems. Through AI-powered algorithms, AR applications can recognize and understand the real-world environment, interpret user interactions and generate responsive and interactive virtual overlays seamlessly integrated with the physical world [52].

  One of the key areas where AI enhances augmented reality is in object recognition and tracking. By leveraging deep learning algorithms, AR systems can accurately identify and track objects in real-time, allowing virtual content to interact and align with the physical environment. This enables users to place virtual objects in their surroundings, creating immersive and interactive experiences [21]. Additionally, AI-powered algorithms can analyze the visual context and semantic understanding of the environment, enabling AR systems to provide relevant and personalized information overlaid on objects or locations.

- **Computer Vision**: AI coupled with computer vision has enabled machines to perceive and understand visual information with human-like capabilities. AI algorithms, particularly deep learning models, have significantly advanced computer vision tasks, allowing machines to analyze and interpret images and videos with remarkable accuracy and efficiency.

One of the key areas where AI enhances computer vision is in object detection and recognition. Deep learning models have demonstrated exceptional performance in detecting and localizing objects in images [48]. By learning from vast amounts of labeled data, these models can identify objects of interest, regardless of variations in scale, orientation and background clutter. AI-powered computer vision systems can accurately label objects, classify them into predefined categories and even generate detailed object segmentation maps [31], enabling a wide range of applications, from autonomous driving to surveillance and robotics. Furthermore, AI contributes to computer vision through image understanding and semantic analysis. By leveraging AI techniques, computer vision systems can extract high-level semantic information from images [31], enabling machines to comprehend scenes, relationships between objects and even human actions. Deep learning models can process sequential data, such as videos, to recognize and predict temporal dynamics, enabling tasks like action recognition, video summarization and anomaly detection. These AI-powered computer vision systems enable machines to understand visual content in a more sophisticated and contextualized manner, opening doors to applications in fields such as healthcare (2.2.1.8), security and content analysis [67]. From object detection and recognition to image understanding and analysis, AI-powered computer vision systems continue to push the boundaries of what machines can perceive and comprehend from visual data. As AI algorithms and models continue to evolve and improve, one can anticipate even more sophisticated computer vision solutions that will have a profound impact on various industries, enhancing efficiency, safety and our overall understanding of the visual world.

2.3 Technologies Used in Information Sharing Platforms

Information sharing empowers individuals to make informed decisions, fosters innovation through collective knowledge, and streamlines processes by eliminating information silos. Effective information sharing strategies typically involve well-defined protocols for access, contribution, and security. The chosen platform depends on the context, with options ranging from internal knowledge bases to collaborative tools.

2.3.1 Zendesk

Zendesk offers several advantages as an information sharing platform. Its comprehensive ticket management system allows businesses to efficiently handle customer inquiries, track tickets and assign tasks to appropriate agents, ensuring prompt and accurate responses. The platform also supports multichannel communication, such as email, chat, phone and social media, consolidating various channels into a single interface for seamless customer interactions [1]. Additionally, Zendesk provides self-service options like knowledge bases and FAQs [6, 69], empowering customers to find answers independently and reducing the workload on support agents. Automation features, robust reporting and analytics, contribute to improved productivity, better insights and data-driven decision-making.

However, there are some disadvantages to consider. The cost of subscription fees and user licenses is significant. Additionally, customization options have limitations, making it challenging for businesses with unique support processes to configure Zendesk to their specific requirements. There may be a steep learning curve involved for administrators and support agents to fully utilize the platform's features.

2.3.2 Discourse

Discourse, as an information sharing platform, offers several advantages for facilitating discussions and knowledge dissemination. One of the main advantages is its user-friendly interface and modern design, which enhances engagement and encourages active participation. Discourse allows users to engage in threaded conversations, providing a structured format for discussions and making it easy to follow conversations. The platform also supports features like mentions, notifications and bookmarks, keeping users informed and enabling them to track relevant topics [66]. This fosters a collaborative environment where users can share their knowledge, ask questions and receive timely responses from community members.

Furthermore, Discourse's categorization and tagging system allows for easy organization and navigation of information. Users can explore different categories or search for specific topics of interest, facilitating efficient access to relevant information. The platform also offers trust levels and moderation tools to maintain a productive and respectful community. By rewarding active and trustworthy contributors while discouraging spam and inappropriate content, Discourse ensures that the information shared is of higher quality and promotes meaningful interactions [18].

However, there are a few potential disadvantages to consider when using Discourse as an information sharing platform. As an open-source solution, it requires technical expertise for installation, configuration and ongoing maintenance. This could pose a challenge for individuals or organizations without the necessary technical resources. Additionally, since the content is largely user-generated, the accuracy and reliability of information can vary. It's essential to have effective moderation in place to ensure the integrity of the shared information and prevent the spread of misinformation or low-quality content.

In summary, Discourse's advantages lie in its user-friendly interface, threaded conversations and categorization features, which facilitate efficient information sharing and collaboration. However, the need for technical expertise and the potential variability of content quality are factors to consider when using Discourse as an information sharing platform.

2.3.3 Bespoke Solutions

When it comes to information sharing pages, opting for bespoke solutions rather than using pre-made tools can offer several advantages as they allow the creation of platforms that are precisely tailored to specific requirements, ensuring a seamless and efficient user experience.

By developing a custom information sharing page, it is possible to design the platform to match specific branding, layout and user interface preferences. This level of customization enables a cohesive and consistent user experience, which can enhance engagement and improve information consumption. The bespoke solution can be built with features and functionalities that are specifically aligned with particular information sharing goals, ensuring that users have access to the most relevant and valuable content.

Another advantage of bespoke solutions for information sharing pages is the flexibility to accommodate unique content types and formats. Unlike pre-made tools that may have limitations on the types of content they can support, a bespoke solution can be designed to handle a wide range of media, such as text, images, videos and documents. This flexibility allows for presentation of information in the most effective and engaging manner, promoting knowledge sharing and collaboration among users.

Furthermore, bespoke solutions provide the opportunity for advanced customization and integration with existing systems or databases, enabling the integration of information sharing platforms with internal databases, external Application Programming Interfaces (APIs), or other software solutions to provide a seamless flow of information and enable real-time updates. This integration capability allows for a more efficient and interconnected ecosystem, enhancing the overall value of the information sharing platform.

While offering extensive customization and integration capabilities, bespoke solutions typically require substantial technical expertise for development and maintenance. This reliance on specialized skills may pose challenges for organizations lacking such resources or facing budget constraints. Additionally, the initial investment in building a bespoke platform can be higher compared to adopting pre-made tools.

2.3.4 Comparing Solutions

In assessing the most suitable platform for information sharing initiatives, it is essential to weigh the advantages and disadvantages of available solutions. Below is a comparative analysis of three prominent approaches: Zendesk, Discourse, and bespoke solutions. Each offers unique features and capabilities tailored to different organizational needs. By examining their strengths and weaknesses, its possible to make informed decisions to optimize information sharing strategies.

Table 2.2: Comparison of Information Sharing Platforms

	Zendesk	**Discourse**	**Bespoke Solutions**
Advantages	- Efficient ticket management - Multichannel support - Self-service options - Automation	- User-friendly interface - Threaded conversations - Categorization and tagging - Trust levels and moderation	- Customizable interface - Tailored to specific requirements - Flexible content presentation
Disadvantages	- Cost of subscription fees and user licenses - Limited customization options	- Requires technical expertise for installation and maintenance - Variable content quality	- Requires technical expertise for development and maintenance - Potentially higher initial cost

2.4 Data Mining - Association Rules

Association Rule Mining (ARM) is one of the most pivotal and extensively studied data mining method [41]. Its significance lies in its ability to unveil intricate relationships, frequent patterns and associations within vast datasets, thereby facilitating the extraction of valuable insights and enabling informed decision-making across diverse domains [41].

At its core, ARM delves into transactional databases or repositories to discern patterns of co-occurrence or correlation among different items. These patterns are encapsulated in association rules, typically depicted in the form of "if-then" statements, indicating the presence of certain items implies the likelihood of other items occurring as well. This process is instrumental in revealing not only what items frequently co-occur but also the strength and significance of these associations.

ARM was initially employed for Market Basket Analysis, aiming to unravel the intricate web of relationships among items purchased by customers. This analytical technique explores transactional data to uncover patterns and correlations, giving a glimpse into how different products are linked in a shopping basket. By carefully analyzing these transactions, ARM helps to better understand consumer behavior, revealing which items tend to go hand in hand and uncovering hidden links behind purchase decisions [3].

The applications of ARM extend far beyond specific domains, permeating various sectors and industries. Its utility lies in its versatility and adaptability to different contexts, enabling the leverage of its capabilities for a wide array of purposes such as: [41, 55]

- **Market Basket Analysis**: uncovers associations between products frequently purchased together, aiding in targeted marketing strategies and store layout optimizations.

- **Telecommunication networks**: helps identify patterns in user behavior, such as call routing preferences or usage trends, assisting in network optimization and resource allocation.

- **Risk Management**: identifies correlations between various risk factors, enabling the development of predictive models to assess and mitigate potential risks effectively.

- **Inventory control**: analyzes purchase patterns to predict demand for different products, optimizing stock levels and reducing excess inventory costs.

- **Medical Diagnosis**: reveals associations between symptoms, diseases and patient characteristics. By analyzing patient data, ARM can aid in identifying common patterns of symptoms associated with specific diseases, assisting healthcare professionals in making accurate diagnoses and treatment decisions.

- **Protein Sequences**: recognizes patterns and associations among protein sequences. By analyzing sequence data, ARM can help researchers discover relationships between protein structures, functions and evolutionary histories, facilitating insights into biological processes and drug discovery efforts.
- **Census Data**: determines associations and trends among demographic variables such as age, income, education and geographic location. By identifying patterns in census data, ARM can assist policymakers and planners in understanding population dynamics, allocating resources efficiently and making informed decisions regarding social and economic policies.
- **Customer Relationship Management**: finds patterns and associations among customer interactions, preferences and behaviors. By analyzing customer data, ARM can assist businesses in segmenting customers, predicting customer needs and preferences and developing targeted marketing campaigns to enhance customer satisfaction and loyalty.

2.4.1 Association Rules Definitions

In order to understand how ARM functions, its necessary to understand the basic definitions. (adapted from [3] and [64])

An **Itemset**, $I = \{i_1, i_2, \ldots, i_n\}$, is a collection of one or more items. In a transactional database context, items could represent products, services, or any other entities that can be associated with transactions.

A **Transaction** T is a set of items associated together (e.g. bought, performed together), such that $T \subseteq I$. Transactions are subsets of the entire itemset.

A **Transaction Database** D is a set of all transactions within a dataset. It serves as the source of data for mining association rules.

An **Association Rule**, $X \Rightarrow Y$, where $X, Y \subseteq I$ is an implication that suggests a relationship between two sets of items. It is in the form of "X implies Y," where X and Y are subsets of the item set I. The rule indicates that if X occurs in a transaction, then Y is likely to occur as well. Note that the sets can be composed of only one item, $X \Rightarrow i_j$, where $i_j \in I$ and $i_j \notin X$, with i_j being a single item in I that is not present in X.

The **Support**, S of an itemset I is the proportion of transactions in the transaction database D that contain X. It indicates the frequency with which the itemset appears in the transactions. Mathematically, it is calculated as the number of transactions containing X divided by the total number of transactions. The support of the rule $X \Rightarrow Y$ in the transaction database D is the support of the combined itemset $(X \cup Y)$ in D. It indicates how often the rule is true for the given dataset.

The **Confidence**, C measures the reliability of the association rule. It is the ratio of the number of transactions in D that contain both X and Y, $(X \cup Y)$ to the number of transactions that contain X. In other words, it represents the probability of occurrence of Y given the occurrence of X.

The structure of association rules follows the pattern: 'If <condition> then <conclusion>' with associated support and confidence metrics. The support represents the percentage of examples that satisfy both the condition and the conclusion, denoted as $p(condition, conclusion)$. On the other hand, confidence evaluates the level of confidence about the conclusion given a condition, expressed as $p(conclusion|condition)$, where:

$$p(conclusion|condition) = \frac{p(conclusion \cap condition)}{p(condition)}$$

These metrics play a crucial role in assessing the significance and reliability of association rules.

2.4.2 Association Rules Simple Example

To understand these concepts, a simple hypothetical example will be presented. A scenario where we have a list of items purchased by five clients at a supermarket will be considered. Table 2.3 presents a brief overview of the purchases made by these users:

Table 2.3: List of items purchased

List of Items
Milk, Yoghurt, Bread
Milk, Eggs, Bread
Milk, Yoghurt
Ham, Yoghurt
Milk, Bread, Ham

From table 2.3, a new table of co-occurrences can be devised to facilitate the determination of associations (table 2.4). This table serves to simplify the analysis by showcasing how frequently pairs of items are purchased together. Each cell in the co-occurrence table represents the number of times two items appear together in a single shopping transaction.

Table 2.4: Table of co-occurrence of items purchased in table 2.3

	Milk	**Ham**	**Eggs**	**Yoghurt**	**Bread**
Milk	4	1	1	2	3
Ham	1	2	0	1	1
Eggs	1	0	1	0	1
Yoghurt	2	1	0	3	1
Bread	3	1	1	1	3

By examining this co-occurrence table, we gain insights into which items tend to be bought together more frequently. This information forms the basis for identifying associations or patterns in shopping behavior. In this example, it is evident that eggs and yogurt are never purchased together, while milk and bread are frequently bought together.

With this information, calculating the support and confidence of association rules becomes easier. For instance, if we want to analyze whether customers who buy milk also purchase bread, we can examine the association rule: "If Milk then Bread." The support for this rule is 3 out of 5 transactions, or 60% (indicating that milk and bread were bought together in 3 out of the 5 transactions), while the confidence is 3 out of 4, or 75% (showing that out of all transactions involving milk, 3 out of 4 also included bread).

Similarly, exploring the association rule "If bread then milk" reveals that the support is also 3 out of 5 transactions, or 60%, indicating that bread and milk were bought together in 3 out of the 5 transactions. The confidence for this rule is 3 out of 3, or 100%, meaning that whenever bread was purchased, milk was also purchased in all 3 instances where bread was bought. This bidirectional analysis provides a comprehensive understanding of the relationship between milk and bread in customer transactions in this particular itemset.

In essence, the association rule problem entails the exploration of all association rules in a transaction dataset, denoted as $X \Rightarrow Y$, where X and Y represent sets of items. This pursuit mandates the specification of minimum support and confidence thresholds. Initially, the focus lies on discovering frequent itemsets that meet or exceed the support threshold. Subsequently, association rules adhering to the confidence criteria are derived from these itemsets.

2.4.3 Association Rules Algorithms

The field of data mining encompasses a plethora of algorithms designed to extract patterns and insights from large datasets. These algorithms employ various strategies and data structures to achieve their objectives. Despite their diversity, they all aim to identify association rules within transactional data.

Given a transaction dataset T and specified minimum support and confidence thresholds, the resulting set of association rules is consistent across different mining algorithms. This means that regardless of the specific algorithm employed, the set of association rules derived from the dataset remains the same.

In other words, the association rules present within the dataset are uniquely determined by its contents and the specified thresholds. Therefore, any mining algorithm applied to the dataset should yield the same set of association rules, even though their computational efficiencies and memory requirements may vary.

While different algorithms may employ distinct methodologies and optimizations, ultimately they converge on identifying the same underlying patterns and relationships within the data. This ensures that the extracted association rules are robust and consistent across various mining techniques, enabling analysts to confidently derive actionable insights from the data. However, it is important to note that certain algorithms may have unique features or capabilities that distinguish them from others. For example, some algorithms may support the generation of more complex association rules involving multiple items in the consequent, while others may focus on specific types of data or patterns. These algorithm-specific features can influence the types of insights that can be extracted from the data and may lead to differences in the resulting association rules.

2.4.3.1 AIS Algorithm

The AIS (Agrawal, Imielinski and Swami) algorithm [3], introduced as the first method for mining association rules, aimed to enhance database quality while providing essential functionality for processing decision support queries, pioneering the exploration of association rules within transactional datasets.

One notable characteristic of the AIS algorithm is its generation of association rules with only one item as the consequent. For instance, it can generate rules such as $X \cap Y \rightarrow Z$, where X and Y are sets of items leading to item Z as the consequent. However, it does not generate rules like $X \rightarrow Y \cap Z$, where X leads to a combination of items Y and Z in the consequent.

This restriction simplifies the rule generation process, focusing on associations between itemsets and individual items rather than complex combinations of items in the consequent. While this approach may limit the breadth of association rules generated, it can streamline the analysis and interpretation of results, particularly in scenarios where clarity and simplicity are paramount.

The process of uncovering frequent itemsets in the AIS algorithm involves iteratively scanning the database multiple times. The algorithmic process is outlined in Tables (I-VI), shown in Figure 2.3. During the initial pass, individual item support counts are accumulated, as shown in Table (II). Assuming a minimum support threshold of 30%, large single-item sets are identified in Table (III). Items with support counts below the threshold (e.g., i_4) are removed as they are considered infrequent.

These frequent individual items serve as the foundation for generating candidate 2-itemsets. This process involves extending them with other items found in the same transaction. To prevent creating duplicate itemsets, items are ordered beforehand. Candidate generation then involves joining large items from the previous pass with any remaining item in the transaction, ensuring it appears after the last item in the existing frequent itemsets. For example, considering the transactions, based on the defined order, candidate 2-itemsets are formed by extending i_1 with i_2, i_2 with i_3 and i_6, following the sequence, resulting in the groupings shown in Table (IV).

Subsequent database scans involve calculating the support count of these candidate 2-itemsets and comparing them to the minimum support threshold. This process continues iteratively, generating candidate (k+1)-itemsets by extending frequent k-itemsets with additional items from the same transaction. The iteration concludes when no further frequent itemsets are discovered. In this example, the final set of frequent itemsets includes no large 3-itemset.

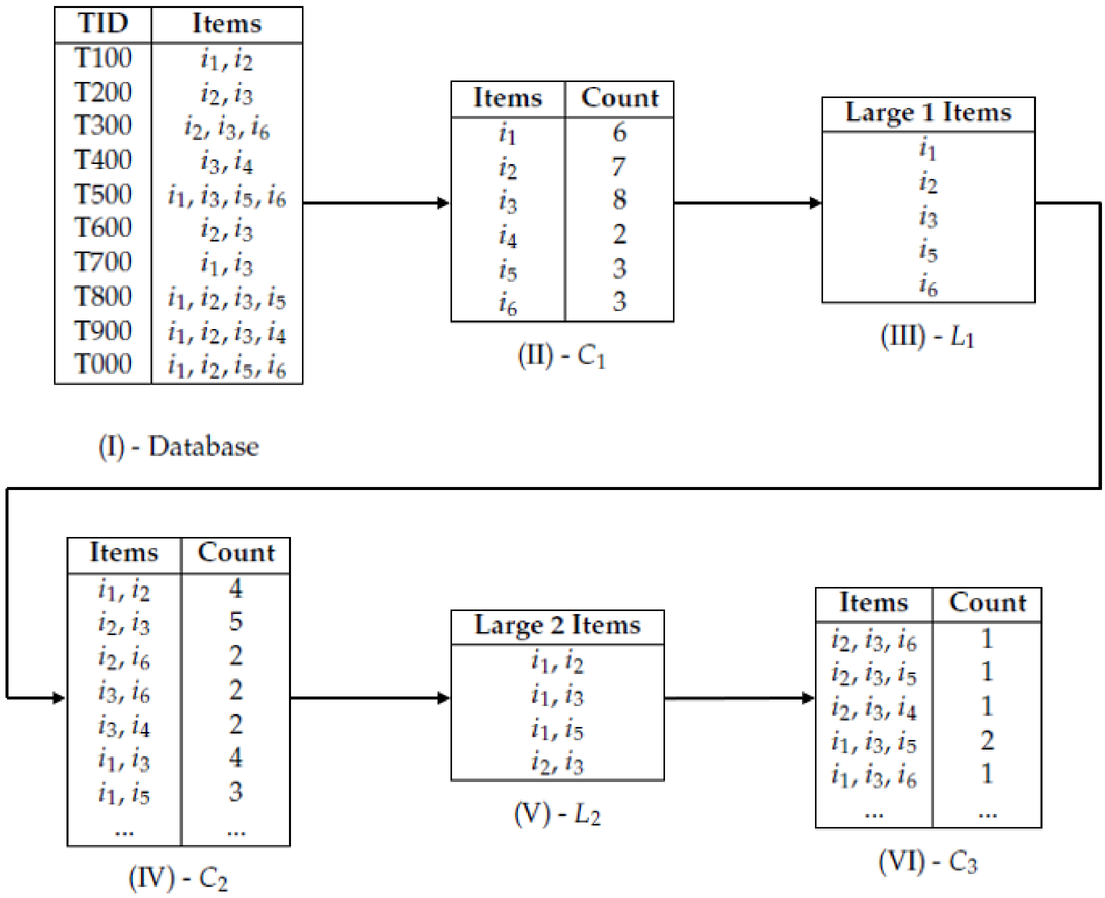

Figure 2.3: AIS Algorithm Process. Adapted from [76]

A key drawback of the AIS algorithm lies in the generation of a large number of candidate itemsets that ultimately turn out to be infrequent. This leads to increased memory requirements and wasted computational effort. Furthermore, the algorithm necessitates multiple scans of the entire database, further adding to its limitations [42].

2.4.3.2 Apriori Algorithm

The Apriori algorithm stands out as a significant advancement in the realm of association rule mining, with its foundational proposal by Agrawal in [4]. In contrast to the straightforward approach of the AIS algorithm, Apriori introduces notable efficiency enhancements by optimizing the candidate generation process and implementing a novel pruning technique.

At the core of the Apriori algorithm lies the Apriori property, a fundamental principle that greatly influences its efficiency and effectiveness. This property states that if an itemset is frequent (i.e., its support exceeds the minimum threshold), then all of its subsets must also be frequent. In other words, if a set of items occurs frequently in the dataset, then every subset of that set must also occur frequently.

Apriori's approach to identifying all large itemsets from the database involves two distinct processes. Initially, candidate itemsets are generated, followed by a database scan to ascertain the actual support count of each corresponding itemset. During the first database scan, the support count of each item is calculated and large 1-itemsets are generated by pruning those itemsets whose support falls below the predefined threshold. Figure 2.4 displays the algorithmic process outlined in Tables (I-VI).

Subsequent passes over the database involve generating candidate k-itemsets by joining frequent (k-1)-itemsets. However, to abide by the Apriori property, candidate k-itemsets are pruned by checking their sub (k-1)-itemsets. If any sub (k-1)-itemset is not in the list of frequent (k-1)-itemsets, the candidate k-itemset is pruned out. This iterative process continues until all frequent itemsets are identified. To illustrate the process, let's consider how candidate 1-itemsets are generated. We begin by joining frequent 1-itemsets. This will result in itemsets such as {(i_1, i_2), (i_2, i_3), (i_2, i_6), (i_3, i_6), (i_3, i_4), ...}.

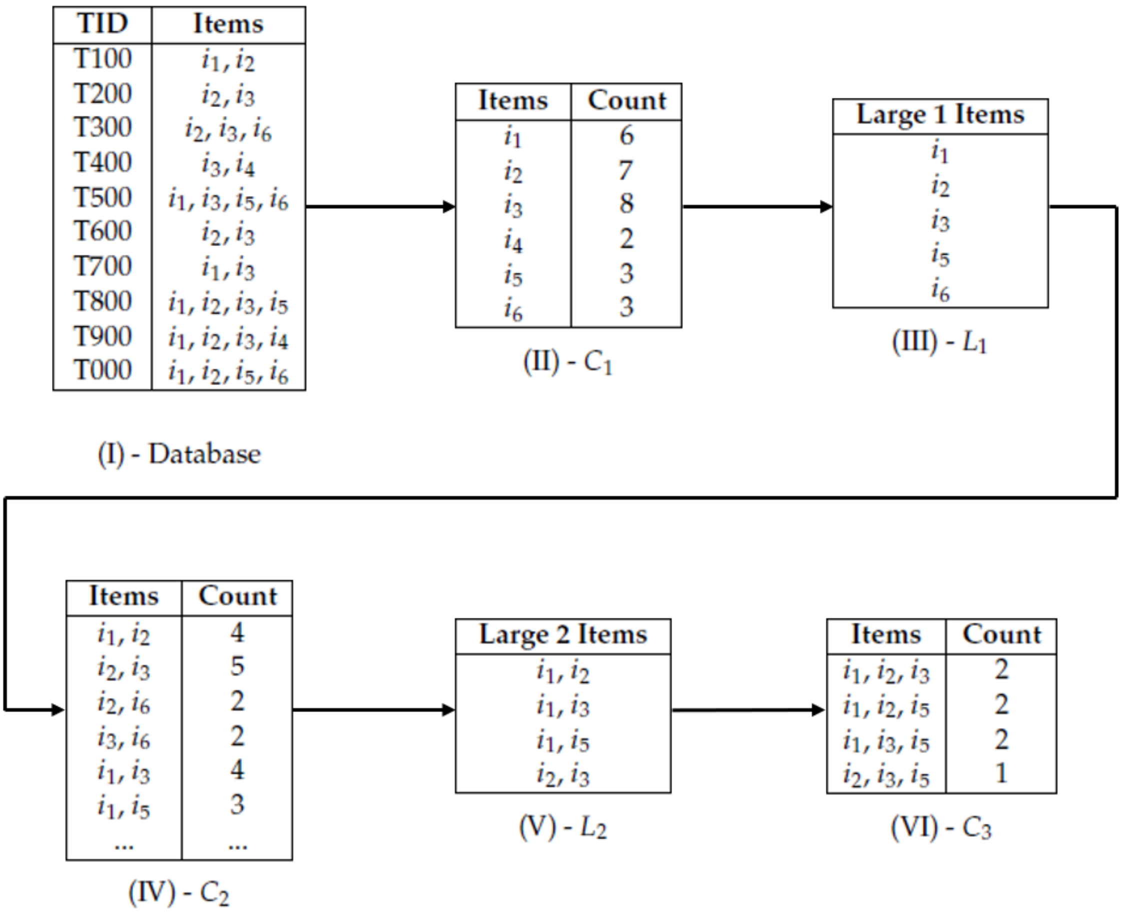

Figure 2.4: Apriori Algorithm Process. Adapted from [76]

Next, we perform a subset check on these candidate itemsets. This ensures that any subset of a frequent itemset must also be frequent. Since i_4 is not a frequent 1-itemset, the candidate itemsets containing it, like {(i_3, i_4), (i_4, i_5), (i_4, i_6), ...} are removed from the list.

This process of joining, checking subsets and removing non-qualifying candidates is repeated iteratively until no more frequent itemsets can be found. This means either no new candidate itemsets can be generated, or all generated candidates are infrequent.

Apriori's strategy mitigates wasteful efforts by avoiding the counting of candidate itemsets known to be infrequent. By generating candidates level-wise from frequent itemsets and pruning them based on the Apriori property, the algorithm significantly reduces computation, I/O costs and memory requirements. Comparing the candidate counts in Tables VI of Figures 2.3 and 2.4 illustrates the dramatic reduction achieved by the Apriori algorithm.

Despite its efficiency improvements, Apriori still grapples with drawbacks such as multiple scans of the database and a complex candidate generation process. However, it employs a level-wise search approach and limits the number of passes over the data to the size of the largest itemset (K), aiming to mitigate its resource consumption.

2.4.3.3 FP-Tree Algorithm

In efforts to overcome the limitations of Apriori series algorithms, [Han et al.] have devised association rule mining techniques employing tree structures. One such breakthrough is the FP-Tree Algorithm [30], a pivotal advancement in association rule mining that effectively addresses Apriori's bottlenecks. Unlike its predecessor, FP-Tree generates frequent itemsets through just two database passes, completely eliminating the need for candidate generation.

By circumventing the candidate generation process and minimizing database passes, the FP-Tree algorithm achieves a speed boost of an order of magnitude compared to Apriori [76, 30]. The frequent pattern generation involves two main steps: constructing the FP-Tree and deriving frequent patterns from it. The FP-Tree construction process unfolds as follows: **(1)** The initial scan of the database collects the support count of each item, generating frequent 1-itemsets, akin to the Apriori algorithm. These frequent itemsets are then sorted in descending order of their supports and an ordered head table is created, as demonstrated in Table III of Figure 2.5. **(2)** A root node labeled "Root" initializes the FP-Tree, denoted as T, depicted in Figure 2.6.

TID	Items
T100	i_1, i_2
T200	i_2, i_3
T300	i_2, i_3, i_6
T400	i_3, i_4
T500	i_1, i_3, i_5, i_6
T600	i_2, i_3
T700	i_1, i_3
T800	i_1, i_2, i_3, i_5
T900	i_1, i_2, i_3, i_4
T000	i_1, i_2, i_5, i_6

(I) - Database

Items	Count
i_1	6
i_2	7
i_3	8
i_4	2
i_5	3
i_6	3

(II) - C_1

Items	Count
i_3	8
i_2	7
i_1	6
i_5	3
i_6	3

(III) - C_1

TID	Ordered Large Items
T100	i_2, i_1
T200	i_3, i_2
T300	i_3, i_2, i_6
T400	i_3
T500	i_3, i_1, i_5, i_6
T600	i_3, i_2
T700	i_3, i_1
T800	i_3, i_2, i_1, i_5
T900	i_3, i_2, i_1
T000	i_2, i_1, i_5, i_6

(IV) - Transactions of Large Items, ordered by Support

Figure 2.5: FP-Tree Data Transformation Process. Adapted from [30].

The database undergoes another scan to construct the FP-Tree utilizing the previously created head table. During this process, the order of frequent items within each transaction is rearranged based on the head table. For instance, consider the transaction T800 (i_1, i_2, i_3, i_5), which is reordered as (i_3, i_2, i_1, i_5) since i_3 has a higher frequency than i_2 and i_2 has a higher frequency than i_1 in the database. Let the items in each transaction be denoted as $[p \mid P]$, where p represents the first frequent item and P represents the remaining items list. Subsequently, the function Insert$[p \mid P]; T$ is invoked. This function operates as follows:

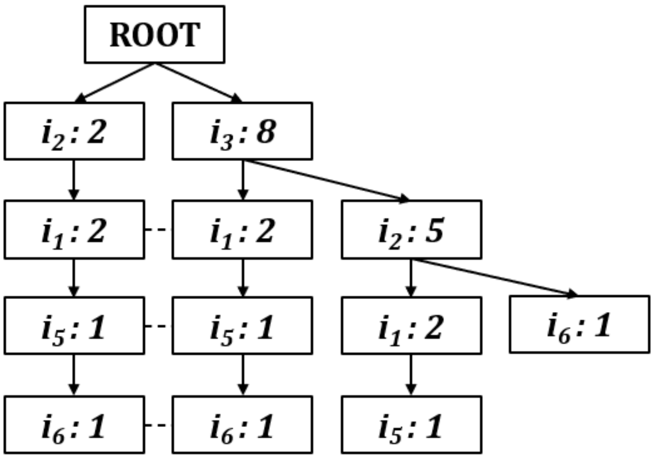

Figure 2.6: FP-Tree with data from Figure 2.5. Adapted from [30].

If T has a child node N such that N.item-name equals p.item-name, then the count of N is incremented by 1. Otherwise, a new node N is instantiated with N.item-name set to p.item-name and a support count of 1. Its parent link is connected to T and its node link is established to the node with the same item-name through a sub-link. This Insert$P;T$ function is recursively invoked until P becomes empty. To illustrate the insertion process and the construction of the FP-Tree mentioned earlier, let's consider the insertion of transaction T300 into the FP-Tree.

After reordering, the transaction becomes (i_3, i_2, i_6), with i_3 designated as p and (i_2, i_6) as P. Subsequently, the insert function is invoked. Initially, we search to determine whether the node i_3 exists in the tree. It turns out that i_3 was already created in transaction T200 (i_3, i_2). Following the protocol, the node named i_3 is incremented with a support count of 1. Since T is the Root in this context, the node i_3 is linked to the Root and the insert function is recursively called again. This time, p is i_2, P is i_6 and T is i_3. The resulting FP-Tree for the database is depicted in Figure 2.6.

Utilizing the head table (Table III of Figure 2.5) and the FP-Tree (2.6) facilitates the generation of frequent patterns. For instance, let's consider the process of obtaining all frequent itemsets related to i_6. Initially, following the head table, we identify the pattern base of this node, comprising all paths that terminate with this node.

For i_6, its pattern base comprises: $(i_3, i_2)(1)$, $(i_3, i_1, i_5)(1)$ and $(i_2, i_1, i_5)(1)$, with the number in parentheses denoting the support of each pattern. Subsequently, the counts of all items in the pattern base are aggregated, resulting in $i_3(2)$, $i_2(2)$ and $(i_5)(1)$ in this case. By comparing the support count with the minimal support threshold of 30%, there are no conditional FP-Tree of i_6 generated. If we assume a minimal support threshold of 20%, the conditional FP-Tree of i_6 is generated $(i_3, i_2)(2)$. Consequently, we derive the frequent itemset/pattern (i_3, i_2, i_6). The mining results align with those of the Apriori algorithm.

The efficiency of the FP-Tree algorithm is attributed to three main factors. Firstly, the FP-Tree serves as a condensed representation of the original database, as it exclusively incorporates frequent items for tree construction, discarding irrelevant data. Additionally, through item ordering based on their support levels, overlapping sections are streamlined to appear just once with distinct support counts.

Secondly, this algorithm necessitates only two scans of the database. Utilizing the FP-growth procedure, frequent patterns are generated by constructing conditional FP-Trees containing patterns with designated suffix patterns, facilitating efficient pattern generation as demonstrated in the aforementioned example. Moreover, this approach significantly reduces computational expenses.

Thirdly, FP-Tree employs a divide-and-conquer approach, differing from the Apriori-like bottom-up generation of combinations of frequent itemsets. This leads to a significant reduction in the size of the conditional pattern base generated at the subsequent search level, as well as the size of its associated conditional FP-tree.

However, like every algorithm, FP-Tree has its limitations, particularly its challenges in adapting to interactive mining systems. In interactive mining, users have the flexibility to adjust the support threshold based on established rules. However, with FP-Tree, modifying the support threshold may necessitate repeating the entire mining process. Additionally, FP-Tree is not well-suited for incremental mining. As databases evolve over time with new data inserts, employing the FP-Tree algorithm may require repeating the entire process due to these changes [76].

2.4.3.4 Comparing the Algorithms

When comparing FP-Tree, Apriori and AIS algorithms, each exhibits distinct strengths and limitations in the context of association rule mining.

AIS algorithm demonstrates effectiveness in generating association rules and in its ease of implementation but encounters challenges related to the sheer volume of candidate itemsets produced. This abundance of candidates, many of which turn out to be infrequent, contributes to heightened memory requirements and computational overhead. Additionally, the necessity for multiple database scans and its restriction of only being able to generate association rules with one item in the consequent further compounds these limitations, potentially hindering the algorithm's practical utility in large-scale mining tasks.

On the other hand, Apriori showcases efficiency by employing a level-wise search approach and constraining the number of database scans to the size of the largest itemset.

Despite these advantages, it still contends with issues like multiple passes over the database and a complex candidate generation process. While its resource consumption is mitigated to some extent, the need for repeated scans and extensive processing can impede its scalability and efficiency. Lastly, the FP-Tree algorithm offers notable efficiency improvements by constructing a compact representation of the database and streamlining the generation of frequent itemsets. However, its adaptability to interactive mining systems is challenged, as adjusting the support threshold may demand a complete re-execution of the mining process. Furthermore, FP-Tree lacks suitability for incremental mining, meaning that database updates necessitate repeating the entire mining procedure, which can be cumbersome and resource-intensive. In summary, each algorithm offers distinct advantages and drawbacks, making it crucial to carefully consider their suitability and trade-offs based on specific mining requirements and dataset characteristics.

2.5 Persuasive Technologies

Persuasive technologies are transforming the way we interact with the world, subtly influencing our attitudes and behaviors. These technologies, designed to change our minds subtly, are seamlessly integrated into our daily lives, from the social media platforms we use to the fitness trackers we wear [23].

Persuasive technologies employ various tactics, drawing from psychology, marketing and computer science, to influence users towards specific behaviors. These tactics can include:

- **Social influence:** leveraging social proof, commitment and authority figures to influence our decisions. By showcasing popular choices, fostering a sense of commitment and presenting authoritative endorsements, these tactics shape our behavior and attitudes.
- **Gamification:** incorporating game mechanics like points, badges and leaderboards to motivate participation. By turning tasks into engaging challenges and rewarding progress with virtual achievements, gamification enhances motivation and encourages desired behaviors.
- **Personalization:** Adaptive learning platforms tailors the learning experience to individual preferences and needs. By analyzing user data and adapting content and pacing accordingly, these platforms optimize engagement and facilitate effective learning outcomes.
- **Scarcity and urgency:** creating a sense of limited time or availability to encourage immediate action. By highlighting limited quantities or time-bound offers, these tactics tap into users' fear of missing out and drive impulse decisions and behaviors.

They also leverage principles [23] such as:

- **Principle of Attractiveness:** This principle emphasizes the importance of visual appeal in computing technology. Research has shown that aesthetically pleasing designs can capture users' attention and evoke positive emotions, making them more receptive to persuasion. By investing in visually appealing interfaces, developers can enhance the persuasive potential of their technologies.
- **Principle of Similarity:** People tend to be more persuaded by technologies that reflect aspects of themselves or their identities. Whether it's through personalized recommendations, relatable content, or user-specific customization options, aligning computing technology products with users' preferences and characteristics can increase their persuasive impact.

- **Principle of Praise:** Offering praise can be a powerful tool in persuasion. Positive reinforcement, conveyed through words, images, symbols, or sounds, can boost users' confidence and self-esteem, making them more receptive to the messages or suggestions presented by computing technologies. Incorporating praise into user interactions can foster a sense of validation and encouragement.

- **Principle of Reciprocity:** Humans have a tendency to reciprocate favors or gestures, even from non-human entities like computing technology. When users perceive that a technology has provided them with value or assistance, they may feel compelled to reciprocate by complying with its requests or recommendations. By offering useful features or helpful support, computing technologies can trigger this sense of obligation and increase their persuasive influence.

- **Principle of Authority:** People are more likely to comply with requests or directives from authoritative sources. In the context of computing technology, assuming roles of authority, such as presenting expert opinions, certifications, or endorsements can enhance credibility and persuasion. Users may be more inclined to trust and follow the guidance of technologies that exude authority or expertise in their respective domains.

Persuasive technologies have a wide range of applications, impacting various aspects of peoples lives. Some prominent examples [28] include:

- **Marketing:** Online platforms employ targeted advertising, personalized recommendations and scarcity tactics to influence buying decisions. By leveraging user data, advertisers tailor ads to specific audiences, maximizing their impact. Scarcity tactics, such as limited-time offers, create a sense of urgency, compelling users to make immediate purchases.

- **Public health:** Gamified apps and reminder systems encourage healthy habits like exercise and medication adherence. Gamification elements, such as challenges and rewards, make behavior change engaging, while reminder systems improve medication compliance through timely notifications.

- **Education:** Adaptive learning platforms personalize the learning experience and provide motivating feedback. By analyzing student performance and preferences, these platforms dynamically adjust content to match individual needs, enhancing engagement and retention. Real-time feedback and progress tracking reinforce positive learning behaviors, fostering a sense of accomplishment and intrinsic motivation.

- **Social media:** Platforms leverage social influence, emotional triggers and endless scrolling to keep users engaged. Social proof and peer influence shape user behavior, while emotional triggers elicit strong reactions and engagement. Endless scrolling mechanisms prolong user sessions, driving higher levels of engagement and ad revenue.

While persuasive technologies can be powerful tools for positive behavior change, they also raise ethical considerations. Concerns include issues of privacy, autonomy and manipulation. Designers and developers have a responsibility to ensure that these technologies are used ethically and transparently, with due consideration given to users' rights and well-being. Persuasive technologies have the potential to shape attitudes and behaviors in various domains. By understanding human psychology and employing effective design strategies, these technologies can be harnessed to promote positive outcomes and improve user experiences. However, it's essential to approach their development and implementation with caution, considering the ethical implications and potential consequences [28].

2.6 Cloud Computing

Cloud computing refers to the delivery of computing services over the internet, providing access to resources such as storage, processing power and applications on a pay-as-you-go basis. This model has revolutionized the way businesses and individuals access and use technology, offering flexibility, scalability and cost-efficiency. Cloud computing represents a significant advancement in information technology, emerging as a predominant model for provisioning IT resources. It facilitates convenient, on-demand access to a shared pool of managed and scalable resources, including servers, storage and applications. In our daily routines, we heavily rely on cloud services for tasks such as data storage, document creation, business management and online gaming [65].

Furthermore, cloud computing underpins crucial digital trends like mobile computing, the Internet of Things, big data and artificial intelligence. This infrastructure accelerates industry evolution, disrupts traditional business models and propels digital transformation forward. Nonetheless, alongside its myriad benefits and opportunities, cloud computing poses challenges and concerns, particularly regarding the security and protection of customer data.

The widespread adoption and transformative impact of this computing paradigm can be owed to its array of defining features [45]:

- **On-Demand Self-Service:** In the realm of cloud computing, users wield unprecedented autonomy, empowered to provision computing resources tailored precisely to their requirements. This liberation from dependency on service providers enables individuals and organizations alike to swiftly respond to evolving needs with agility and efficiency, unencumbered by bureaucratic hurdles or delays.

- **Broad Network Access:** The essence of cloud services lies in their omnipresence across the digital landscape, seamlessly accessible via the internet from an array of devices including laptops, smartphones and tablets. This ubiquitous accessibility transcends geographical barriers, fostering collaboration, enhancing productivity and fueling innovation on a global scale.

- **Resource Pooling:** At the heart of cloud infrastructure lies a sophisticated multi-tenant model, enabling providers to judiciously allocate and reallocate resources among diverse users based on fluctuating demand. This pooling mechanism not only optimizes resource utilization but also engenders substantial cost savings and heightened operational efficiency, benefiting both providers and consumers alike.

- **Rapid Elasticity:** A hallmark feature of cloud resources is their innate elasticity, facilitating instantaneous scaling in response to dynamic fluctuations in demand. This inherent flexibility empowers businesses to seamlessly accommodate sudden surges in traffic or workload without the need for costly investments in additional infrastructure, thereby ensuring uninterrupted service delivery and optimal performance.

- **Measured Service:** Central to the ethos of cloud computing is meticulous metering of resource consumption, affording customers unparalleled transparency and control over their expenditure. This pay-as-you-go model not only fosters cost-effectiveness but also enables businesses to accurately align their expenses with actual resource utilization, thereby maximizing efficiency and optimizing budget allocation.

Cloud services can be generally divided into three categories, depending on the service they provide. These categories represent distinct layers of abstraction in the cloud computing ecosystem, each catering to specific user needs and preferences [45]:

- **Infrastructure as a Service (IaaS):** Representing the foundational layer of cloud computing, IaaS bestows users with virtualized computing resources accessible over the internet. These resources encompass a comprehensive suite of offerings, including virtual machines, storage and networking capabilities. Notably, users retain full autonomy and control over their chosen operating systems, applications and middleware, thus facilitating unparalleled customization to accommodate specific requirements and preferences. With IaaS, organizations can effortlessly scale their infrastructure in tandem with evolving demands, all while minimizing operational overheads and maximizing resource utilization efficiency.

- **Platform as a Service (PaaS):** Positioned as an intermediary layer between infrastructure and software, PaaS offers a holistic platform that streamlines the development, deployment and management of applications. By abstracting away the intricacies of underlying infrastructure provisioning and maintenance, PaaS liberates developers from the complexities traditionally associated with software development. This enables them to focus their efforts solely on coding, testing and refining applications, thereby accelerating the application development lifecycle and fostering a culture of innovation and agility. With PaaS, organizations can rapidly bring new products and services to market, capitalize on emerging opportunities and remain at the forefront of technological innovation.

- **Software as a Service (SaaS):** Represents the pinnacle of convenience and accessibility, delivering software applications directly to end-users over the internet on a subscription basis. This revolutionary deployment model eliminates the need for users to install, maintain, or update applications locally, thereby alleviating the associated administrative burdens and costs. Instead, users gain instant access to a vast range of applications via web browsers or APIs, seamlessly integrating into their existing workflows and processes. Moreover, with the provider assuming responsibility for maintenance and support, users can rest assured knowing that their software applications are always up-to-date, secure and optimized for peak performance. SaaS empowers organizations to leverage cutting-edge software solutions without the constraints of traditional licensing models, enabling them to enhance productivity, streamline operations and drive business growth with unparalleled efficiency and ease.

Given the remarkable success and widespread adoption of cloud technology, there has been a significant proliferation of cloud providers and serverless platforms. As such, newer and more specialized terms like Function as a Service (FaaS) and Backend as a Service (BaaS) have come to fruition and have become increasingly prevalent.

- **BaaS:** Rrefers to services that serve as substitutes for traditional backends, handling tasks such as data storage, querying, authentication processes and notifications. Google Firebase[2] is a prominent example of such a service, simplifying tasks like user management from an administrative standpoint.
- **FaaS:** Pertains to services that execute code in response to events, often serving as backend services that respond to incoming Hypertext Transfer Protocol (HTTP) requests from users. Google Cloud Functions[3] is Google's offer for FaaS.

Most cloud providers offer comprehensive solutions that encompass all aspects of cloud computing, catering to the diverse needs and requirements of their users. By offering a wide range of services under one umbrella, cloud providers aim to streamline the deployment and management of cloud-based resources for their customers. This approach not only simplifies the procurement process but also promotes interoperability and consistency across different components of the cloud ecosystem.

For example, major cloud providers like Google Cloud Platform, Amazon Web Services and Microsoft Azure offer extensive catalogs of services that span the distinct layers of cloud computing, IaaS, PaaS, SaaS, etc.

Overall, by offering comprehensive solutions that cover all aspects of cloud computing, cloud providers empower organizations to build, deploy and scale applications with ease, not only simplifies the management of IT infrastructure but also accelerates time-to-market for new products and services.

[2]https://firebase.google.com/
[3]https://cloud.google.com/functions/

3 Development

3.1 Overview

After analyzing and studying possible solutions and different approaches in the State of the Art chapter, Chapter 2, it became clear that a bespoke knowledge base platform needed to be built from the ground up, tailored to the user's specific needs, as the functionalities provided by the platforms examined are either limited or unnecessary.

By developing a platform from the ground up, one gains the advantage of customization and fine-tuning every aspect to align perfectly with specific objectives. One can carefully design and implement functionalities that directly address particular challenges and provide a tailored solution for the users.

In pursuit of this objective, three key modules were created for an ongoing project known as CORE[1] Cardio-Oncology, a web application that is currently being developed by Holos, S.A.[2] in close collaboration with the medical personnel from the Department of Oncology at Puerta de Hierro Majadahonda University Hospital[3]. These components aim to augment the project's functionality and expand its capacity to facilitate interactive engagement and knowledge dissemination for its users.

In order to create these modules, a comprehensive analysis of the project needed to be conducted, examining both its back-end and front-end systems, overall architecture and the programming languages employed. This involved studying the existing infrastructure, understanding the data flow between different components and identifying the key functionalities required for the new modules. Additionally, thorough research was conducted to evaluate the most suitable technologies and frameworks to ensure compatibility and seamless integration with the existing platform.

[1] https://core.holos.pt/
[2] https://www.holos.pt
[3] https://www.comunidad.madrid/hospital/puertadehierro/

3.2 Project Base

Project CORE is an ongoing development initiative, being constructed utilizing the React JavaScript[4] library. React is a popular and widely adopted open-source JavaScript library that facilitates the creation of dynamic and interactive user interfaces [2]. Its component-based architecture promotes modular development, allowing developers to build reusable and encapsulated User Interface (UI) components. React's virtual Document Object Model (DOM) efficiently updates and renders UI elements, resulting in enhanced performance and responsiveness [2].

React's focus on performance optimization ensures that the application remain fast and responsive even as it scales in complexity. Through techniques such as:

- **Virtual DOM Diffing:** Where React creates virtual representations of the DOM as trees, Virtual DOM trees. When changes occur in the aplication state, React generates a new Virtual DOM tree representing the updated UI. To update the actual DOM, React performs a process called *diffing* between the previous Virtual DOM tree and the new one. This involves comparing the two trees and identifying the minimal set of DOM operations required to synchronize the real DOM with the virtual representation.

- **Memoization:** A technique that optimizes the performance of functions by caching the results of expensive function calls and returning the cached result when the same inputs occur again. By *memoizing* the results of function calls, React avoids recomputing values when the inputs remain unchanged.

By utilizing these techniques, React minimizes unnecessary DOM updates and re-renders, resulting in a smoother user experience.

Additionally, Project Core leverages Google App Engine[5], part of Google's PaaS offerings, a serverless platform, for streamlined and efficient web application deployment. Google App Engine enables developers to build and deploy web applications with ease, without the need for managing servers or scaling infrastructure.

In tandem with React, CORE uses several products from the Google Cloud suite for the backend, from Firebase's user authentication, to BigQuery's transactional database Cloud SQL storage. The combined utilization of React and Google's Cloud Suite of products exemplifies the project's commitment to leveraging cutting-edge technologies for building a highly dynamic, performant and manageable system. This approach ensures that the project not only meets its objectives but also delivers a superior user experience to its intended users.

[4]https://react.dev/

[5]https://cloud.google.com/appengine

3.3 Developed Components

As previously mentioned, the focus of this thesis was the development and integration of essential modules: the Questions and Answers module, the Medication Management module and the Quality of Life module.

The **Questions and Answers Module** was engineered to provide users with a robust platform for users seeking information and insights on a diverse array of topics, ranging from dietary concerns to managing work-life balance amidst cancer treatments. Its functionalities include features for efficient question management and submission, ensuring users could easily access relevant information. Moreover, the module facilitated interactions with healthcare professionals, who are the only users who can provide answers to the questions, thereby guaranteeing the accuracy and reliability of the information shared.

With a firm commitment to fostering a dynamic and interactive community-driven environment, the module incorporated user engagement mechanisms to encourage active participation and knowledge sharing among users, and recommendation of questions based on user views using AI knowledge discovery algorithms, further personalizing the learning experience and promoting deeper engagement.

The primary objective of the module is to empower patients and individuals with medical-related inquiries, enabling them to deepen their understanding of their condition and make informed decisions regarding their health and well-being. By facilitating access to expert advice and fostering meaningful interactions, the Questions and Answers module aims to enhance users' overall well-being and quality of life.

The **Medication Management Module** was meticulously designed to simplify the multiple tasks related to medication management. It displays a comprehensive list of the user's various medications, with information such as start and end dates, dosages, administration routes and designated intake times. Alongside this, there is an intuitive tracker functionality that enables users to seamlessly update their medication intake records, ensuring accurate and up-to-date information regarding their dosage adherence. This tracker not only aids in monitoring medication consumption but also provides valuable insights into adherence patterns and trends.

Additionally, the module includes a proactive alert system that highlights instances of missed medication intake, thereby promoting accountability and adherence to prescribed regimens. By flagging missed doses, users are empowered to take timely corrective actions, fostering a proactive approach to medication management and enhancing overall treatment outcomes.

One of the standout features of this module is its integration with Google Calendar, enhancing user convenience by delivering timely notifications directly to their devices. By synchronizing medication schedules with users' calendars, this integration helps users to never miss a dose and remain consistently adherent to their prescribed treatment plans.

Through the integration of these features, the Medication Management module stands as a comprehensive solution tailored to meet the diverse needs of users in managing their medication routines effectively and efficiently.

The **Quality of Life Module** fosters a collaborative approach to well-being management. This user-centric module empowers individuals to actively participate in their health journey by providing a comprehensive platform for self-monitoring and gaining insights into various aspects of well-being.

Through intuitive interfaces, users can effortlessly record daily data points encompassing sleep patterns, blood pressure, heart rate, exercise routines, pain levels, and overall well-being. This data can be visualized with interactive graphs, revealing trends and patterns over time not only by users, but also healthcare professionals. This allows for collaborative goal setting, informed decision-making, and personalized guidance based on the user's unique health picture.

Together, these components significantly contributed to the project's functionalities, serving as essential components for its operations. The modules were developed with a strong focus on user-centric design and functionality, aiming to provide users with the tools and resources they need to lead healthier and more informed lives. Through their integration into the project, these modules significantly contributed to the project's overarching mission of enhancing users' overall health and life satisfaction.

Additionally, it's noteworthy that these modules are designed to accommodate users across diverse linguistic backgrounds, with support for English, Spanish and Portuguese languages. This multilingual approach ensures inclusivity and accessibility for a broader user base, allowing individuals from different regions and language preferences to engage seamlessly with the platform. By offering support for multiple languages, the project aims to break down language and regional barriers and provide equitable access to valuable healthcare resources and information.

Currently, and throughout their development, these modules were actively tested by patients and healthcare professionals. This user-centered testing approach provided invaluable insights that directly shaped the design and functionality of the modules.

3.4 Modules Architecture

For the integration and correct implementation of these modules in the project, a clear and modular architecture needed to be followed, not only to provide seamless interaction between modules but also between pacients and healthcare professionals with the web application.

The overall architecture of the system follows a modular design, where each component serves a specific function while maintaining interoperability with other components. Figure 3.1 illustrates this architecture.

The modules utilize the React framework as their foundation. Data storage and retrieval are managed through Google Cloud Platform Firestore and BigQuery. Additionally, specialized tools and services, such as Google Calendar integration and Python scripts deployed on Google Cloud Functions, are incorporated to extend the core functionality of the modules. Further explanations regarding the architecture of each module will be provided in the subsequent chapters.

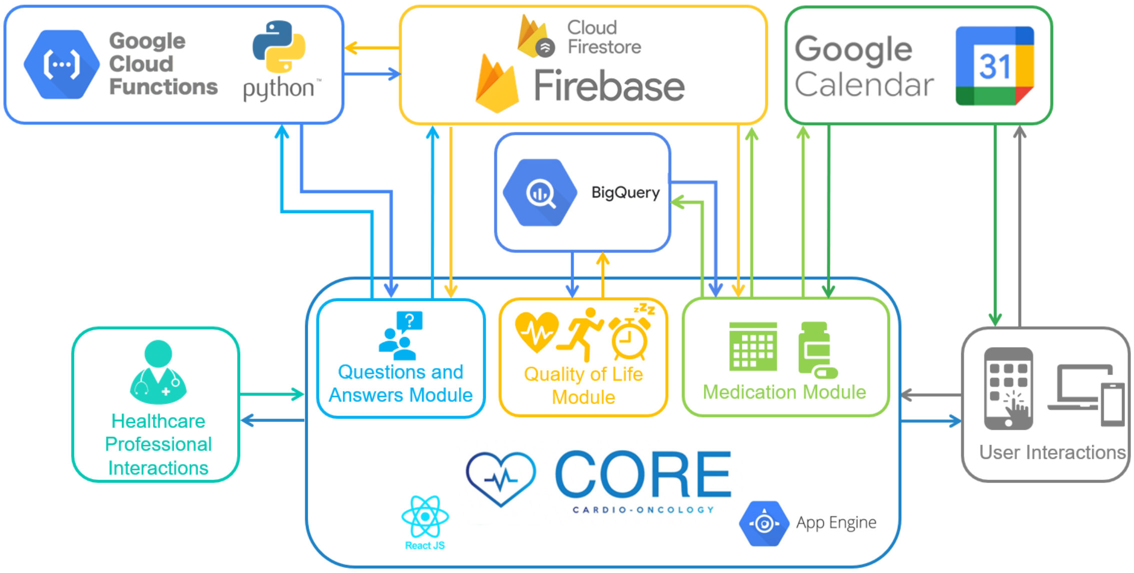

Figure 3.1: Modules Architecture.

3.5 Questions and Answers Module

The Questions and Answers module is designed to cater to the diverse needs of pacients seeking reliable information and expert guidance on a plethora of topics. In this module, users can pose questions, examine answers provided by healthcare professionals and learn valuable insights from curated content. At the heart of the module's functionality lies the ability for users to browse through the wealth of information available, whether through exploring trending topics, browsing popular questions, or discovering tailored recommendations selected to cater to their interests and needs. By fostering an environment conducive to collaborative learning and knowledge exchange, the Questions and Answers module aims to empower users in their quest for knowledge and to further enhance their understanding of various healthcare-related matters.

As Project CORE is still in development, with a prototype already in operation, it's important to note that the data available within the Questions and Answers module represents a base set of questions and answers. These questions were sourced from real patients from the Puerta de Hierro Majadahonda University Hospital and the responses were provided by qualified healthcare professionals from the hospital's Oncology Department. While the platform evolves, this initial dataset serves as a robust starting point, laying the groundwork for future expansion and refinement as the application continues to grow.

3.5.1 Architecture

At the core of this module lies React, which serves as its foundational framework and, as referenced in Chapter 3.4 the module relies on Firestore to handle data storage and retrieval. Firestore was specifically chosen for its capabilities in near real-time synchronization and its flexibility in handling structured data. The interactions with Firestore are facilitated through the utilization of the official Firestore NPM package[6] that handles communication between React and the Firestore database, ensuring reliable and secure data handling processes.

Moreover, to enhance the module's functionality, Google Cloud Functions running on Python 3.9 are incorporated into the system. These functions are triggered through HTTP requests, and, based on the data received in the payload, these functions generate personalized recommendations tailored to users' preferences and needs. These recommendations are then sent back to the React component, to be displayed to the users. The Functions also access Firestore, querying the database through the Google Cloud Firestore Python library[7]. Figure 3.2 displays the isolated architecture of the module.

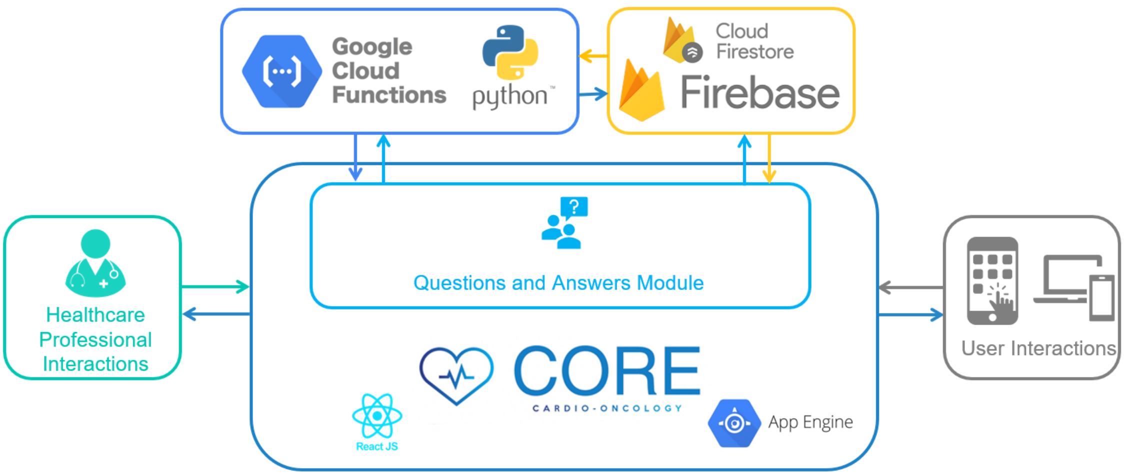

Figure 3.2: Questions and Answers Module Architecture.

3.5.2 Interface

By closely following the front-end templates designed by the development team in charge of Project CORE, it was possible to create the questions and answers module and integrate it to the web application ensuring a polished and intuitive user interface. The template is based on Material UI[8], a React component library and it served as a starting point, providing a pre-designed structure and visual components that can be customized and tailored to meet the specific requirements of the page functionality. This structure consists

6 https://www.npmjs.com/package/@firebase/firestore
7 https://pypi.org/project/google-cloud-firestore/
8 https://mui.com/

of a front page with four tabs, "Popular", "Recent", "Search" and "Recommended" that organize the questions according to different needs. Figure 3.3 shows this structure.

As referenced in chapter 2.5, according to the Principle of Attractiveness, aesthetically pleasing designs can capture users' attention and evoke positive emotions. Taking this into consideration, the interface was designed to be visually appealing in order to provide a comforting and supportive environment for patients seeking answers to their questions and concerns. Details such as the use of calming color schemes, soft edges and intuitive navigation elements supports patients in their information-seeking process, ultimately contributing to a more positive and empowering user experience.

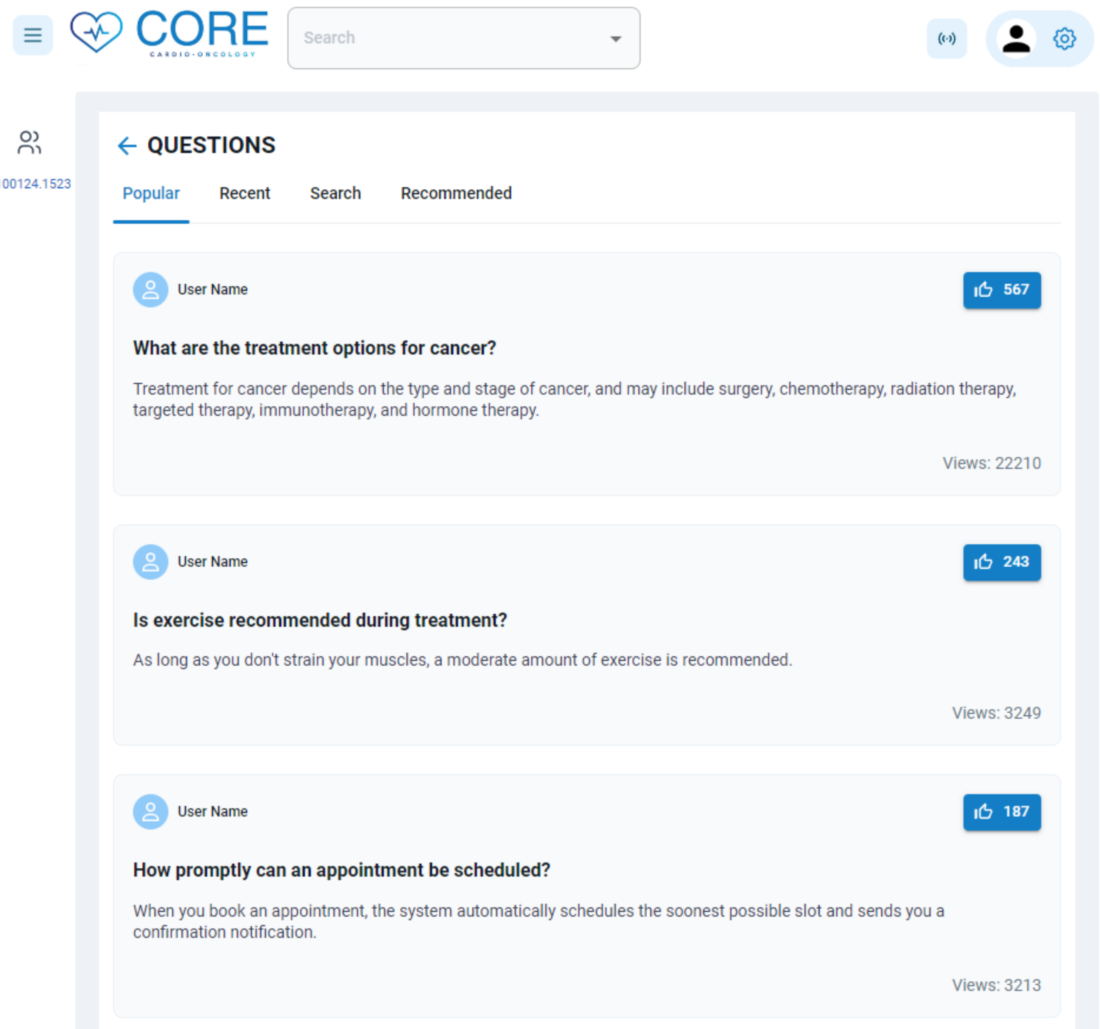

Figure 3.3: CORE's Questions and Answers page.

The module's front page displays a comprehensive list of questions alongside their respective answers, views, like counters and buttons, complemented by the usernames and avatars of the posters. Among the initial implementations were the integration of a like button functionality and a counter to track the number of likes garnered by user-generated content. Additionally, a views counter was introduced to monitor the number of times a question has been viewed. These functionalities not only enable users to express their appreciation for valuable contributions but also facilitate the identification of popular or highly regarded posts within the platform. Moreover, they provide valuable insights into the visibility and engagement levels of specific posts, enriching the overall user experience.

3.5.3 Data Storage and Retrieval

With the foundations of the interface and its core features implemented, the next step in the process was to transition static, hard-coded data to a dynamic and scalable solution for data storage and retrieval. In order to achieve this, Google Firebase's Cloud Firestore[9] NoSQL cloud database was chosen to to store and sync data for client and server-side in near real-time.

Unlike traditional relational databases, Firestore follows a NoSQL data model, allowing for flexible document-oriented storage and efficient querying of data. Firestore organizes data into collections and documents, where each document contains a set of key-value pairs representing the data. Collections can contain multiple documents and Firestore allows for hierarchical structuring of data to accommodate complex relationships.

One of the key features of Firestore is its near real-time data synchronization capabilities. Changes made to data in Firestore are automatically propagated to all connected clients in near real-time, ensuring that applications remain up-to-date and responsive. This near real-time synchronization is achieved through the use of web sockets, enabling efficient communication between clients and the Firestore backend.

Furthermore, Firestore offers robust security features to protect data integrity and privacy with Identity and Access Management (IAM). IAM allows administrators to define fine-grained access controls and permissions at the level of Firestore collections, documents and individual fields. This ensures that only authorized users or services can access and modify specific data, reducing the risk of unauthorized access or data breaches.

To accomplish this, a new collection was created, solely for the Questions and Answers (Q&A) module.

Figure 3.4 shows this structured collection in Firestore. In this collection, every question is a document whose name is the question's unique ID. Each document contains fields for the question, the answer, information about the user who posted it, a timestamp, counter for likes and views, associated language and other pertinent information. Coupled to these question documents are subcollections dedicated to tracking user engagement metrics. Specifically, two subcollections, one for likes and another for views, complement the main question collection. These subcollections enable the storage and management of data related to user interactions with the questions. Each interaction, whether it be a like or a view, is recorded within these subcollections. Each document essentially represents a JavaScript Object Notation (JSON) object, with fields representing different attributes of the question and its associated information (Figure 3.5 shows a simplified example of the

[9] https://firebase.google.com/docs/firestore

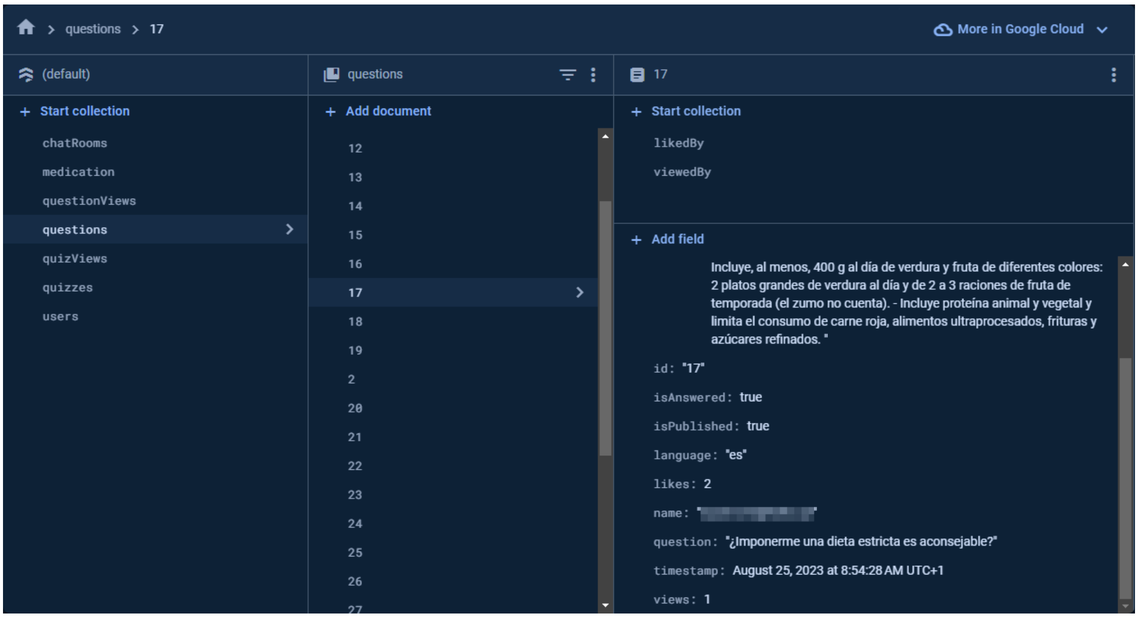

Figure 3.4: Firestore Questions and Answers Collection.

document's JSON representation). In each subcollection, a document is created, for each user that interacts with the question, in this document, the user ID and email is stored, Figure 3.6 shows this structure.

```
questions = {
    "questionID1": {
        "question": "Should I exercise?",
        "answer": "Exercise is always recommended.",
        "id": "questionID1",
        "isAnswered": true,
        "isPublished": true,
        "language": "en",
        "likes": 2,
        "name": "userName",
        "timestamp": "July 2, 2023 at 9:13:21 AM UTC+1",
        "views": 5
    }
}
```

Figure 3.5: Simplified example of the documents' JSON representation.

Firestore's JSON-like structure lends itself well to storing varied data such as this within documents. Additionally, the support for indexing and querying enables efficient retrieval and manipulation of data stored within these documents.This functionality enables the

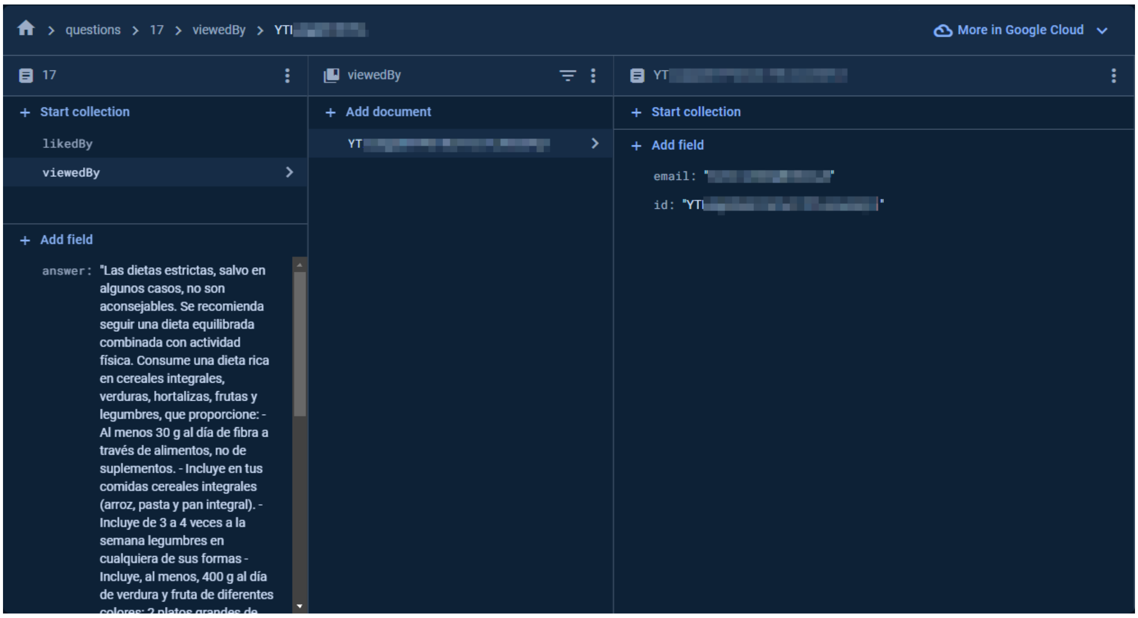

Figure 3.6: Firestore Question Document Subcollections.

execution of intricate queries to filter, sort and aggregate data based on diverse criteria, enhancing responsiveness and flexibility in data management.

3.5.4 User's Questions Management

After assuring the correct retrieval and exposition of questions with Firebase, it was deemed necessary to create a distributed system for platform administrators in order to curate questions, answer them and finally publish the respective answers. Figure 3.7 shows the interface where users can ask their questions. As any user can post questions, it was necessary to implement measures to filter the submission of duplicate or off-topic inquiries that would otherwise inundate the platform with irrelevant data which would diminish the overall quality and reliability of the platform.

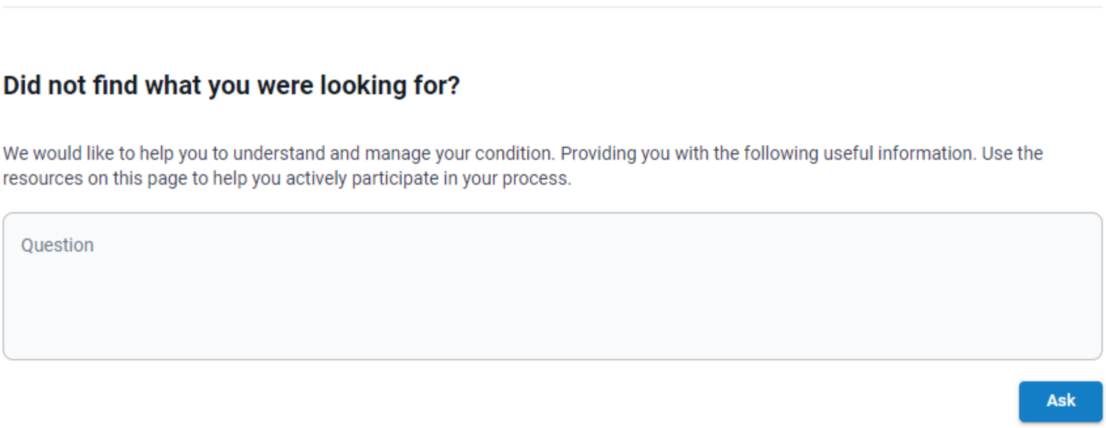

Figure 3.7: Question Asking Functionality.

In order to uphold the cohesion and validity of responses, the answers to users' questions are provided by qualified medical personnel. This not only ensures the accuracy and reliability of the information imparted but also instills trust and confidence among users seeking authoritative guidance and support, remaining aligned to the Principle of Authority referenced in chapter 2.5.

These administrators can have specialist and editor privileges, where specialists can answer questions and editors can publish or reject answers. These functionalities are offered in additional tabs that are hidden to regular users, showing the possible options according to the user's privileges. Figures 3.8 and 3.9 show the capabilities offered to specialists and editors respectively. Given the implementation of user permissions, a dedicated page was created to facilitate the management of these authorizations. Figure 3.10 shows this service.

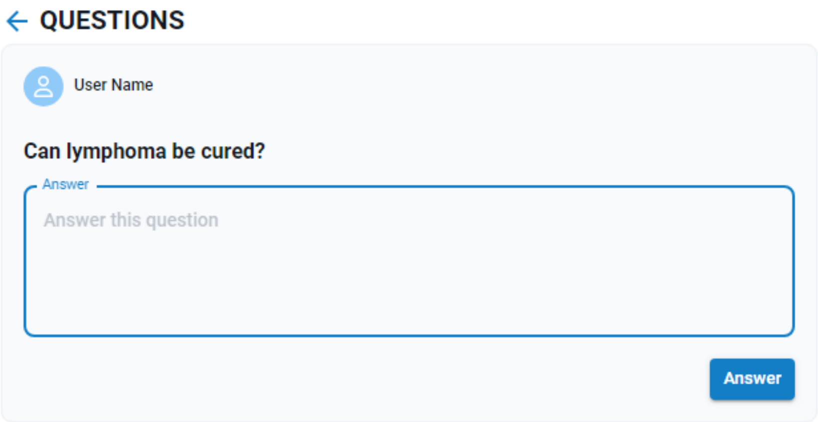

Figure 3.8: Question answering functionality.

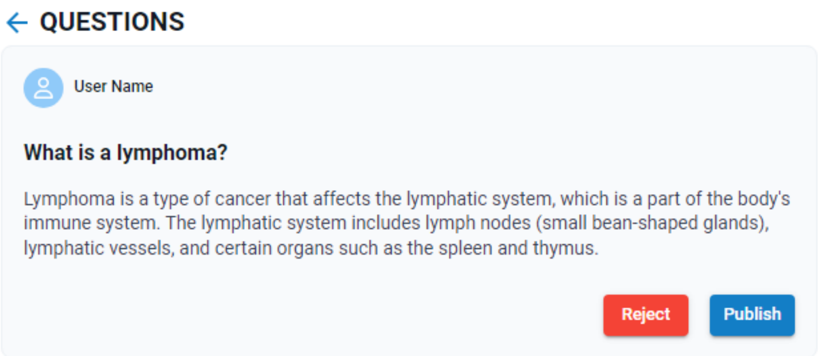

Figure 3.9: Question publishing functionality.

As the platform supports multiple languages, the data displayed was limited according to the user's selected language, rather than solely translating all of the questions and

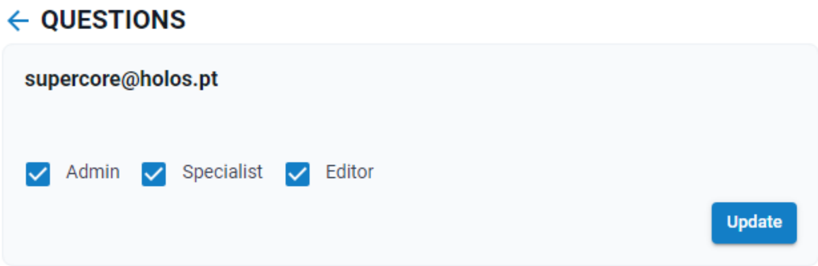

Figure 3.10: Authorizations modification functionality.

answers. By restricting questions to specific languages, we can better address cultural sensitivities and nuances inherent in patients' inquiries and ensure accurate communication of complex medical terminologies, avoiding potential misunderstandings or inaccuracies that may arise from direct translations.

3.5.5 Question Recommendations

With the foundation of the Questions and Answers page established, the focus shifted to the more promenant feature of this page, the ability to show recommendations of questions based on users' prior views. This functionality aims to personalize the user experience, offering tailored suggestions that align with individual preferences and browsing history. By analyzing users' interactions with the platform, the recommendation system facilitates the discovery of relevant content, enriching engagement and fostering a more intuitive and rewarding user journey. Figure 3.11 displays this tab.

As patients navigate the page, the recommended information will progressively align more closely with their individual needs and preferences. This personalized curation extends to accommodating diverse stages of treatment and addressing varied interests, spanning from dietary considerations to exercise regimens and managing work-life balance. Whether individuals are seeking advice on symptom management, treatment options, or lifestyle adjustments, the recommendation system adapts to offer pertinent resources tailored to their unique circumstances based on user interactions with the available data. This approach aims to provide support across all the possible steps of their treatment journey, empowering them to make informed decisions and enhance their overall well-being.

When a user enters the Q&A page, a HTTP request is created using JavaScript's Fetch API with a JSON payload containing the user's ID which is then sent to a Google Cloud Function tasked with generating the recommendations.

This Google Cloud function creates question recommendations using Association rules mining algorithms. In order to achieve this, the function is running Python 3.9 with the following libraries:

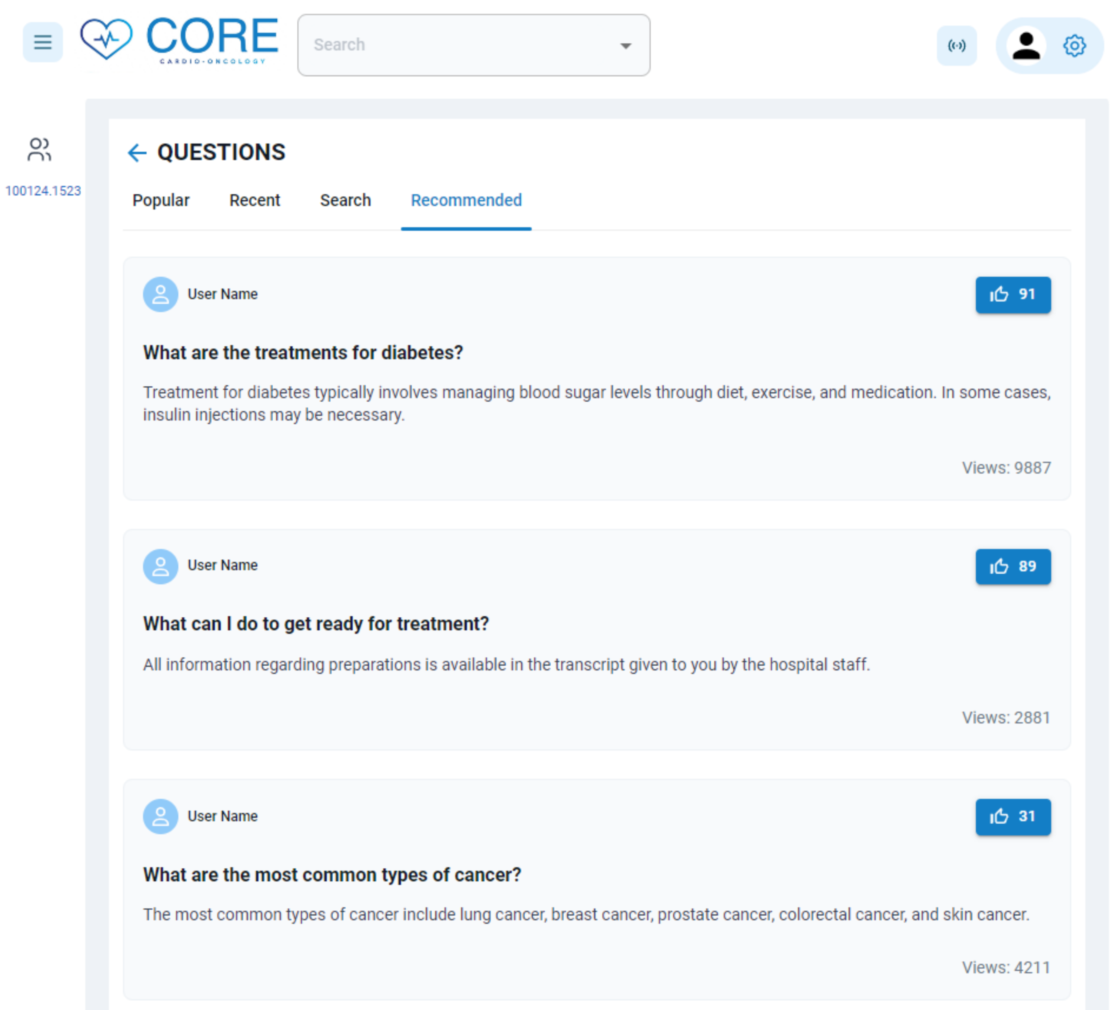

Figure 3.11: Recommended tab of Questions and Answers page.

- **Apyori:** chosen for its efficient implementation of the Apriori Algorithm.
- **Firebase:** used in order to retrieve data from Firestore.

The Apriori Algorithm was selected as the prime candidate for association rules generation due to several reasons mentioned in chapter 2.4.3. The AIS method was ruled out as it exhibits lower efficiency and slower performance while also limiting itself to generating solely one-item consequent association rules, the Apriori Algorithm stands out for its overall speed and capability to generate multi-item consequent association rules efficiently. Although the FP-Tree method is faster and more efficient, it lacks adaptability for incremental mining, making it less suitable for dynamic datasets requiring continuous updates.

Upon initialization, the function first extracts the user ID from the request payload. Subsequently, it retrieves the question views for each user, storing this information in an object. Concurrently, when retrieving this information, when users collection user ID is captured, its data is stored in another object. The function then proceeds to apply the

Apriori Algorithm, which generates a list of recommended questions. From this list, the questions that were originally viewed by the user are filtered out, leaving behind only the new recommendations. This pruned list is then returned to React for presentation to the user. It's important to note that if the user has already viewed all recommended questions, an empty list will be returned. In such cases, the module will dynamically switch to displaying the most popular questions, ensuring continued engagement and content relevance for the user. Figure 3.12 shows the process with simplified data.

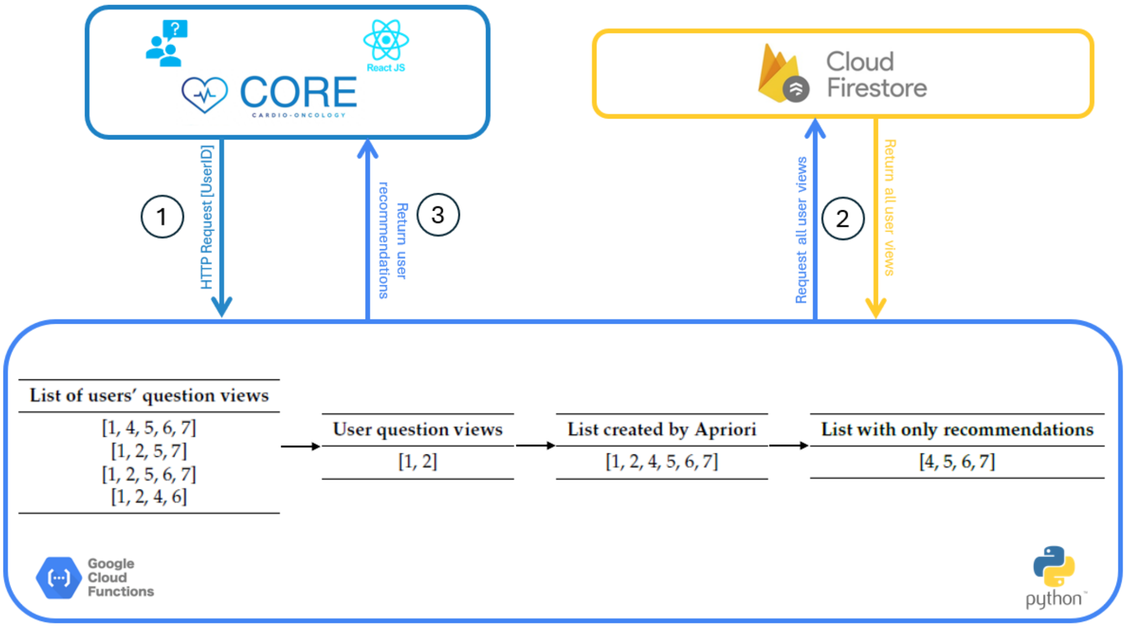

Figure 3.12: Simplified example of the Recommendation Google Cloud Function Process.

For the Apriori Algorithm, a minimum support of 5% and a minimum confidence of 20% were chosen as they strike a balance between generating relevant recommendations while also ensuring that the recommendations are sufficiently reliable. These thresholds were selected after careful consideration of the dataset and user requirements.

3.6 Medication Management Module

The Medication Management module is designed to aid users in the processes associated with medication-related tasks, catering to the specific needs of patients navigating complex treatment regimens. Within this module, users gain access to create a comprehensive list of their medications, complete with details such as start and end dates, dosages, administration routes and intake times. Additionally, the module features a medication intake tracker, allowing users to intuitively monitor their medication adherence and track their progress over time. By adding integration with Google Calendar, the module also provides timely notifications and reminders to help users ensure that they stay on track with their medication schedules.

3.6.1 Architecture

Similarly to the Questions and Answers module, this component is also built upon the React Framework, benefiting from its versatility and robustness in creating dynamic user interfaces. Furthermore, database functionalities are established by Firestore, chosen for its near real-time synchronization capabilities and flexibility in handling structured data, as detailed in Chapter 3.2 for the Q&A component. However, in addition to Firestore, the medication management module incorporates the use of BigQuery for the storage and retrieval of medication intake records data. This addition of BigQuery enhances the module's capabilities by providing a scalable solution for managing large datasets related to medication intake.

Additionally, integration with Google Calendar was established, through the use of the official NPM Google Calendar API Package[10]. Figure 3.13 shows the isolated architecture of the module.

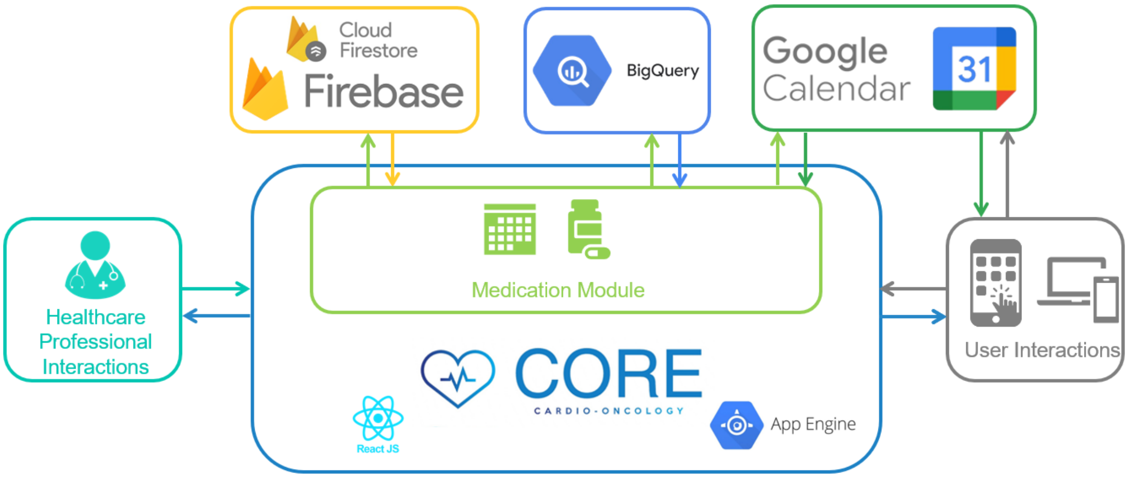

Figure 3.13: Medication Management Module Architecture.

3.6.2 Interface

The design principles of this module follow what was described in the previous module, utilizing the Material UI React library and adhering to the Principle of Attractiveness explained in Chapter 2.5.

This module's front page displays a comprehensive list of medication added by the user with all of its associated information, as well as a summary of the user's adherence to their medication intake routine in the last 7 days. Figure 3.14 displays this page.

In order to display this summary, a custom component was created to render the medication intake data in a visually appealing and user-friendly format. This custom component, utilizes React to dynamically generate a week calendar layout and populate it with symbols that represent medication intake adherence information for each day. Three

[10] https://www.npmjs.com/package/react-google-calendar-api

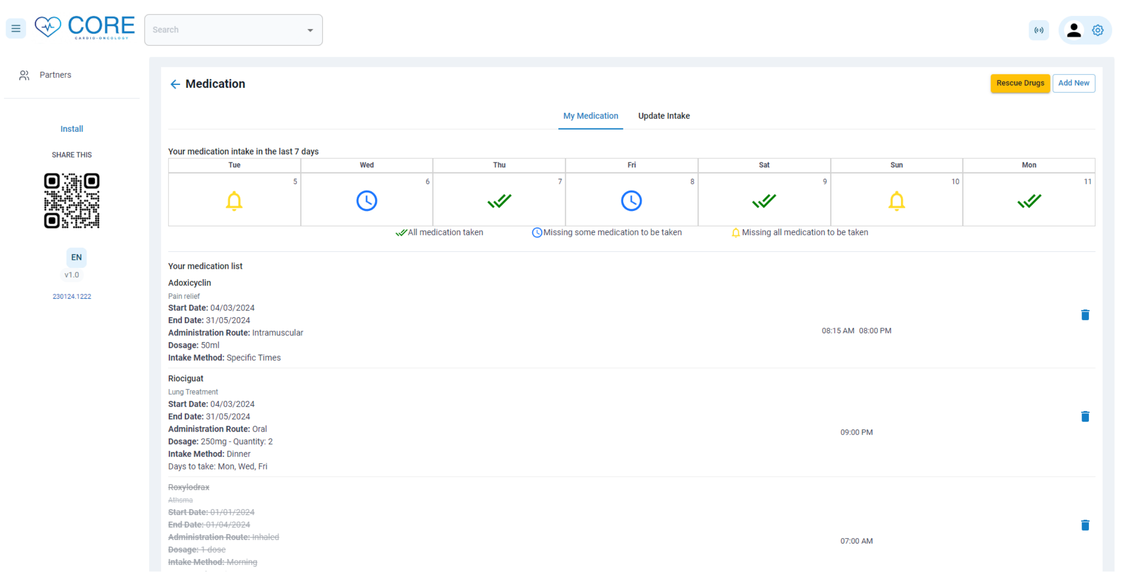

Figure 3.14: Medication Management Module Front Page.

icons represent medication adherence status: a yellow bell icon indicates no medication was taken that day, a blue clock signifies that at least one of the medicaments was taken and a green checkmark indicates all prescribed medications were taken. Figure 3.15 shows this behaviour.

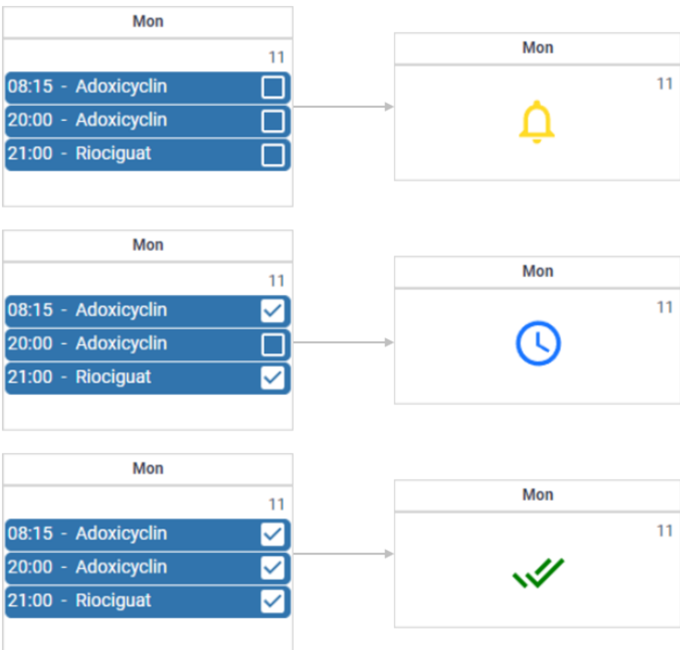

Figure 3.15: Symbols displayed according to daily medication adherence.

To add new medication to the list, users can click on the "Add New" button, which then redirects them to the corresponding page for medication addition (Figure 3.16). In this page, pacients are prompted to input details such as the start and end dates for the treatment, the name of the drug, the appropriate administration route (Oral, Intravenous, Intramuscular, Subcutaneous, Transdermal and Inhaled). If the selected administration route chosen is oral, a drop down menu is available in order to select the number of pills to be taken. Other input fields include the associated dosage (50mg, 15ml, etc), a brief description for the treatment, followed by a checkmark that indicates whether the medications is supposed to be taken in specific days of the week (only mondays and fridays) instead of everyday. Figure 3.17 shows this selector.

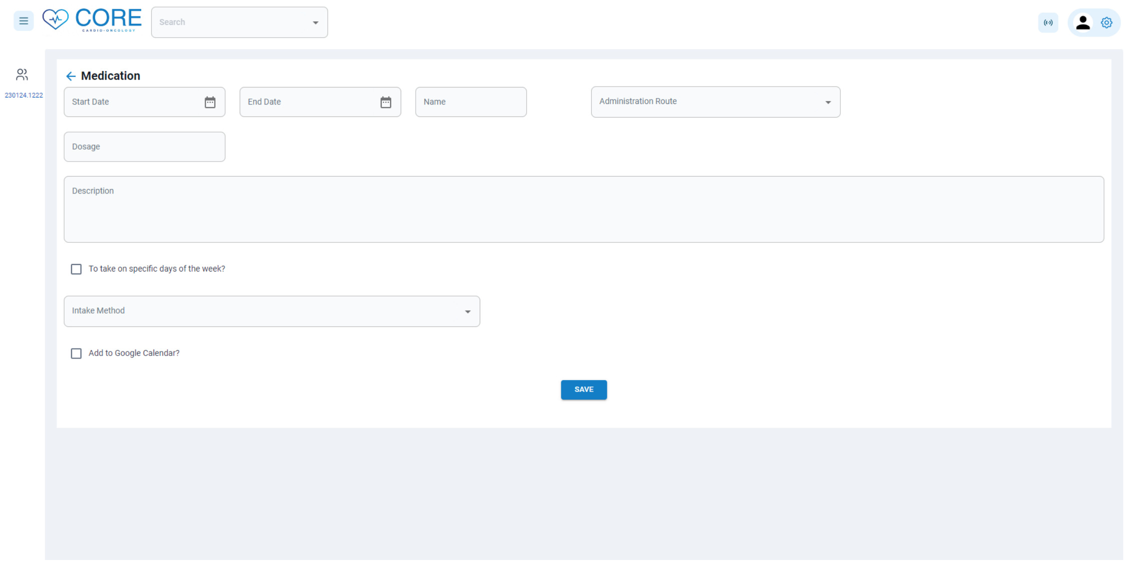

Figure 3.16: Medication Management Module Add New Medication Page.

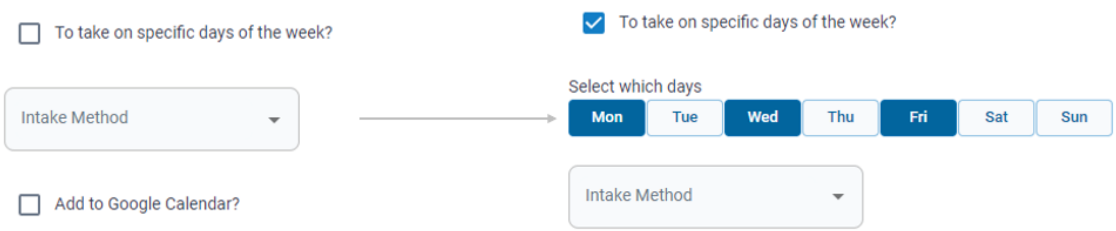

Figure 3.17: Dynamic selector for specific days.

The next input field is a selection for intake methods, to choose from a list of (specific times, morning, breakfast, lunch, dinner, before sleep and to take on an empty stomach). The specific times option allows the user to select multiple times for a certain medication to be taken (e.g. 08h15 and 20h00), for this to be accomplished, a dynamic input is displayed, where the patient can increment or decrement the number of intakes for the day, then, for each, a time selection box is presented. For the remainder of the options, the selection box is only presented once. Figure 3.18 shows this dynamic input.

Lastly, a checkmark is present to allow the user to add a reminder to Google Calendar. If selected, when the medication is saved, a popup window will appear (Figure 3.19) that prompts the user to sign in with a Google account in order to add the the reminder to their calendar. When the medicine is added, the user is brought back to the front page

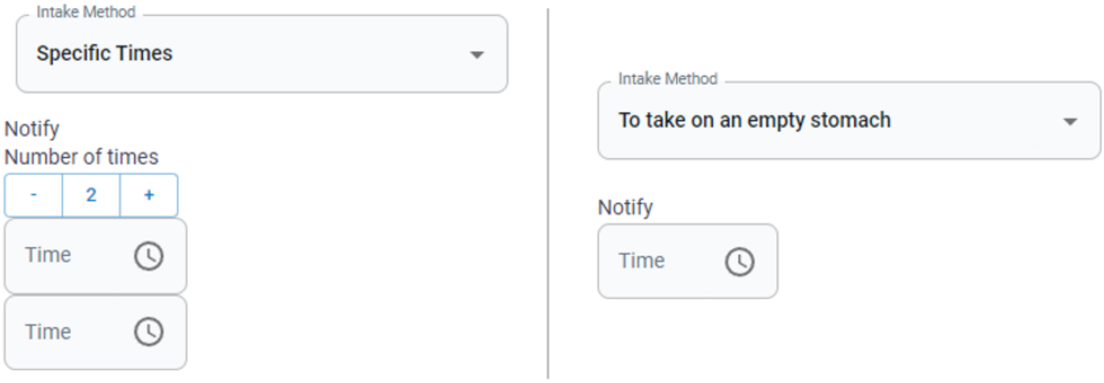

Figure 3.18: Dynamic selector for specific times.

and events are created (according to the chosen days and times) in the patient's Google account calendar. Figure 3.20 shows an example of an event and Figure 3.21 shows how the user's calendar will look with multiple reminders for different medications.

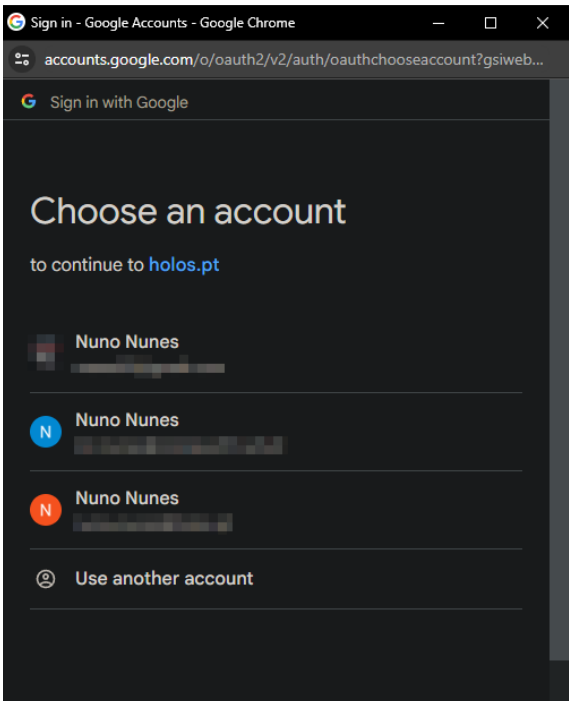

Figure 3.19: Google Sign In Popup Page.

To keep track of medication intake, a dedicated tab was created where the user is shown the scheduled intakes for the last 7 days (Figure 3.22). Here, the previously introduced custom week calendar component is also used, but, opposed the summary displayed in the front page, in this tab individual intakes are shown. The individual intakes displayed in each day show the time of intake, the medication name and an interactive checkmark indicating whether the medicine was taken. However, presenting this information on mobile devices in portrait mode posed a challenge due to limited horizontal real-estate. To address this, a separate component (Figure 3.23) was used specifically for mobile users,

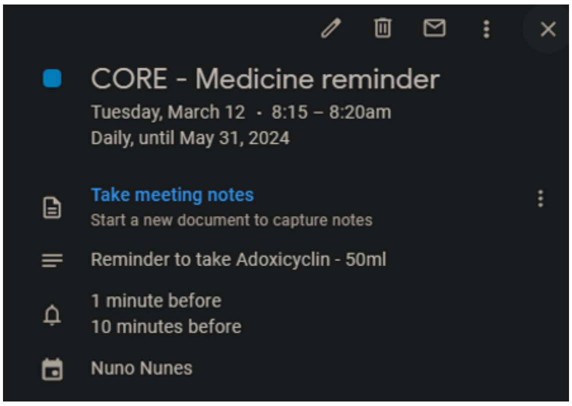

Figure 3.20: Medication Reminder Calendar Event.

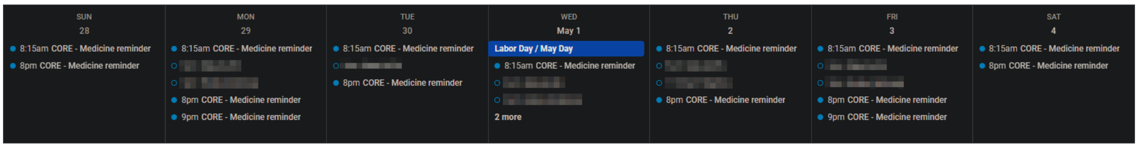

Figure 3.21: Medication Reminders in Google Calendar.

organizing the data vertically to better display the week calendar. Users can modify the intake status by toggling the checkmarks to reflect their adherence and then save the changes by pressing the update button. When the intakes are updated, the patient is brought back to the front page, where the summary is already updated.

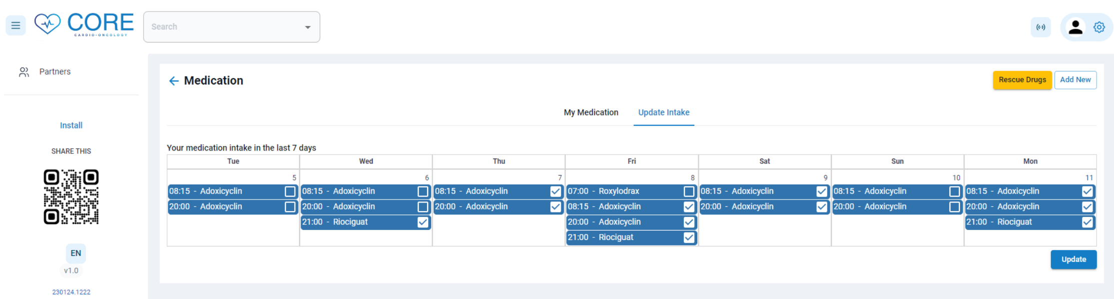

Figure 3.22: Medication Management Module Medication Intake Update Tab - PC.

For healthcare professionals to be able to browse patient's medication information and add new treatments, a new tab was created, whose access is restricted to healthcare professionals. In this tab, the patient's medication list and medicine intake adherence over the past 28 days are displayed. The medication list showed is precisely the same as the one displayed to the users in their medication page, however the intake adherence calendar differs, the custom calendar component is now extended to be able to display 28 days and it cannot be interacted with, as it is read-only. In order to add new medication to a user, Medical personel can press the 'Add New' button, redirecting them to the previously mentioned medication addition page. To note that the ability to synchronize with the Google Calendar is then up to the user, as when new medicine is added, a notification appears and the user can choose if they would like to create calendar events or not.

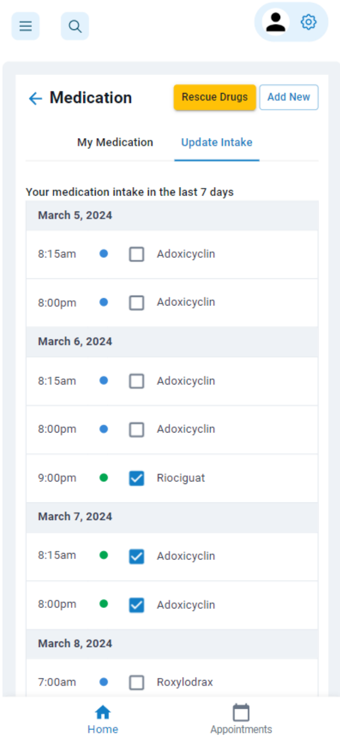

Figure 3.23: Medication Management Module Medication Intake Update Tab - Mobile.

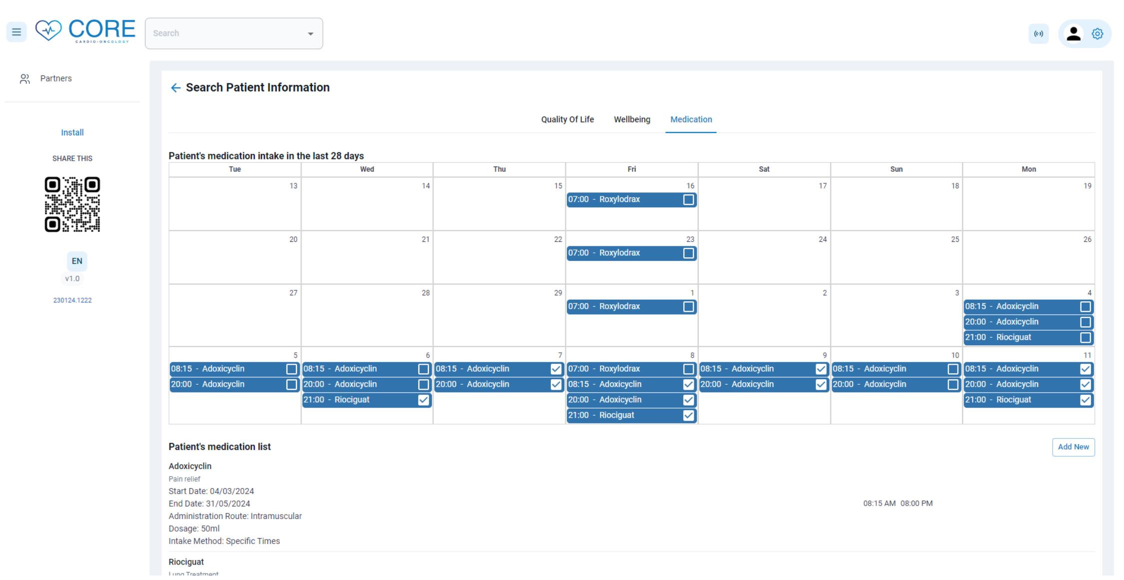

Figure 3.24: Medication Management Module Doctor Preview Page.

3.6.3 Data Storage and Retrieval

For user data storage and retrieval, Firestore was chosen as the primary database solution within the medication management module, aligning with the rationale described in Chapter 3.5.3 for the Q&A module. For this data, a new collection was created in Firestore database dedicated to medication management. Within this collection, a document is created for each user utilizing the module, containing essential user identifiers such as the user ID and email. Additionally, a subcollection is nested within each user document to organize their medication data effectively (Figure 3.25). Within this medication sub-collection (Figure 3.26), crucial information pertaining to the treatment plan is stored, including details filled by the user in the 'Add New' page, as outlined in Chapter 3.6.2. This data encompasses various fields such as medication name, dosage, administration route and intake schedule. Additionally, each medication entry is associated with a unique identifier for the drug and Google Calendar IDs. These Calendar IDs serve a crucial role in managing treatment reminders, allowing the automatic removal of calendar events associated with treatments when they are discontinued.

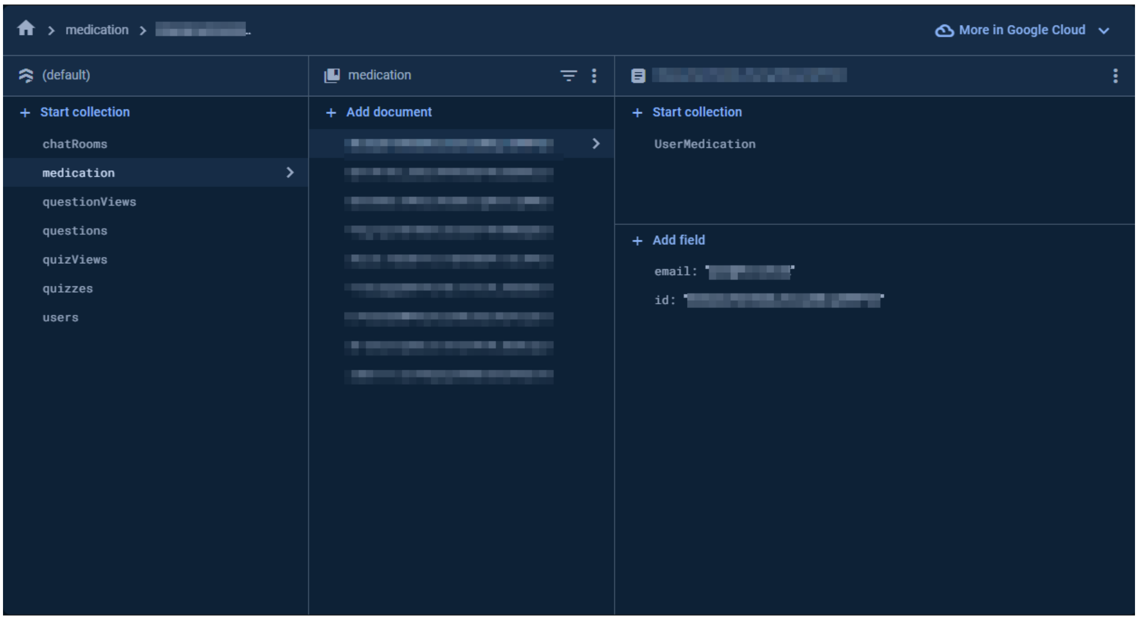

Figure 3.25: Firestore Medication Management Collection.

For monitoring users' medication adherence, Google BigQuery was selected due to the specific characteristics of the stored data. As oncological patients often have long-term drug treatments spanning several years, the volume of data entries for each user could potentially reach thousands. Given this context, BigQuery's SQL structure provides a scalable and efficient solution for storing and retrieving large volumes of structured data. With BigQuery, the focus is primarily on data storage, ensuring that medication intake records are securely and reliably stored for future reference and analysis as needed.

Entries for medication adherence are simple, they consist of the user's ID, the drug's name and ID, the date and time and a flag that indicates wether it was taken or not. Figure 3.28 shows the BigQuery interface with a few example entries.

3.6.4 Limitations

Although the integration with Google Calendar proved successful, a noteworthy limitation arises due to the ongoing development status of Project Core. Given its current developmental phase, only a maximum of 100 users can utilize the calendar reminders

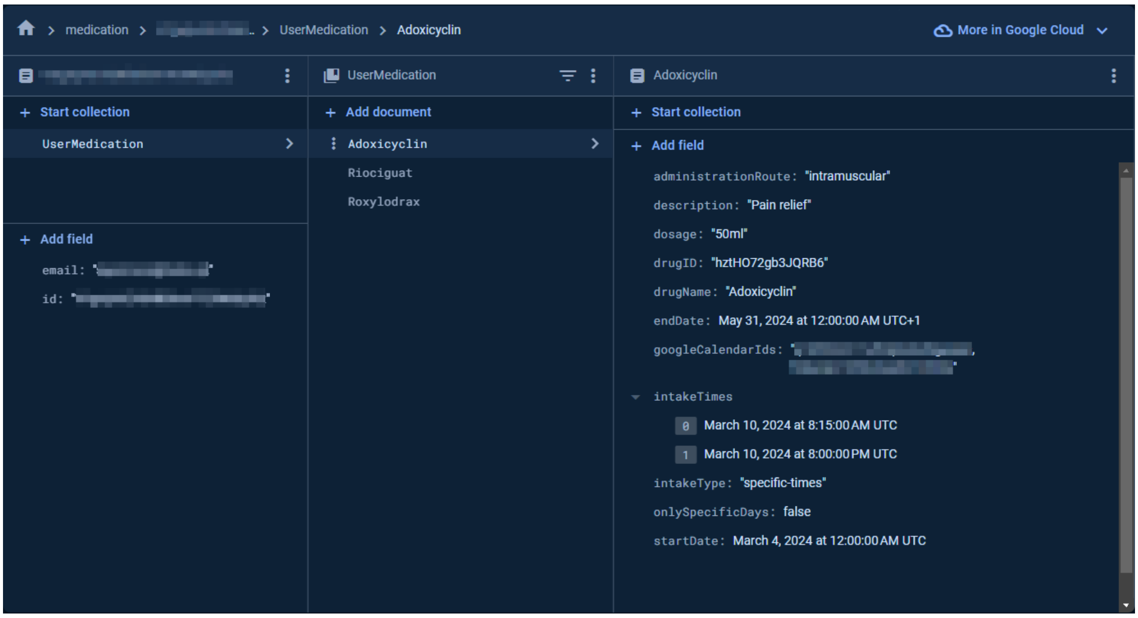

Figure 3.26: Firestore User Medication Subcollection.

feature and these users must be manually added to a whitelist. This restriction is imposed by Google for testing purposes while applications are in development. For broader usage beyond this limit, the web application must go through a verification process conducted by Google.

This restriction is enforced due to the nature of the calendar API, which involves handling sensitive personal data. In our case, we only write data and do not read it, but this access is still considered sensitive. During the development phase, if a new user attempts to activate reminders, a notification screen, similar to the one depicted in Figure 3.29, will appear. This screen informs the user that the app is not verified and that if the user wishes to proceed, they have the option to send an email to the provided address. Then, a developer must manually add the user's email to the whitelist, which will then enable them to freely use the calendar reminder integration.

3.7 Quality of Life Module

The Quality of Life module has been developed to offer users a platform to monitor and manage various aspects of their health and daily living. With this component, patients can input and track key indicators such as exercise routines, pain levels, sleep patterns, cardio health, and general subjective well-being. By enabling users to actively monitor these indicators, the module promotes self-awareness and facilitates communication between users and healthcare providers.

By incorporating visually engaging graphs and simplifying the process of adding new entries, patients are motivated to engage in regular tracking and adopt healthier behaviors. Subsequently, this data becomes accessible to healthcare professionals, in a dedicated page

```
medication = {
    "UserID1": {
        "UserMedication": {
            "AdoxiCyclin": {
                "administrationRoute": "intramuscular",
                "description": "Pain relief",
                "dosage": "50ml",
                "drugID": "genericUID1",
                "drugName": "Adoxicyclin",
                "endDate": "May 31, 2024 at 12:00:00 AM UTC+1",
                "googleCalendarIds": "genericUID2, genericUID3",
                "intakeTimes": [
                "March 10, 2024 at 8:15:00 AM UTC",
                "March 10, 2024 at 8:00:00 PM UTC"
                ],
                "intakeType": "specific-times",
                "onlySpecificDays": false,
                "startDate": "March 4, 2024 at 12:00:00 AM UTC"
            }
        },
        email: "user@email.com",
        id: "UserID1"
    },
    "UserID2": {"data": "..."}
}
```

Figure 3.27: Simplified example of the collection's JSON representation.

with restricted access, providing valuable insights into patients' health and well-being.

3.7.1 Architecture

Unlike the Q&A and Medication Management modules, this component exclusively relies on BigQuery for data storage. The repetitive nature of the data makes it suitable for BigQuery's columnar storage format, optimizing storage efficiency and query performance, ensuring robust and scalable data storage and retrieval.

Data exchange in this module happens through HTTP requests. When data needs to be stored, an HTTP request is sent to the BigQuery API, specifying the data to be stored and its destination within the BigQuery dataset. Similarly, when data needs to be retrieved, HTTP requests are made to the BigQuery API, specifying the query parameters such as filters, sorting criteria, and aggregation functions. Figure 3.30 displays the component's architecture.

Row	userID	drugID	date	time	taken	drugName
1	[illegible]	TwHRIZrPsvINSzb	2023-11-18	06:15	true	Ibuprofeno
2	[illegible]	TwHRIZrPsvINSzb	2023-11-16	06:15	false	Ibuprofeno
3	[illegible]	TwHRIZrPsvINSzb	2023-11-19	06:15	true	Ibuprofeno
4	[illegible]	TwHRIZrPsvINSzb	2023-11-15	06:15	false	Ibuprofeno
5	[illegible]	TwHRIZrPsvINSzb	2023-11-11	06:15	false	Ibuprofeno
6	[illegible]	TwHRIZrPsvINSzb	2023-11-17	06:15	true	Ibuprofeno
7	[illegible]	TwHRIZrPsvINSzb	2023-11-21	06:15	true	Ibuprofeno
8	[illegible]	TwHRIZrPsvINSzb	2023-11-09	06:15	true	Ibuprofeno
9	[illegible]	TwHRIZrPsvINSzb	2023-11-07	06:15	true	Ibuprofeno
10	[illegible]	TwHRIZrPsvINSzb	2023-11-12	06:15	false	Ibuprofeno
11	[illegible]	TwHRIZrPsvINSzb	2023-11-06	06:15	true	Ibuprofeno
12	[illegible]	TwHRIZrPsvINSzb	2023-11-22	06:15	true	Ibuprofeno
13	[illegible]	TwHRIZrPsvINSzb	2023-11-08	06:15	true	Ibuprofeno
14	[illegible]	TwHRIZrPsvINSzb	2023-11-20	06:15	true	Ibuprofeno
15	[illegible]	TwHRIZrPsvINSzb	2023-11-10	06:15	true	Ibuprofeno
16	[illegible]	bHaOCEF9OYKfvFp	2024-03-01	07:00	false	Roxylodrax
17	[illegible]	bHaOCEF9OYKfvFp	2024-01-05	07:00	false	Roxylodrax
18	[illegible]	bHaOCEF9OYKfvFp	2024-03-08	07:00	false	Roxylodrax
19	[illegible]	bHaOCEF9OYKfvFp	2024-02-23	07:00	false	Roxylodrax
20	[illegible]	bHaOCEF9OYKfvFp	2024-03-22	07:00	false	Roxylodrax
21	[illegible]	bHaOCEF9OYKfvFp	2024-01-12	07:00	false	Roxylodrax
22	[illegible]	bHaOCEF9OYKfvFp	2024-02-09	07:00	false	Roxylodrax
23	[illegible]	bHaOCEF9OYKfvFp	2024-02-02	07:00	false	Roxylodrax
24	[illegible]	bHaOCEF9OYKfvFp	2024-01-26	07:00	false	Roxylodrax
25	[illegible]	bHaOCEF9OYKfvFp	2024-02-16	07:00	false	Roxylodrax

Figure 3.28: BigQuery interface displaying medication intake entries.

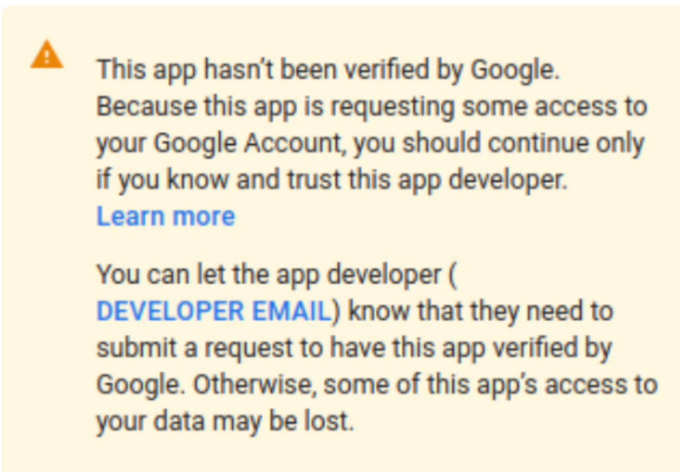

Figure 3.29: Alert screen for unverified app.

3.7.2 Interface

When users enter the quality of life page, they are shown four options, pain, cardio, sleep and exercise. Figure 3.31 shows the page. From here, they can choose which aspect they would like to analyse or register new data.

After selecting one of the options, users are presented with an intuitive graph that displays the latest data according to the aspect chosen, Table 3.1 shows the data displayed for each aspect. Each graph is dynamically generated using React's ApexCharts package[11]

[11] https://apexcharts.com/

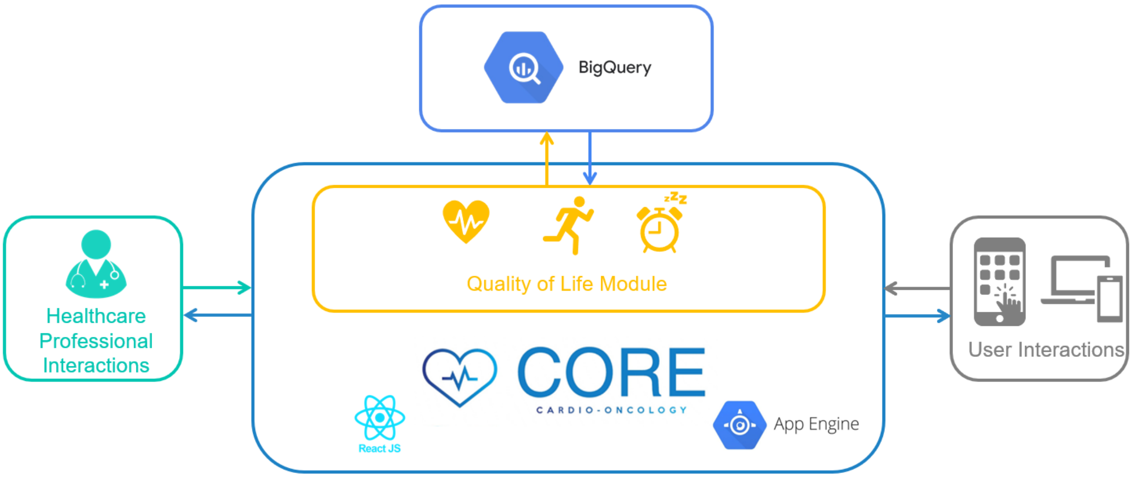

Figure 3.30: Quality of Life Module Architecture.

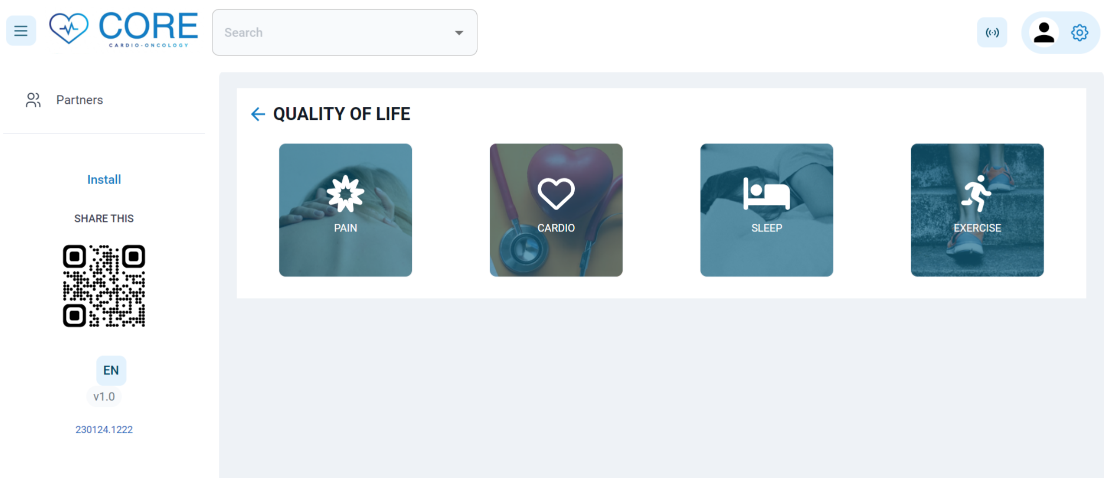

Figure 3.31: Quality of Life Module Front Page.

and provides a visual representation of the relevant data. When hovered over with the mouse, the graph displays more information, offering additional insights into the data points.

Table 3.1: Details displayed for each corresponding aspect

Aspect	Details
Pain	Pain Scale, Location, Type
Cardio	Systolic Blood Pressure, Diastolic Blood Pressure, Heart Rate
Sleep	Quality Level, Duration, Start Time, End Time
Exercise	Weight, Duration, Type, Quantity

The pain graph exhibits a gradient according to the pain level, with colors ranging from blue to red to signify the severity of pain, with blue indicating lower levels and red indicating higher levels.

In the cardio graph, distinct lines represent each metric, including diastolic blood pressure, systolic blood pressure, and heart rate, allowing users to monitor their cardiovascular health trends over time.

The sleep graph illustrates sleep patterns, displaying times and durations, with colors indicating the quality of sleep.

The exercise graph features different bars representing various types of exercises, each color-coded to denote the type of exercise performed. Additionally, a line graph tracks weight evolution, providing users with insights into their exercise progress and weight management journey.

Figure 3.34 shows an example of the graphs of all four options.

These graphs are presented in a timeline format, allowing users to navigate forward and backward in time using the arrow buttons. Furthermore, users can opt to view the data in tabular form by clicking the "Show Table" button, facilitating a detailed examination of individual entries and enabling users to delete selected entries if necessary.

In order to add new data, patients can press the 'Add New' button which redirects them to a dedicated input page where they can conveniently input new data corresponding to the selected aspect. These input pages are built using MUI components, ensuring a consistent and user-friendly interface across all aspects of data entry.

3.7.3 How I Feel

An additional aspect absent from the quality of life page is the "How I Feel" metric, designed to gauge patients' subjective well-being. This metric serves as a pivotal tool in understanding patients' overall health perception and is strategically featured on the front page of the application, serving as a gentle reminder for patients to actively engage in self-assessment by rating their well-being on a scale ranging from 1 to 5 stars. By capturing these insights, patients contribute valuable data points that offer deeper insights into their holistic health journey and aid healthcare providers in delivering personalized care tailored to individual needs and experiences. This data can inform decision-making processes, facilitate early interventions when necessary, and contribute to the ongoing improvement of healthcare services. Figure 3.32 shows the component, displayed in CORE's front page and Figure 3.33 shows the associated history graph.

Figure 3.32: How I Feel Component displayed in CORE's front page.

3.7.4 Care Management Dashboard

To effectively monitor patient quality of life and subjective well-being, a dedicated page has been developed, presenting a comprehensive array of the previously mentioned graphs

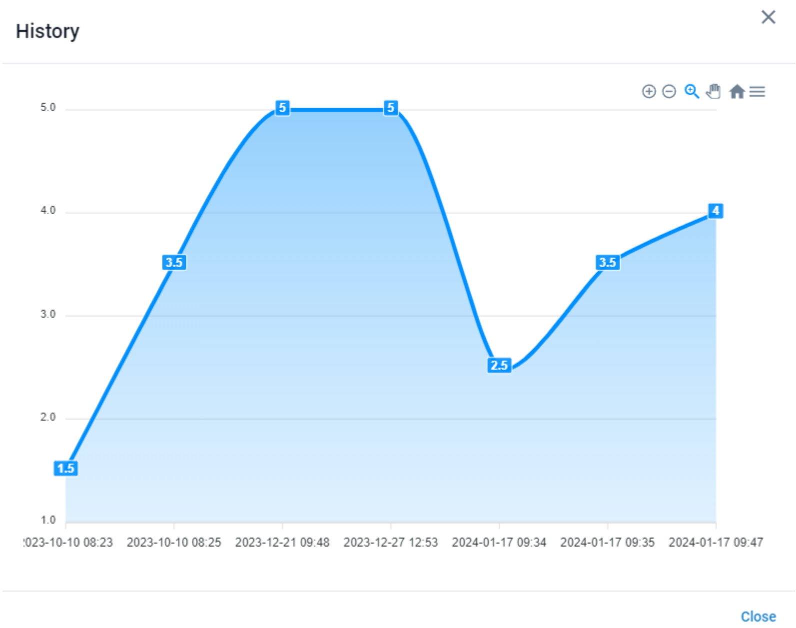

Figure 3.33: Wellbeing Graph.

encompassing various aspects pertinent to patient care. The page offers an intuitive interface, with the same functionalities that are available to the patients, such as navigating through different timeframes to analyze trends over time and the hover-over tooltips that provide additional information.

Accessible exclusively to healthcare professionals, this page serves as a pivotal tool for gaining valuable insights into patient health status and overall well-being. By aggregating multiple aspects onto a single interface, professionals can efficiently assess various dimensions of patient health in a cohesive manner.

In essence, the quality of life monitoring page represents a cornerstone in the provision of patient-centered care, empowering healthcare professionals with the tools and insights necessary to deliver personalized and effective healthcare interventions. Through its comprehensive features and user-friendly design, this page exemplifies a commitment to optimizing patient outcomes and enhancing the quality of care delivery.

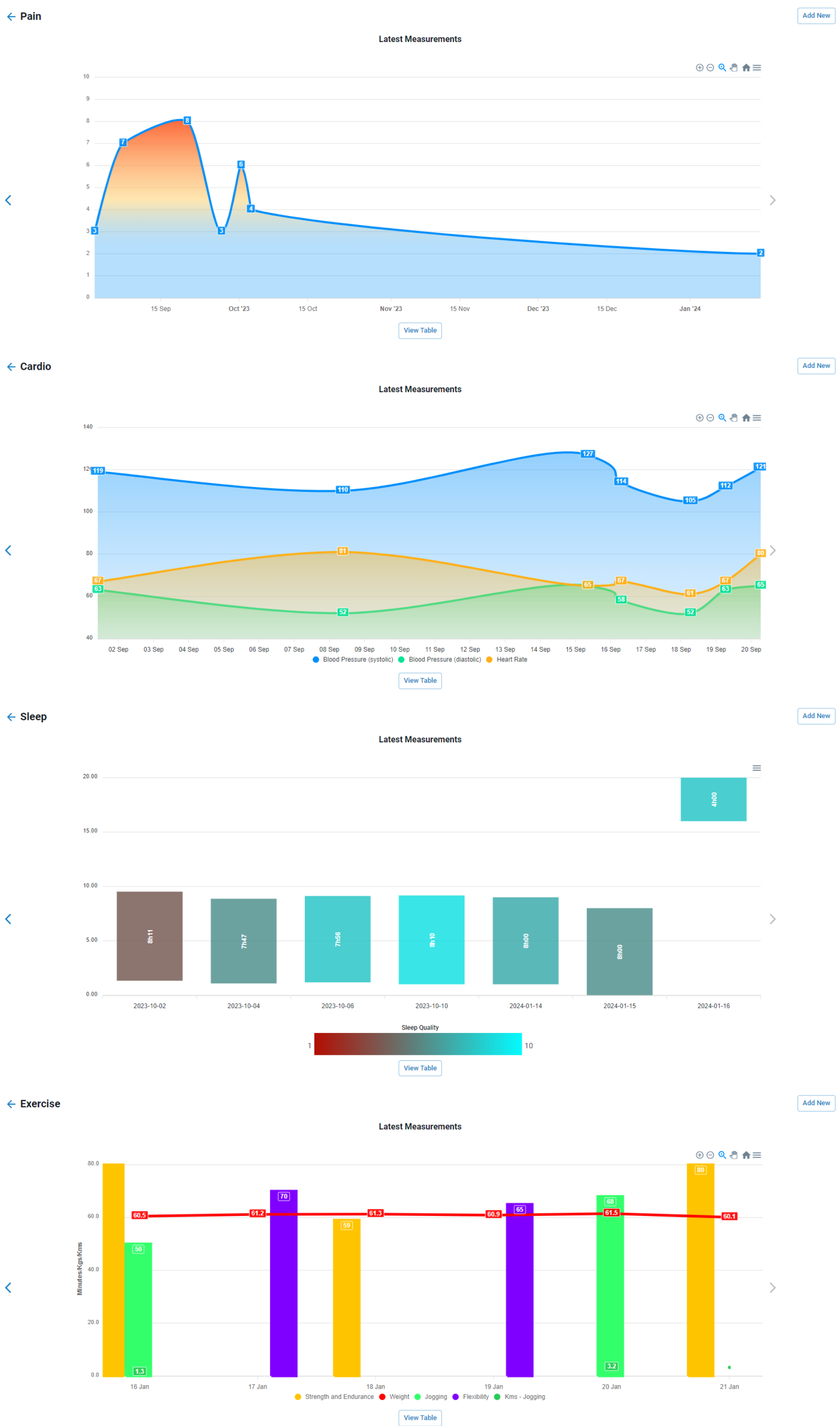

Figure 3.34: Quality of Life Graphs.

4

Conclusion

The development of these knowledge base modules for the CORE web application is driven by the overarching goal of providing comprehensive support for patients and healthcare professionals across various medical domains. These modules aim to empower patients to actively engage in their healthcare journey while also equipping healthcare professionals with valuable insights into patient health and well-being.

By offering a diverse range of resources, including tailored educational materials, medication intake reminders, medication management solutions and registry and monitoring of quality of life indicators, these modules strive to facilitate informed decision-making and enhance the overall quality of care across different medical specialties. Through this approach, the modules aim to foster collaboration and communication between patients and healthcare providers, ultimately improving patient outcomes and contributing to advancements in healthcare practices.

The development of the modules was closely followed by Dra. Maria Torrente of the Oncology Department of the Puerta de Hierro Majadahonda University Hospital. Her expertise and insights played a pivotal role in ensuring that the modules met the specific needs and requirements of oncology patients. Through her guidance, the modules were tailored to address key challenges faced by patients undergoing treatments, including the management of medication regimens and the need for reliable information and support resources.

Throughout the development of the modules, a test base comprising 20 individuals, including both patients and healthcare professionals, was involved in testing this prototype. This inclusive approach ensured that the modules were rigorously evaluated from various perspectives, thereby enhancing their effectiveness and relevance. The feedback received from this diverse group was overwhelmingly positive, with participants expressing high levels of satisfaction with the overall functionality and usability of the modules.

Additionally, CORE was selected to be used as a base product in the following projects:

- "Programa multidisciplinar basado en inteligencia artificial para la rehabilitación cardio-oncológica de los pacientes con cáncer de pulmón" - Ayudas Santander-UFV

- "Programa multidisciplinar basado en ejercicio físico, control de riesgo cardiovascular e inteligencia artificial en pacientes con cáncer de pulmón y mama" - Ayuda Estrategica en Salud 2024

And is being used, for exploration and testing in:

- "Efectos de una atención multidisciplinar en la calidad de vida y riesgo cardiovascular de pacientes con cáncer de mama precoz" - Beca SEOM 2023

4.1 Future Work

After successfully implementing the recommendation features in the Q&A section, it is necessary to apply these steps to the quiz component to ensure consistency and a seamless user experience throughout the platform. By extending the same level of interactivity and moderation to the quiz, users can actively participate, provide feedback and acquire knowledge related to the quiz's topics.

One limitation of the medication management module is its reliance on Google accounts for medicine reminders. This implies that users who do not possess a Google account will not fully benefit from this feature, missing out on the convenience and assistance provided by the reminder functionality, potentially impacting their medication adherence and overall treatment experience. Expanding the accessibility of reminder features to accommodate users without Google accounts could enhance the inclusivity and effectiveness of the module, ensuring that all patients receive comprehensive support in managing their medication regimens. To address this issue, initial efforts were undertaken to integrate the reminder system with Google Firebase Cloud Messaging. This integration allows for the delivery of notifications directly to users' browsers, circumventing the need for a Google account to receive reminders. By leveraging this approach, the module aims to enhance accessibility and ensure that all users can benefit from timely medication reminders.

Another aspect for improvement of the medication management module lies in enhancing the calendar display functionality to allow for more dynamic views of medication adherence, rather than being limited to fixed timeframes such as the last 7 or 28 days. This enhancement would provide users with greater flexibility in monitoring their medication intake over customizable time periods, thereby offering a more comprehensive overview of their adherence behavior.

For the Quality of Life module, motivational elements can be integrated to enhance user engagement and adherence. This includes incorporating rewards or motivational badges as incentives based on progress towards goals. Users could receive recognition or celebrate milestones when they achieve certain amounts of registrations or demonstrate consistent engagement with the module. These motivational features aim to encourage users to actively participate in tracking their quality of life indicators and promote a sense of accomplishment as they work towards their health and wellness goals.

Additionally, integrating with popular quality of life monitoring solutions such as health tracking bands and smart watches can provide valuable data sources for the recorded parameters. These platforms offer comprehensive monitoring capabilities and are trusted by users for tracking various aspects of their well-being, including physical activity, sleep patterns, and vital signs. Leveraging these platforms as data sources can enhance the functionality of the Quality of Life module by providing users with a seamless experience and access to a wider range of health metrics.

Bibliography

[1] *5 Ways a Customer Interaction Can Improve Your Business*. https://www.zendesk.com.br/blog/5-ways-customer-interaction-can-improve-business/. Website: Zendesk Blog. 2023 (cit. on p. 21).

[2] A-Team Global. *Why is React so Popular?* https://a-team.global/blog/why-is-react-so-popular/. In this article, A-Team Global explains in detail why React is so popular. We will examine more closely what makes creating applications with React such a fantastic experience. 2023 (cit. on p. 42).

[3] R. Agrawal, T. Imieliński, and A. Swami. "Mining association rules between sets of items in large databases". In: *Proceedings of the 1993 ACM SIGMOD international conference on Management of data*. 1993, pp. 207–216 (cit. on pp. 24, 25, 28).

[4] R. Agrawal, R. Srikant, et al. "Fast algorithms for mining association rules". In: *Proc. 20th int. conf. very large data bases, VLDB*. Vol. 1215. Santiago, Chile. 1994, pp. 487–499 (cit. on p. 29).

[5] T. B. Alakus and I. Turkoglu. "Comparison of deep learning approaches to predict COVID-19 infection". In: *Chaos, Solitons & Fractals* 140 (2020), p. 110120 (cit. on p. 18).

[6] *Answer the Right Questions: FAQ Software*. https://www.zendesk.com.br/help-center/knowledge-base-software/faq-software/. Website: Zendesk Help Center. 2023 (cit. on p. 21).

[7] J. Arents and M. Greitans. "Smart industrial robot control trends, challenges and opportunities within manufacturing". In: *Applied Sciences* 12.2 (2022), p. 937 (cit. on p. 19).

[8] L. Arnold et al. "An introduction to deep learning". In: *European Symposium on Artificial Neural Networks (ESANN)*. 2011 (cit. on p. 17).

[9] G. Aruni, G. Amit, and P. Dasgupta. "New surgical robots on the horizon and the potential role of artificial intelligence". In: *Investigative and clinical urology* 59.4 (2018), pp. 221–222 (cit. on pp. 9, 13).

[10] N. Bhirud et al. "A literature review on chatbots in healthcare domain". In: *International journal of scientific & technology research* 8.7 (2019), pp. 225–231 (cit. on p. 7).

[11] M. T. Brown et al. "Medication adherence: truth and consequences". In: *The American journal of the medical sciences* 351.4 (2016), pp. 387–399 (cit. on p. 2).

[12] R. Y. Choi et al. "Introduction to machine learning, neural networks, and deep learning". In: *Translational vision science & technology* 9.2 (2020), pp. 14–14 (cit. on p. 17).

[13] M. S. Cousins, L. M. Shickle, and J. A. Bander. "An introduction to predictive modeling for disease management risk stratification". In: *Disease Management* 5.3 (2002), pp. 157–167 (cit. on p. 18).

[14] A. Das et al. "Towards intelligent extended reality in stroke rehabilitation: Application of machine learning and artificial intelligence in rehabilitation". In: *Augmenting Neurological Disorder Prediction and Rehabilitation Using Artificial Intelligence*. Elsevier, 2022, pp. 309–329 (cit. on p. 14).

[15] T. Davenport and R. Kalakota. "The potential for artificial intelligence in healthcare". In: *Future healthcare journal* 6.2 (2019), p. 94 (cit. on pp. 9, 15).

[16] D. Demner-Fushman, W. W. Chapman, and C. J. McDonald. "What can natural language processing do for clinical decision support?" In: *Journal of biomedical informatics* 42.5 (2009), pp. 760–772 (cit. on pp. 9, 10).

[17] A. M. Dinani, K. V. Kowdley, and M. Noureddin. "Application of artificial intelligence for diagnosis and risk stratification in NAFLD and NASH: the state of the art". In: *Hepatology* 74.4 (2021), pp. 2233–2240 (cit. on p. 9).

[18] *Discourse Top Features*. `https://www.discourse.org/features`. Website: Discourse. 2023 (cit. on p. 22).

[19] R. M. Epstein and R. L. Street. *The values and value of patient-centered care*. 2011 (cit. on p. 1).

[20] A. Esteva et al. "Deep learning-enabled medical computer vision". In: *NPJ digital medicine* 4.1 (2021), p. 5 (cit. on pp. 9, 11).

[21] M Eswaran and M. R. Bahubalendruni. "Challenges and opportunities on AR/VR technologies for manufacturing systems in the context of industry 4.0: A state of the art review". In: *Journal of Manufacturing Systems* 65 (2022), pp. 260–278 (cit. on p. 19).

[22] D. Feldman-Stewart et al. "What questions do patients with curable prostate cancer want answered?" In: *Medical Decision Making* 20.1 (2000), pp. 7–19 (cit. on p. 1).

[23] B. J. Fogg. "Persuasive technology: using computers to change what we think and do". In: *Ubiquity* 2002.December (2002), p. 2 (cit. on p. 35).

[24] C. Friedman et al. "Automated encoding of clinical documents based on natural language processing". In: *Journal of the American Medical Informatics Association* 11.5 (2004), pp. 392–402 (cit. on pp. 9, 10).

[25] G. Frija et al. "How to improve access to medical imaging in low-and middle-income countries?" In: *EClinicalMedicine* 38 (2021) (cit. on p. 15).

[26] M. J. V. Gil et al. "Augmented reality in physical therapy: systematic review and meta-analysis". In: *JMIR Serious Games* 9.4 (2021), e30985 (cit. on p. 14).

[27] D. Grewal et al. *The future of technology and marketing: A multidisciplinary perspective*. 2020 (cit. on p. 18).

[28] J. Hamari, J. Koivisto, and T. Pakkanen. "Do persuasive technologies persuade?-a review of empirical studies". In: *Persuasive Technology: 9th International Conference, PERSUASIVE 2014, Padua, Italy, May 21-23, 2014. Proceedings 9*. Springer. 2014, pp. 118–136 (cit. on pp. 36, 37).

[29] P. Hamet and J. Tremblay. "Artificial intelligence in medicine". In: *Metabolism* 69 (2017), S36–S40 (cit. on p. 9).

[30] J. Han, J. Pei, and Y. Yin. "Mining frequent patterns without candidate generation". In: *ACM sigmod record* 29.2 (2000), pp. 1–12 (cit. on pp. 31, 32).

[31] E. A. Holm et al. "Overview: Computer vision and machine learning for microstructural characterization and analysis". In: *Metallurgical and Materials Transactions A* 51 (2020), pp. 5985–5999 (cit. on p. 20).

[32] A. J. Holmgren and J. Adler-Milstein. "Health information exchange in US hospitals: the current landscape and a path to improved information sharing". In: *Journal of hospital medicine* 12.3 (2017), pp. 193–198 (cit. on p. 1).

[33] T. Irizarry, A. DeVito Dabbs, and C. R. Curran. "Patient portals and patient engagement: a state of the science review". In: *Journal of medical Internet research* 17.6 (2015), e148 (cit. on p. 5).

[34] M. N. Islam et al. "A systematic review on the use of AI and ML for fighting the COVID-19 pandemic". In: *IEEE Transactions on Artificial Intelligence* 1.3 (2020), pp. 258–270 (cit. on p. 12).

[35] Y. Jeon, E. J. Park, and S. H. Baik. "Robotic surgery for rectal cancer and cost-effectiveness". In: *Journal of minimally invasive surgery* 22.4 (2019), p. 139 (cit. on p. 14).

[36] X. Jiang et al. "Towards an artificial intelligence framework for data-driven prediction of coronavirus clinical severity". In: *Computers, Materials & Continua* 63.1 (2020), pp. 537–551 (cit. on pp. 9, 12).

[37] D. C. Kaelber and D. W. Bates. "Health information exchange and patient safety". In: *Journal of biomedical informatics* 40.6 (2007), S40–S45 (cit. on p. 1).

[38] N. S. Khan, A. Abid, and K. Abid. “A novel natural language processing (NLP)–based machine translation model for English to Pakistan sign language translation”. In: *Cognitive Computation* 12 (2020), pp. 748–765 (cit. on p. 16).

[39] Z. Khorgami et al. “The cost of robotics: an analysis of the added costs of robotic-assisted versus laparoscopic surgery using the National Inpatient Sample”. In: *Surgical endoscopy* 33 (2019), pp. 2217–2221 (cit. on p. 14).

[40] N. Kim et al. “Probing what different NLP tasks teach machines about function word comprehension”. In: *arXiv preprint arXiv:1904.11544* (2019) (cit. on p. 16).

[41] S. Kotsiantis and D. Kanellopoulos. “Association rules mining: A recent overview”. In: *GESTS International Transactions on Computer Science and Engineering* 32.1 (2006), pp. 71–82 (cit. on p. 24).

[42] T. A. Kumbhare and S. V. Chobe. “An overview of association rule mining algorithms”. In: *International Journal of Computer Science and Information Technologies* 5.1 (2014), pp. 927–930 (cit. on p. 29).

[43] B. D. Lam et al. “A Patient-Centered Approach to Writing Ambulatory Visit Notes in the Cures Act Era”. In: *Applied Clinical Informatics* 14.01 (2023), pp. 199–204 (cit. on p. 1).

[44] A. Lin et al. “Artificial intelligence in cardiovascular imaging for risk stratification in coronary artery disease”. In: *Radiology: Cardiothoracic Imaging* 3.1 (2021), e200512 (cit. on p. 12).

[45] D. C. Marinescu. *Cloud computing: theory and practice*. Morgan Kaufmann, 2022 (cit. on pp. 38, 39).

[46] B. Martínez-Pérez, I. De La Torre-Díez, and M. López-Coronado. “Mobile health applications for the most prevalent conditions by the World Health Organization: review and analysis”. In: *Journal of medical Internet research* 15.6 (2013), e120 (cit. on p. 7).

[47] J. Monge et al. “AI-Based Smart Sensing and AR for Gait Rehabilitation Assessment”. In: *Information* 14.7 (2023), p. 355 (cit. on p. 14).

[48] A. Morar et al. “A comprehensive survey of indoor localization methods based on computer vision”. In: *Sensors* 20.9 (2020), p. 2641 (cit. on p. 20).

[49] H. Muaddi et al. “Clinical outcomes of robotic surgery compared to conventional surgical approaches (laparoscopic or open): a systematic overview of reviews”. In: *Annals of surgery* 273.3 (2021), pp. 467–473 (cit. on p. 13).

[50] P. M. Nadkarni, L. Ohno-Machado, and W. W. Chapman. “Natural language processing: an introduction”. In: *Journal of the American Medical Informatics Association* 18.5 (2011), pp. 544–551 (cit. on p. 16).

[51] A. Palanica et al. “Physicians’ perceptions of chatbots in health care: cross-sectional web-based survey”. In: *Journal of medical Internet research* 21.4 (2019), e12887 (cit. on pp. 6, 7).

[52] P. Phupattanasilp and S.-R. Tong. “Augmented reality in the integrative internet of things (AR-IoT): Application for precision farming”. In: *Sustainability* 11.9 (2019), p. 2658 (cit. on p. 19).

[53] I. M. Pires et al. “A research on the classification and applicability of the mobile health applications”. In: *Journal of personalized medicine* 10.1 (2020), p. 11 (cit. on p. 8).

[54] ProviderTech. *7 Must-Have Features of Patient Engagement Platforms*. https://www.providertech.com/patient-engagement-platforms-7-must-have-features/. 2023 (cit. on p. 5).

[55] A. Rajak and M. K. Gupta. “Association rule mining: applications in various areas”. In: *Proceedings of international conference on data management, Ghaziabad, India*. 2008, pp. 3–7 (cit. on p. 24).

[56] A. Reynolds. “Patient-centered care”. In: *Radiologic technology* 81.2 (2009), pp. 133–147 (cit. on p. 1).

[57] S. J. Rolnick et al. “Patient characteristics associated with medication adherence”. In: *Clinical medicine & research* 11.2 (2013), pp. 54–65 (cit. on p. 2).

[58] G. K. Savova et al. “Use of natural language processing to extract clinical cancer phenotypes from electronic medical records”. In: *Cancer research* 79.21 (2019), pp. 5463–5470 (cit. on p. 16).

[59] K. Schouten, F. Frasincar, and F. de Jong. “Ontology-Enhanced Aspect-Based Sentiment Analysis”. In: 2017-06, pp. 302–320. ISBN: 978-3-319-60130-4. DOI: 10.1007/978-3-319-60131-1_17 (cit. on p. 17).

[60] J. S. Shapiro et al. “Using health information exchange to improve public health”. In: *American journal of public health* 101.4 (2011), pp. 616–623 (cit. on p. 1).

[61] F. Shi et al. “Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19”. In: *IEEE reviews in biomedical engineering* 14 (2020), pp. 4–15 (cit. on p. 15).

[62] A. Singh, S. Sengupta, and V. Lakshminarayanan. “Explainable deep learning models in medical image analysis”. In: *Journal of imaging* 6.6 (2020), p. 52 (cit. on p. 18).

[63] M. Sostero. “Automation and robots in services: review of data and taxonomy”. In: (2020) (cit. on p. 19).

[64] P. A. C. Sousa and J. P. Pimentão. *Knowledge Discovery: Association Rules*. FCT NOVA Knowledge Discovery Class T_12 Association Rules. 2022 (cit. on p. 25).

[65] A. Sunyaev and A. Sunyaev. “Cloud computing”. In: *Internet Computing: Principles of Distributed Systems and Emerging Internet-Based Technologies* (2020), pp. 195–236 (cit. on p. 38).

[66] *Thriving Communities Made Easy*. https://www.discourse.org/. Website: Discourse. 2023 (cit. on p. 21).

[67] Y. Tian, G. Pan, and M.-S. Alouini. “Applying deep-learning-based computer vision to wireless communications: Methodologies, opportunities, and challenges”. In: *IEEE Open Journal of the Communications Society* 2 (2020), pp. 132–143 (cit. on p. 20).

[68] M. Torrente et al. “An artificial intelligence-based tool for data analysis and prognosis in cancer patients: Results from the Clarify study”. In: *Cancers* 14.16 (2022), p. 4041 (cit. on p. 9).

[69] *Using the Knowledge Base in Zendesk*. https://support.zendesk.com/hc/en-us/sections/4405303738906-Using-the-knowledge-base-in-Help-Center. Website: Zendesk Support. 2023 (cit. on p. 21).

[70] G. Velikova et al. “Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial”. In: *Journal of clinical oncology* 22.4 (2004), pp. 714–724 (cit. on p. 2).

[71] M. Venkatesan et al. “Virtual and augmented reality for biomedical applications”. In: *Cell reports medicine* 2.7 (2021), p. 100348 (cit. on p. 9).

[72] R. Wootton. “Telemedicine”. In: *Bmj* 323.7312 (2001), pp. 557–560 (cit. on p. 7).

[73] R. Wootton, J. Craig, and V. Patterson. *Introduction to telemedicine*. CRC Press, 2017 (cit. on p. 7).

[74] Z. Xu. “I don’t understand you but I trust you: Using computer-aided text analysis to examine medical terminology use and engagement of vaccine online articles”. In: *Journal of Communication in Healthcare* 14.1 (2021), pp. 61–67 (cit. on p. 2).

[75] K.-H. Yu, A. L. Beam, and I. S. Kohane. “Artificial intelligence in healthcare”. In: *Nature biomedical engineering* 2.10 (2018), pp. 719–731 (cit. on p. 9).

[76] Q. Zhao and S. S. Bhowmick. “Association rule mining: A survey”. In: *Nanyang Technological University, Singapore* 135 (2003) (cit. on pp. 29–31, 33).

2024 Knowledge Base for Oncology Patient Care Nuno Nunes

NOVA SCHOOL OF SCIENCE & TECHNOLOGY