SaúdeLink: AI-Enhanced Health Data Integration to Support Healthcare Providers

Maurício de M. dos Santos[1], Evilasio C. Junior[1], Rossana M. C. Andrade[1], Pedro A. M. Oliveira[2]
[1]*Group of Networks, Software Engineering and Systems (GREat), Federal University of Ceará (UFC), Sobral and Fortaleza, Brazil*
[2]*Lab. of Innovation and Scientific Computing (LJCC), Federal Institute of Maranhão (IFMA), Pedreiras, Brazil*
{*mauricioms@alu, evilasiojunior, rossana*}*@ufc.br, pedro.oliveira@ifma.edu.br*

Keywords: Digital Health, Wearable Devices, Internet of Things (IoT), Large Language Models (LLMs), Contextual Analysis, Technical Action Research

Abstract: This paper presents the development of *SaúdeLink*, an intelligent platform designed to integrate physiological and behavioral data from wearable devices with large language models (LLMs), specifically leveraging the Gemini language model (gemini-2.5-flash), aiming to support contextual analysis and decision-making in digital health. The proposal explores the potential of data collected from *wearable* technologies, such as heart rate, sleep, and physical activity levels, to generate personalized and context-aware insights about patients' health and well-being. The development followed the *Technical Action Research* (TRA) methodology, which enables iterative cycles of design, technical implementation, and empirical refinement. The proposed architecture combines vector-based representations (*embeddings*), semantic search, and prompt engineering, allowing clinical information to be interpreted adaptively according to each patient's history and profile. For the initial evaluation, two synthetic profiles, *Sedentary* and *Athlete*, were created to represent contrasting lifestyle patterns. Results demonstrated that the system can distinguish between physiological profiles and generate recommendations tailored to each context. Moreover, the platform was designed to support continuous expansion, allowing daily data ingestion and the integration of external documents, such as medical guidelines and scientific publications, to enrich the model's contextual understanding. The obtained results indicate that *SaúdeLink* provides a solid foundation for future applications involving longitudinal patient monitoring, validation with healthcare professionals, and integration with existing clinical systems. The platform highlights the potential of combining Internet of Things (IoT) data and generative artificial intelligence to enhance clinical screening, monitoring, and decision support.

1 INTRODUCTION

The advancement of digital health technologies and the widespread adoption of wearable devices have transformed the way physiological and behavioral parameters are continuously monitored. Metrics such as heart rate, sleep stages, caloric expenditure, and physical activity levels are now routinely captured by commercial devices, offering unprecedented opportunities for longitudinal health tracking (Li et al., 2017; Mary, 2020). These technologies have become valuable tools for promoting personalized care and improving quality of life (Kang and et al., 2022). At the same time, the global phenomenon of population aging and the increasing prevalence of chronic diseases have reinforced the need for preventive and continuous monitoring strategies in healthcare systems (Organization, 2018). In this context, digital health solutions emerge as essential allies for both patients and professionals.

Despite the availability of data generated by wearables, their practical use in clinical practice remains limited. Healthcare professionals often face challenges in accessing, integrating, and interpreting information due to interoperability issues, device heterogeneity, and the fragmented nature of data streams (Canali et al., 2022). Consequently, many consultations still rely primarily on subjective reports, without fully leveraging the potential of objective sensor-based measurements (Contribuciones, 2024). This gap underscores the importance of developing platforms that bridge raw sensor data with clinically relevant insights.

Recent research has sought to address this challenge by proposing frameworks that combine Internet of Health Things (IoHT), machine learning, and semantic analysis. For instance, (Oliveira et al., 2025) introduced the dataset presented in *Healful Dataset*, which integrates wearable data with self-reported quality of life assessments, enabling the correlation between objective health indicators and subjective well-being measures. Complementarily, the dataset presented in Healful platform was proposed as an IoHT-based solution to infer quality of life using machine learning models applied to data collected from smartphones and wearables (Oliveira and Andrade, 2024). These initiatives demonstrate the feasibility of continuous monitoring and contextual analysis in real-world scenarios. In addition, the process put forward by Costa Junior's doctoral work on the MOTION process highlights methodological advances in the development of adaptive IoHT applications based on movement patterns, reinforcing the maturity of this research field (Junior, 2023).

Building on these foundations, SaúdeLink emerges as an intelligent platform that integrates physiological and behavioral data from wearables with large language models (LLMs). The Gemini language model (gemini-2.5-flash) was used, enabling contextualized and adaptive interpretation of patient information. By combining vector-based representations, semantic search, and prompt engineering, SaúdeLink provides healthcare professionals with personalized insights that go beyond raw metrics. Initially, the platform operates with the dataset presented in *Healful: Wearable Data vs Self-Reported QoL dataset* (Oliveira et al., 2025), leveraging parameters such as height, weight, heart rate, steps, calories, and sleep stages (Awake, Light, Deep, and REM). The results of preliminary evaluations demonstrate the system's ability to distinguish lifestyle profiles and generate tailored recommendations.

SaúdeLink offers a structured, comprehensible, and extensible framework for clinical support that can be used for patient triage, longitudinal monitoring, and integration with electronic health records. This contribution underscores the promise of combining wearable technologies and language models for more effective, preventive, and personalized healthcare.

2 DATASET

The *SaúdeLink* platform is powered by data derived from the dataset presented in *Healful Dataset*. This comprehensive resource integrates physiological metrics from wearable devices with self-reported quality of life assessments. Developed by (Oliveira et al., 2025), the dataset aims to support research in digital health, wearable-based monitoring, and computational modeling of behavioral and physiological patterns.

The dataset was built from a data collection protocol involving participants equipped with commercially available wearable devices capable of tracking sleep stages, heart rate, steps, and caloric expenditure. Alongside these objective measurements, participants also provided daily self-reported information regarding lifestyle habits and perceived well-being. This dual structure enables the dataset to be used in studies correlating sensor-based data with subjective assessments, evaluating digital biomarkers, and developing intelligent systems for health monitoring.

Although the dataset presented in *Healful Dataset* contains multiple components, this work focuses specifically on the file `20230120-data-collector-dailyRegister.csv`, which aggregates daily health and behavior metrics for each participant. This subset includes the fields most relevant to the present study, such as height, weight, steps, caloric expenditure, heart rate, and sleep stages (Awake, Light, Deep, and REM). These metrics form the foundation for the semantic representations and contextual analyses performed by the *SaúdeLink* platform.

By relying on this structured, high-quality dataset, *SaúdeLink* can simulate realistic patient scenarios and evaluate its ability to interpret physiological patterns, generate contextualized summaries, and provide differentiated recommendations based on distinct health profiles.

3 RELATED WORK

The growing field of digital health has been strongly influenced by the increasing availability of wearable devices that continuously monitor physiological and behavioral parameters, such as heart rate, sleep stages, and physical activity levels. This surge in sensor-based data has driven extensive research into transforming raw measurements into actionable health insights. The integration of Artificial Intelligence (AI) and Machine Learning (ML) techniques has played a central role in this transformation, enabling the automatic extraction of meaningful patterns and the prediction of health-related outcomes from high-dimensional time-series data.

The study by (Nurmi et al., 2021) investigated the application of machine learning algorithms to health data from wearable and sensor-based systems. The review identified a wide range of techniques employed for processing data from accelerometers, gyroscopes, electrocardiograms (ECG), electroencephalograms (EEG), and glucose sensors. Classical algorithms such as Support Vector Machines (SVM), Decision Trees, and Random Forests were among the most frequently adopted. At the same time, neural networks and deep learning approaches have recently gained prominence for their superior ability to capture temporal dependencies and nonlinear relationships in physiological data. Despite these advancements, the authors emphasize persistent challenges related to generalization across populations, sensor heterogeneity, and data imbalance. These limitations hinder the translation of research findings into consistent real-world clinical use. These findings highlight the need for platforms like *SaudeLink*, designed to organize heterogeneous wearable measurements and convert them into structured, clinically relevant interpretations that support professional decision-making.

While algorithmic advancements continue to expand the analytical potential of wearable data, clinical adoption remains limited. (Baig et al., 2019) explored the barriers preventing the widespread use of wearable sensors and Internet of Things (IoT) solutions in healthcare environments. The authors identified critical obstacles, including device accuracy, limited battery life, lack of interoperability across platforms, and data privacy concerns. Moreover, they highlighted the absence of practical frameworks for integrating wearable data into Electronic Health Records (EHRs) and clinical workflows. These limitations limit healthcare professionals' ability to contextualize and interpret patient information derived from continuous monitoring devices. The insights from this study align closely with the motivation behind *SaudeLink*, which seeks to bridge the gap between raw sensor data and its clinical interpretation by providing a structured, human-centered interface for health professionals.

More recently, the research community has begun exploring the application of Large Language Models (LLMs) to enhance the interpretability and contextual understanding of wearable sensor data. The work by (Zhang et al., 2024), titled *PhysioLLM: Supporting Personalized Health Insights with Wearables and Large Language Models*, proposed a framework where LLMs were fine-tuned on multimodal physiological data to generate personalized textual feedback for users. This approach demonstrated the potential of LLMs to reason about temporal trends and provide human-understandable summaries from high-frequency wearable streams. Similarly, in the study *Health-LLM: Large Language Models for Health Prediction via Wearable Sensor Data* (Wang et al., 2024), the authors combined sensor-derived embeddings with language model reasoning to improve predictive tasks such as stress detection and sleep quality estimation. Both studies highlight the growing trend of leveraging LLMs' reasoning and generative capabilities to improve interpretability in digital health, marking a significant step beyond purely numerical ML methods.

The reviewed literature demonstrates considerable progress in data acquisition, modeling, and interpretation of wearable health information. Nevertheless, persistent gaps remain in how these insights are delivered to and utilized by healthcare professionals. While prior research has focused either on algorithmic innovation (as in (Nurmi et al., 2021) and (Wang et al., 2024)) or on technical and adoption barriers (as in (Baig et al., 2019)), there is a lack of intermediary systems that translate processed data into clinically relevant narratives. The *SaudeLink* platform addresses this gap by combining structured metric aggregation with natural language interaction. By providing a layer of contextualization between raw sensor outputs and clinical interpretation, *SaudeLink* aims to enhance the accessibility and practical value of wearable-derived health data in real-world healthcare environments. Furthermore, although it currently operates on an existing dataset (the dataset presented in (Healful: Wearable Data vs Self-Reported QoL, 2024)), its architecture and methodology are designed to be extensible to live, continuously updated data sources in the future.

Overall, integrating LLMs with wearable health data presents a promising research avenue. However, as the reviewed works and current limitations suggest, true clinical utility will depend not only on algorithmic sophistication but also on the ability to bridge communication gaps between patients, data, and professionals.

4 METHODOLOGY

For our research, we used Technical Action Research (TAR). This methodological approach combines scientific investigation with iterative cycles of technical action and critical reflection to develop and evaluate solutions to real-world problems (Wieringa, 2014). This methodology is widely used in Software Engineering and Information Systems research, as it enables the technological development process to be conducted empirically, collaboratively, and adaptively, as illustrated in Figure 1.

In accordance with the three-level structure of Technical Action Research (TAR) proposed by Wieringa (Wieringa, 2014), the development of the

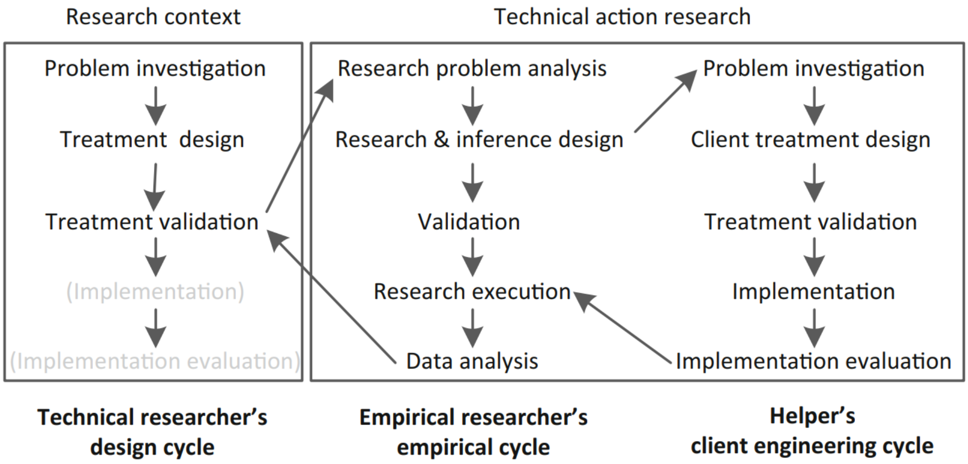

Figure 1: Methodology Technical Action Research (TAR) (Wieringa, 2014).

SaúdeLink platform was organized into iterative cycles distributed across the *problem investigation*, *treatment design*, *treatment validation*, and *implementation* layers illustrated in Fig. 1. Each stage of the project corresponded to one of these levels, ensuring methodological coherence between the theoretical foundations and the practical construction of the solution.

Planning (Problem Investigation and Treatment Design): This phase corresponds primarily to the upper levels of TAR, where the problem context is investigated and an initial treatment is conceptualized. Here, the central challenge was identified: the underutilization of physiological and behavioral data collected from wearable (IoT) devices in clinical practice. A literature review was conducted to examine the use of generative artificial intelligence, retrieval-augmented generation (RAG), and semantic indexing strategies in healthcare (Yang et al., 2025). In parallel, the functional and non-functional requirements of a system intended to support clinical decision-making were analyzed, forming the initial design of the treatment.

Implementation (Treatment Implementation and Local Evaluation): Aligned with the middle level of TAR, this phase involved the construction and iteration of the proposed treatment. The main modules of the platform were developed, including the data processing and indexing pipeline, the embedding-based retrieval layer, the backend responsible for orchestrating semantic queries, and the frontend interface for interaction with healthcare professionals. Development proceeded in iterative cycles, with continuous implementation, integration, and internal evaluation. Each cycle allowed design assumptions to be tested locally, refining performance, usability, and the consistency of the system's generated responses, with feedback from the project supervisors guiding improvements.

Observation and Reflection (Global Reflection and Learning): This final phase corresponds to the reflective layer of TAR, in which the outcomes of the implemented cycles are examined and insights are consolidated. A critical analysis of previous iterations' results identified opportunities to improve data flow, the accuracy and contextualization of generated responses, and the projected user experience for healthcare professionals. This stage also enabled refinement of technical decisions, such as adjustments to data aggregation strategies, prompt engineering employed in interactions with the language model, and optimization of the vector retrieval process. The reflective insights gathered in this stage informed the planning of subsequent iterations and delineated directions for future system expansion.

The application of TAR enabled the platform's development to evolve in a manner guided by empirical evidence and grounded in real clinical needs, even before conducting validation with end users.

5 SAÚDELINK

The development of the *SaúdeLink* platform was structured into three main stages: (i) selection and preprocessing of data from the dataset (Oliveira et al., 2025), (ii) transformation of the data into textual representations and embedding generation for subsequent indexing, and (iii) development of the *SaúdeLink* platform itself, composed of a backend for data management and querying, and a frontend designed for visualization and interaction with healthcare professionals, Figure 2 illustrates the data flow across the different components of the process.

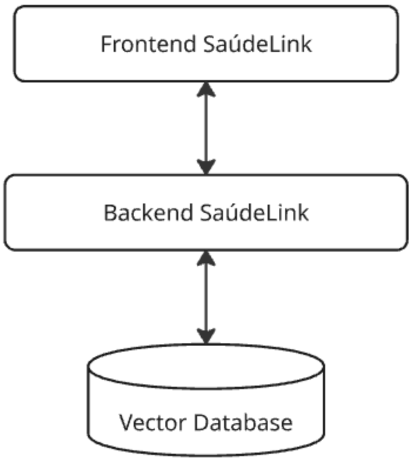

Figure 2: SaúdeLink data flow.

5.1 Data Selection and Preprocessing

The Healful dataset contains a wide range of physiological and psychological metrics collected from wearable devices, including information about sleep, physical activity, heart rate, weight, height, and calories. For this study, the selected metrics were calories, heart rate, steps, weight, height, and sleep stages (awake, light, deep, and REM). This choice aimed to balance data representativeness with analytical feasibility, focusing on parameters directly related to sleep quality, physical activity level, and overall quality of life.

The preprocessing and data analysis were carried out in the Google Colab[1] environment, leveraging its computational resources to handle the volume and complexity of the information. In this environment, the raw data (initially stored in CSV format) were systematically organized and processed. First, the columns corresponding to the selected metrics were extracted and grouped into seven-day temporal windows, allowing aggregated analysis (Barnett et al., 2020) and reducing processing load. Such weekly aggregation is widely adopted in time-series health studies, as it preserves behavioral trends and facilitates clinical interpretation.

Subsequently, descriptive statistical metrics(mean, median, maximum, minimum, and mode) were calculated for each parameter and time window. This transformation served two primary purposes: (1) to provide a statistical summary that facilitates contextual analysis, and (2) to reduce the cognitive load on the language model (LLM) during data interpretation. Although these models can handle numerical information, they tend to exhibit limitations in direct calculations and quantitative comparisons; therefore, including pre-calculated metrics enhances both the accuracy and interpretability of results (Devlin et al., 2019).

[1] https://colab.research.google.com/

5.2 Data-to-Text Conversion and Embedding Generation

With the data already consolidated and processed, a stage was developed to convert numerical records into structured text, enabling them to be used as input for the embedding generation process. This approach allows semantic indexing and subsequent contextual information retrieval through textual queries, a well-established technique in systems based on Retrieval-Augmented Generation (RAG) (Lewis et al., 2020).

Each generated textual document contains three main sections that together describe the physiological and behavioral data of each patient over a given period:

- **Period Information**: Includes the patient identifier, start and end dates of data collection, as well as static attributes such as height and weight.
- **Weekly Statistics**: Presents descriptive summaries (mean, median, minimum, maximum, and mode) for the monitored parameters: heart rate, steps, calories, and sleep stages (light, deep, REM, and awake). These metrics provide a quantitative overview of the patient's weekly behavior.
- **Daily Records**: Contain the raw daily values for each variable, allowing finer-grained analyses such as the correlation between physical activity levels and sleep quality over the same period.

This structured format ensured that each document captured both aggregated and temporal aspects of the patient's data, enabling the embedding generation process to preserve contextual nuances essential for accurate interpretation.

After this textual transformation, the documents were submitted to the embedding generation process within the Colab environment, where the vector database was also implemented using ChromaDB. This tool was chosen because it supports vector indexing with metadata, such as patient ID and time range, and allows filtered searches based on these attributes, improving query efficiency and precision. The result was an organized, semantically enriched vector base that served as the foundation for the information retrieval mechanism used by the backend.

5.3 Backend Development

The platform backend was developed to integrate with and efficiently provide access to the vector database generated in the previous stage. This database contains semantic representations of patient data in an embedding format, enabling contextual, semantically relevant searches via textual queries.

Development was carried out in a local environment, using Python as the primary programming language and the FastAPI framework to build web services. This combination was chosen for its high performance, ease of creating asynchronous APIs, and excellent compatibility with modern AI libraries. The backend architecture was designed in a modular way, allowing independent updates and expansions of the data repository. Additionally, the adopted structure enables the integration of external services that periodically process raw data and update the vector representations, ensuring continuous incorporation of new information into the system.

Since the dataset is anonymized, each patient identifier (ID) was associated with a randomly generated name, forming a *Patient* object that contains the ID, fictitious name, and available data periods. These data are provided to the interface modules and query services, allowing the healthcare professional to select the desired patient and time interval without exposing sensitive information.

The primary function of the backend is to process the questions submitted by healthcare professionals and return contextualized answers based on patient data. For this purpose, the system integrates with the Gemini language model (gemini-1.5-flash) via the Google GenAI library, enabling interpretative and context-aware responses grounded in the retrieved data. This process is described below and illustrated in the Figure 3.

The operational flow follows these steps:

1. **Request reception:** the server receives as parameters the patient identifier, the analysis period, and the question formulated by the healthcare professional.
2. **Search in the clinical vector database:** using metadata (ID and time interval), the system queries the ChromaDB vector database to retrieve only the documents corresponding to the specified patient and period.
3. **Retrieval of auxiliary documents:** the system then performs an additional search to retrieve reference documents that serve as background knowledge for interpreting the metrics. These documents include definitions and formulas of derived indices, such as the following example:

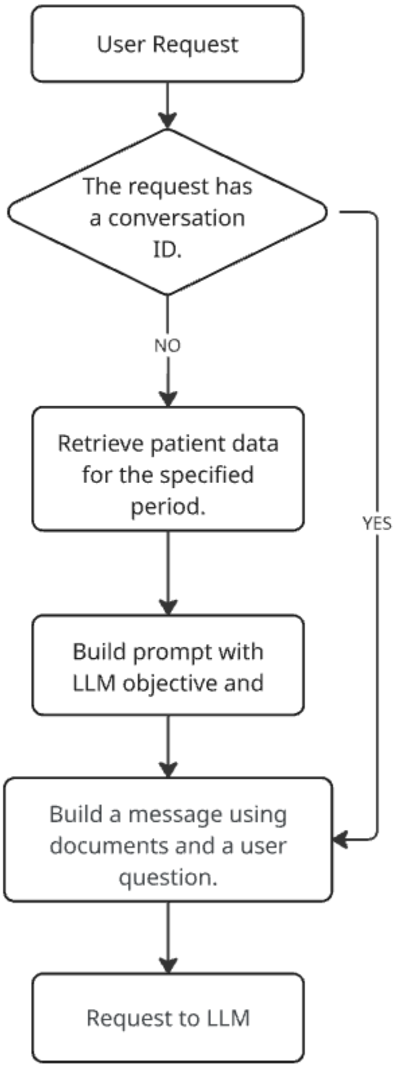

Figure 3: Backend Request Processing Flow.

The sleep architecture index can be defined as:

$$Index = \frac{DeepSleep + SleepREM}{SleepLight + DeepSleep + SleepREM}$$

Classification: $\geq 0.4 \rightarrow$ *Restorative sleep ;* < 0.4 $\rightarrow$ *Non-restorative sleep*

4. **Prompt construction:** after retrieving the necessary information, the backend builds the input prompt to be sent to the language model. This prompt is structured into three main parts: (i) definition of the system's role and objective, (ii) description of the patient data format being sent, and (iii) inclusion of the retrieved data and auxiliary documents.

To optimize performance and maintain conversational context, the server generates a unique conversation identifier (*conversation ID*) for each session.

This identifier is used in subsequent requests, allowing conversation continuity without new searches in the patient's vector database or full prompt reconstruction. Thus, only relevant auxiliary documents are retrieved in later interactions, reducing computational cost and improving the smoothness of the dialogue between the healthcare professional and the system.

5.4 Frontend Development

The platform's frontend was developed with React and TypeScript, following a component-based architecture to promote reusability and facilitate future expansion. This choice also enabled the efficient use of states and properties, which are essential in continuously interactive applications such as *SaúdeLink*. Development was conducted locally, prioritizing rapid prototyping and direct integration with the backend.

The user interface was designed with a focus on simplicity and fluid interaction. The application is structured as a chat, where the healthcare professional asks questions and receives natural-language responses based on patient data. The home screen features only a button that directs the user to the main chat environment.

In the dialogue interface, the system guides the user step by step: first, it prompts the user to select a patient (based on backend-provided data); next, it asks for the analysis period; and finally, it confirms the information before enabling the question field. The system also automatically suggests the most recent valid seven-day interval available for the patient, preventing queries for periods with missing data.

Figure 4 shows the main interface of the platform, illustrating the conversational flow between the healthcare professional and the system, including initial patient selection, period definition, and an example natural-language query.

6 RESULTS AND DISCUSSION

This section presents the results from simulating two synthetic patient profiles, **Sedentary** and **Athlete**, to evaluate the *SaúdeLink* platform's ability to distinguish between physiological and behavioral patterns and to provide personalized recommendations based on these profiles. These profiles represent two extreme but clinically relevant lifestyle categories commonly described in the literature, allowing controlled assessment of the system's contextual reasoning capabilities.

The synthetic values were derived from ranges reported in scientific studies and international guidelines, including references cited in (World Health Organization, 2024; Foundation, 2015; American College of Sports Medicine, 2021). Each profile was defined using the same set of metrics present in the dataset: height, weight, daily steps, caloric expenditure, heart rate, and sleep stages (awake, light, deep, and REM). The **Athlete** profile reflects an individual with high cardiorespiratory fitness, characterized by a taller and leaner body composition (height 1.65–1.90 m, weight 60–90 kg), high daily activity levels (12,000–25,000 steps/day), and elevated caloric expenditure (2,500–4,000 kcal). Sleep patterns follow expected distributions for highly active individuals, with efficient nocturnal recovery: 9,000–15,000 seconds of light sleep, 5,000–8,000 seconds of deep sleep, and 6,000–8,000 seconds of REM sleep, combined with minimal nocturnal awakenings (300–1,200 seconds). Resting heart rate values range from 40–65 bpm, with peaks of 160–190 bpm during training—a typical physiological response in conditioned athletes with enhanced autonomic regulation.

In contrast, the **Sedentary** profile represents an individual with low physical activity and higher metabolic risk. Anthropometric parameters include a height of 1.55–1.75 m and a weight of 75–110 kg, consistent with those reported for sedentary populations in global health studies. Daily steps range from 1,000 to 5,000, accompanied by reduced caloric expenditure (1,400–2,000 kcal). Sleep structure reflects common patterns associated with inactivity and poorer recovery quality: high proportions of light sleep (18,000–25,000 seconds), reduced deep sleep (600–2,000 seconds), and limited REM sleep (1,200–3,600 seconds), along with prolonged awake periods (2,000–4,000 seconds). Heart rate values align with expected sedentary physiology, with resting ranges between 70–100 bpm and modest activity peaks around 120 bpm.

Collectively, these ranges allow the simulation of two clearly distinguishable behavioral and physiological contexts. They serve as a controlled foundation for evaluating how effectively *SaúdeLink* interprets wearable-derived metrics and adapts its recommendations based on individualized health profiles.

6.1 Experimental Scenario

Two patients were simulated using weekly data covering the period from **September 1st, 2025, to September 7th, 2025**. For both profiles, three automated queries were submitted to the *SaúdeLink* platform with the following questions:

1. How was the patient's sleep pattern this week?
2. Is there anything that needs attention regarding

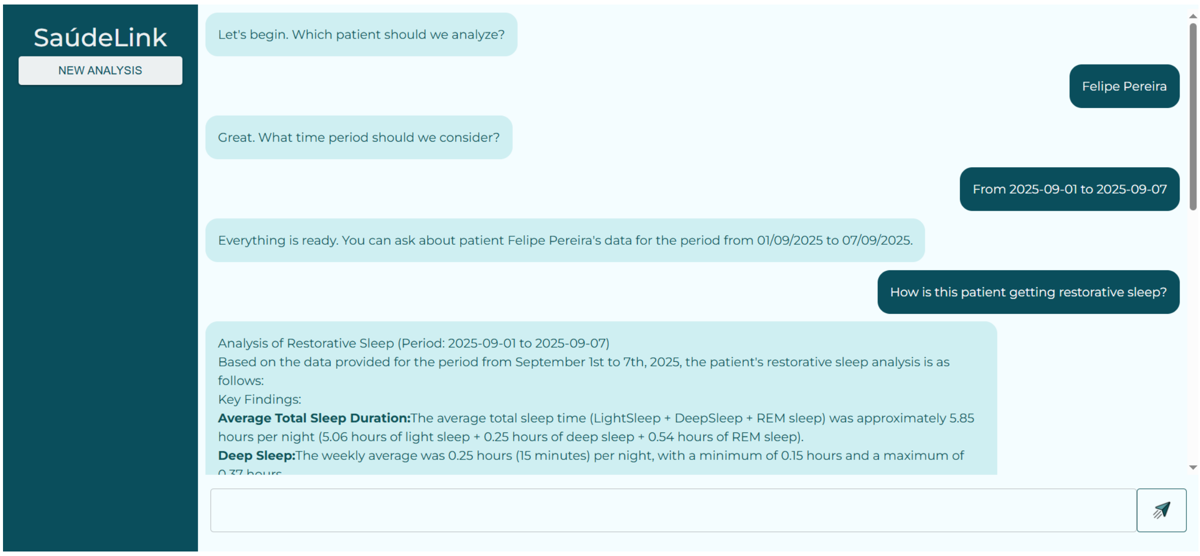

Figure 4: Chat interface of the *SaúdeLink* platform, displaying the patient and period selection stages, as well as an example of a query made by the healthcare professional.

the patient's sleep?

3. Regarding physical activity, should this patient increase or decrease activity frequency, or is the current level already considered healthy?

The responses were analyzed comparatively between the two profiles, highlighting differences in reasoning and recommendations generated by the platform.

6.2 Comparison of Generated Responses

Table 1 summarizes the main differences observed in the system responses for each profile, grouped by category of analysis.

The results show that the platform **consistently distinguished** between the two simulated profiles, correctly associating each profile's physiological behavior with the expected recommendations.

For the **Athlete** profile, the system recognized typical patterns among individuals with high physical activity levels, including a lower resting heart rate and a higher number of daily steps. The recommendations focused on *improving sleep duration*, which aligns with athletes who often have good-quality but shorter sleep due to intense training routines.

For the **Sedentary** profile, the system correctly identified *poor sleep quality* despite adequate duration and recommended an increase in physical activity and investigation of possible sleep disturbances. These recommendations are consistent with findings in the literature linking sedentary lifestyles and overweight to fragmented, inefficient sleep patterns (World Health Organization, 2024; National Heart, Lung, and Blood Institute (NHLBI), 2023).

These results demonstrate that the *SaúdeLink* platform can integrate objective data (steps, calories, sleep, heart rate) with contextually informed interpretive analyses, enabling significant differentiation among patient profiles. This feature reinforces the platform's potential as a triage and monitoring support tool in digital health contexts.

7 CHALLENGES AND LIMITATIONS

Although the *SaúdeLink* platform has demonstrated promising results, its development is accompanied by a series of challenges and limitations that are essential to acknowledge when considering future improvements. A primary concern is the system's technological dependency on external APIs, such as Google's services, and the *Gemini 2.5 Flash* language model. This reliance makes the platform susceptible to changes in provider policies, performance fluctuations, and potentially high operational costs at scale, while also requiring continuous adaptation whenever access rules or pricing models are modified.

Another critical limitation stems from the dataset used. Since the platform operates solely on the *Healful* dataset—an experimental, non-clinical collection—the analyses are based on a restricted and non-representative sample that lacks the diversity

Table 1: Comparison of *SaúdeLink* responses for the Athlete and Sedentary profiles.

Evaluated Aspect	**Athlete Profile**	**Sedentary Profile**
Average Sleep Duration	6.75 h/day — considered insufficient	7.03 h/day — considered adequate
Sleep Architecture	Index of 0.54 — restorative sleep with good proportions of deep and REM stages	Index of 0.13 — non-restorative sleep, predominance of light sleep (87%) with little deep or REM sleep
Awake Time During Sleep	0.28 h/day (4.15%) — low sleep fragmentation	0.84 h/day (11.97%) — fragmented sleep
Overall Sleep Interpretation	Qualitatively good sleep, but insufficient duration	Adequate duration but poor restorative quality
Average Daily Steps	18,617 steps/day — very high physical activity level	2,999 steps/day — low physical activity level (sedentary)
Average Heart Rate	52.7 bpm — indicative of good cardiovascular conditioning	70–100 bpm — within the expected range for sedentary individuals
System Recommendation	Maintain physical activity pattern and improve sleep duration	Increase physical activity and investigate sleep quality

found in real-world populations. Additionally, the dataset originates from a single type of wearable device. It does not include more complex physiological signals such as ECG, heart rate variability, or blood oxygenation, thereby limiting the depth and clinical sophistication of the generated insights.

The simulation of patient profiles adds further constraints. The *Athlete* and *Sedentary* profiles, while helpful for evaluating the system's interpretability, were manually constructed from literature-based ranges and lack clinical validation. These values represent conceptual approximations rather than real physiological distributions, which restricts the generalization of the results.

From a technical standpoint, the system's architecture also presents challenges. The reliance on a single language model and the RAG pipeline's sensitivity to the quality of textual representations mean the output may vary across interactions, and there is no internal mechanism to ensure semantic consistency over time. Moreover, the platform has not yet been optimized for large-scale datasets, limiting its applicability to clinical environments that demand high data throughput and integration with multiple information systems.

Finally, the data pipeline remains constrained to static CSV files and daily-level granularity, without automated continuous ingestion or finer temporal resolution. The absence of automated data cleaning mechanisms increases the risk of inconsistencies and introduces reliance on manual verification.

Together, these limitations highlight the need for further advancements before *SaúdeLink* can reach clinical maturity, particularly regarding data robustness, diversity of information sources, and the strengthening of its underlying AI infrastructure.

8 FINAL REMARKS

Despite the challenges and limitations discussed in the previous section, the development of the *SaúdeLink* platform demonstrated the potential to integrate physiological and behavioral data from wearable devices into an intelligent, language-based contextual analysis architecture. The use of vector representations (*embeddings*) and semantic search enabled clinical information to be interpreted in a context-sensitive manner, allowing the system to provide responses that are coherent with each patient's profile and history.

The results obtained with the synthetic profiles, *Sedentary* and *Athlete*, highlighted the platform's ability to distinguish between distinct physiological patterns and to adapt its recommendations based on the analyzed data. This behavior suggests that the model can understand complex relationships among variables, such as sleep, heart rate, and physical activity level, reinforcing the system's potential as a tool for clinical screening and follow-up.

One of the main differentiators of the proposed system is its ability to expand continuously. The daily insertion of new patient data into the vector database would enable automatic updates to semantic representations, enabling near real-time querying and interpretation of information. This feature supports a longitudinal monitoring model, enabling healthcare professionals to observe trends, detect anomalies, and assess patient evolution over time.

Another relevant aspect was the continuous improvement achieved through *prompt* engineering and data aggregation. Throughout development, adjustments to the structure of the instructions sent to the model and the consolidation of input data resulted in more precise, consistent, and interpretable responses. This evolution underscores the importance of iterative refinement of interaction layers in LLM-based systems.

Furthermore, although the current implementation already supports the use of auxiliary documents in the second retrieval stage, these documents are limited to simple textual descriptions that summarize or interpret the metrics employed in this study. As future work, the platform will be extended to allow healthcare institutions or clinical professionals to upload specialized documents, such as official guidelines, scientific articles, institutional protocols, or tailored patient-care procedures, thereby expanding the semantic context available to the model and enabling richer, more evidence-grounded responses aligned with each organization's knowledge base.

In addition, future developments will include validations with real healthcare professionals interacting with the system through real-world data. This evaluation may be supported by integrating the platform with mobile applications capable of collecting health metrics from wearable devices, such as Aruanã[2], enabling a continuous and automated flow of physiological data into the *SaúdeLink* ecosystem.

In summary, *SaúdeLink* is a promising prototype that integrates Internet of Things (IoT) data with generative artificial intelligence for digital health. Even in its experimental phase, the obtained results indicate that the proposed architecture provides a solid foundation for future implementations involving real-user validation, integration with existing clinical systems, and the expansion of knowledge sources for medical decision support.

REFERENCES

American College of Sports Medicine (2021). *ACSM's Guidelines for Exercise Testing and Prescription*. Wolters Kluwer, Philadelphia, PA, 11th edition.

Baig, M. M., GholamHosseini, H., Moqeem, F., et al. (2019). Current challenges and barriers to the wider adoption of wearable sensors and iot in healthcare. In *Proceedings of the 2019 International Conference on Information Systems (CONF-IRM)*, pages 1–13.

Barnett, J. et al. (2020). Weekly aggregation as a method to reduce data complexity in wearable health data. *Journal of Health Data Science*.

[2] Aruanã: https://aruana.life/

Canali, S. et al. (2022). Challenges and recommendations for wearable devices in clinical settings. *npj Digital Medicine*, 5(1):202.

Contribuciones, R. (2024). Integração de wearables e monitoramento remoto de pacientes. *Contribuciones a las Ciencias Sociales*.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. *NAACL-HLT*.

Foundation, N. S. (2015). How much sleep do we really need? https://www.sleepfoundation.org/how-sleep-works/how-much-sleep-do-we-really-need. Accessed: 2025-11-06.

Junior, E. C. (2023). *MOTION - Development process for self-adaptive internet of health things applications based on movement patterns*. Doctoral thesis, Universidade Federal do Ceará.

Kang, H.-S. and et al. (2022). Wearing the future—wearables to empower users to take health actions. *Journal of Medical Internet Research*, 24:e40123.

Lewis, P., Oguz, E., Rinott, R., and Riedel, S. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. In *NeurIPS*.

Li, X., Dunn, J., Salins, D., Zhou, G., Zhou, W., Schüssler-Fiorenza Rose, S. M., and et al. (2017). Digital health: Tracking physiomes and activity using wearable biosensors reveals useful health-related information. *PLoS Biology*, 15(1):e2001402.

Mary, X. A. (2020). Physiological parameter measurement using wearable sensors. In *Wearable Sensors and Technologies*. Elsevier.

National Heart, Lung, and Blood Institute (NHLBI) (2023). Calorie requirements and tips for healthy eating. Accessed: 2025-11-06.

Nurmi, J., Lohan, E., et al. (2021). Systematic review on machine-learning algorithms used in health data from sensors. *IEEE Access*, 9:112247–112270.

Oliveira, P. A. M. and Andrade, R. M. C. (2024). Internet of health things platform to monitor quality of life. In *Anais do Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS)*. Sociedade Brasileira de Computação.

Oliveira, P. A. M., Andrade, R. M. C., et al. (2025). Healful dataset: Integrating wearable data with self-reported quality of life assessments. In *Proceedings of the 17th International Conference on Health Informatics (HEALTHINF)*. SciTePress.

Organization, W. H. (2018). *World Health Statistics 2018: Monitoring Health for the SDGs*. WHO.

Wang, H., Chen, Y., et al. (2024). Health-llm: Large language models for health prediction via wearable sensor data. *arXiv preprint arXiv:2401.06866*.

Wieringa, R. J. (2014). Technical action research. In *Design Science Methodology for Information Systems and Software Engineering*, chapter 19, pages 269–292. Springer.

World Health Organization (2024). Physical activity. https://www.who.int/news-room/fact-sheets/detail/physical-activity. Accessed: 2025-11-06.

Yang, R., Ning, Y., Keppo, E., Liu, M., Hong, C., Bitterman, D. S., Ong, J. C. L., Ting, D. S. W., and Liu, N. (2025). Retrieval-augmented generation for generative artificial intelligence in health care. *npj Health Systems*, 2(1):2.

Zhang, W., Li, J., et al. (2024). Physiollm: Supporting personalized health insights with wearables and large language models. *IEEE Journal of Biomedical and Health Informatics*. doi: 10.1109/JBHI.2024.10913781.