ISSN 2464-4617 (print)
ISSN 2464-4625 (online)
Computer Science Research Notes - CSRN 3301
http://www.wscg.eu
WSCG 2023 Proceedings

Real-Time Visual Analytics for Remote Monitoring of Patients' Health

Maryam Boumrah[1]
Boumrah.maryam@inpt.ac.ma

Samir Garbaya[2]
Samir.garbaya@ensam.eu

Amina Radgui[1]
Radgui@inpt.ac.ma

[1] Centre d'études doctorales Télécoms et Technologies de l'Information (CEDOC-2TI), INPT, Av.Allal Alfassi,Rabat, Morocco

[2] Laboratoire END-ICAP, INSERM UMR1179, Arts et Metiers Institute of Technology, CNAM, LIFSE, HESAM University, F-75013

ABSTRACT

The recent proliferation of advanced data collection technologies for Patient Generated Health Data (PGHD) has made remote health monitoring more accessible. However, the complex nature of the big volume of medical generated data presents a significant challenge for traditional patient monitoring approaches, impeding the effective extraction of useful information. In this context, it is imperative to develop a robust and cost-effective framework that provides the scalability and deals with the heterogeneity of PGHD in real-time. Such a system could serve as a reference and would guide future research for monitoring patient undergoing a treatment at home conditions. This study presents a real-time visual analytics framework offering insightful visual representations of the multimodal big data. The proposed system was designed following the principles of User Centered Design (UCD) to ensure that it meets the needs and expectations of medical practitioners. The usability of this framework was evaluated by its application to the visualization of kinematic data of the upper limbs' movement of patients during neuromotor rehabilitation exercises.

Keywords

Visual analytics; Patient Remote Monitoring; Brain Stroke; Rehabilitation; User-Centered Design; Kafka.

1. INTRODUCTION

The exponential growth of digital technologies designed for health data capture, such as sensors and wearable devices, has enabled real-time collection of big amounts of PGHD during remote monitoring. This big data has the potential to provide valuable insights to various healthcare stakeholders, particularly clinicians. The clinicians must effectively gather and integrate data from multiple sources, synthesize a comprehensive medical interpretation based on the patient's medical record and make informed decisions [1]. However, to take full benefits of this data, it is crucial to address the challenges of its complexity, including the need for rapid processing, analysis, interpretation, and understanding. The development of new methods and tools to address these challenges is of utmost importance in this era of rapidly advancing technology.

Thus, by implementing a visual analytics solution, the massive amount of data is transformed into meaningful knowledge [2]. The use of visual analytics of medical Big Data plays an important role for taking proactive medical decision and providing better healthcare of patients. Visual analytics is the concept of combining data analysis, visual representation, and human interaction to extract insights and make decisions. Multiple areas of expertise are involved in the visual analytics process including data mining, machine learning and advanced graphic representations. Visual Analytics first emerged from the necessity of upgrading from confirmatory data analysis, which was used to represent the results, to the exploratory analysis involving the interaction with the data [3]. As defined by Keim et al., visual analytics is the integration of computerized analytics methods and visual interaction tools to gain helpful perception and decision based on big and heterogeneous data [4]. Furthermore, the advent of real-time data processing has given rise to dynamic visual analytics, which refers to the integration of knowledge discovery and interactive visual interfaces. This facilitates the analysis of data streams and enables situational awareness in real-time [5]. Over the past decade, dynamic visual analytics has garnered significant attention from research teams across multiple fields due to its potential for supporting immediate decision-making and raising awareness.

Despite the notable efforts in recent years to implement visual analytics technologies in healthcare sector, the adoption of real-time visual analytics solutions in patient health monitoring remains a challenge. The high pace and time sensitivity of medical data, along with the significant impact of any delay on clinical decisions and the safety of patients, requires that visual analytics systems must operate with minimal latency. The current state of the art falls short in offering a comprehensive solution that encompasses all essential features such as scalability, capability to handle heterogeneous data, low latency, independence from treatment systems, user-friendly visualization, cost-effectiveness, and extensiveness.
Taking into consideration the important features for real-time visual analytics in patient health monitoring, this paper proposes a unified framework based on optimized open-source technologies such as Apache Kafka and Dash [6][7]. This framework is presented to efficiently manage real-time monitoring of a considerable number of patients, each with the necessary number of sensors. The usability of the proposed approach was proved by its application for monitoring post-stroke patients during in-home rehabilitation therapy. The system was developed following the User-Centered Design (UCD) approach, with three iterations for refining the user requirements.

The research reported in this paper highlights the potential of visual analytics technologies for improving the quality of patient health monitoring and allowing appropriate clinical decision making. The proposed unified framework offers a promising solution to meet the challenges facing real-time patient health monitoring. The paper is organized as the following: Section 2 provides a comprehensive review of the state-of-the-art research in visual analytics solutions, the challenges faced in the interpretation of health data and the UCD approach for patient health monitoring. Section 3 describes the proposed framework. The development process is described in the section 4. Finally, the conclusion and the research perspectives are presented in section 5.

2. RELATED WORK

The application of advanced visual analytics, incorporating big data analysis, interactive graphical representations, and AI algorithms plays a vital role in monitoring patient's health. Despite its potential benefits, challenges arise in gaining valuable insights from patient-generated data. These challenges encompass data, human, and tool-related limitations. In particular, the quality of health data, restrictions in data access, diversity of data sources, and scarcity of emergency data pose significant difficulties [8]. Moreover, the aggregation of massive data sources into a unified platform, maintenance and storage of growing data volume, integration and interoperability of diverse data types and structures, and increased data analysis time with respect to data volume are key challenges in extracting information from big health data [9]. The complexity of health data and the lack of data standardization may also pose a challenge to the user and could result in misinterpretation [10]. On the user side, the understanding and interpretation of visual representations of health data are subject to the personal perception, leading to differences in interpretation of data and feedback on the visualized information between patients, clinicians, and data analysts [11]. Additionally, the insufficiency of data analysts and the lack of appropriate IT expertise among healthcare practitioners for processing, visualizing, and interpreting patient-generated data is a critical issue that needs to be addressed [12]. Furthermore, health stakeholders may not be prepared to adopt new systems and acquire the necessary skills for data handling [13]. The implementation of visual analytics tools for healthcare data, their development, and the creation of different methods of healthcare data representation remain significant challenges in this field.

In the domain of health monitoring and assessment, the implementation of secure and reliable information systems is of utmost importance. To achieve this, it is necessary to include in these systems appropriate management, analysis, and visualization tools [14] [15]. Feller et al. presented a visual analytics tool for pattern recognition in patient-generated data, which aimed to aid clinicians in identifying systematic and clinically meaningful patterns, and reducing perceived information overload [16]. Similarly, Vu et al. proposed a visual analytics approach for identifying informative temporal signatures in continuous cardiac monitoring alarms, using retrospective evaluation of a middleware alarm escalation software database in conjunction with visualization [17]. Dagliati et al. introduced the MOSAIC dashboard system, they used predictive modeling and longitudinal data analytics to support clinical decision-making. Their system integrated multiple-source data and uses visual and predictive analytics to enhance the management of chronic diseases, such as type two diabetes, through the successful implementation of the learning cycle of healthcare system [18].

The accuracy of health data exploration results relies heavily on the visualization tools and the used data processing interfaces. To address this, the field of human-machine interface has adopted the approach of User-Centered Design (UCD) which prioritizes the end user's requirements and needs in the development cycle of these tools. UCD follows an iterative design process, where end users are involved in the evaluation of the outcome of the system. This approach has been widely adopted in health data processing and representation, as evidenced by several studies in the literature. For instance, Hobson et al. applied UCD to develop a telehealth system named TiM, for collecting information from motor neuron disease patients which was reviewed by care providers [19]. Similarly, Griffin et al. [20] used UCD to develop an mHealth approach for colorectal cancer monitoring in elderly patients using virtual human technology. Raghau et al. developed SMARTHealth, a UCD-based mobile application for Clinical Decision Support in cardiovascular disease risk [21]. Backonja et al. used UCD to study the visualization of health data [22], while Petersen et al. developed a mobile system using UCD for collecting and analyzing data from older adults [23]. David et al. developed a virtual reality-based interactive system for upper extremity rehabilitation of post-stroke patients [24]. Caporaso et al. studied the usefulness of biomechanics and neuroscience in designing a personalized monitoring system for hand rehabilitation [25]. Osborne et al. proposed a UCD-based mobile health application for neuro rehabilitation monitoring of stroke survivors [26]. Wentink et al. used UCD to analyze the requirements of users in eRehabilitation for post-stroke patients [27].

Despite existing visual analytics solutions in patient health monitoring, there's a need for a framework that effectively addresses the limitations of the existing systems and challenges. By adopting user-centered design principles and incorporating open-source technologies, the platform could be reliable and cost-effective solution for patient health monitoring, handle diverse data types and structures, integrate multiple data sources, and prioritize user needs.

The objective of the study reported in this paper is to develop a reliable architecture for a unified real-time patient health monitoring framework using dynamic visual analytics. To design the system and determine its components, a thorough review of the literature was conducted to answer two main research questions:

1. What are the main features of an efficient visual analytics system for real-time in-home patient monitoring?
2. What are the fundamental components of a new system that integrates the identified features?

A systematic review [28] of the most relevant literature on visual analytics for real-time monitoring of stroke patient health was conducted as part of this project. The literature was analyzed to identify the most critical features of an efficient system that aligns with the objectives of this study. The main identified features were scalability, independency, extensiveness, the ability to handle heterogeneous data, real-time interaction, and patient health status prediction. Additionally, the system should be accessible on a distributed platform to users who might not have access to high performance computing devices. Thus, by adopting user-centered design principles and incorporating open-source technologies, the platform could be reliable and cost-effective solution for patient health monitoring, handle diverse data types and structures, integrate multiple data sources, and prioritize user needs.

The results of the review led to the development of a new framework based on Lambda architecture [29]. This framework is described in the following section.

3. PROPOSED FRAMEWORK

The proposed framework combines both streamed and batch data into a unified pipeline, allowing real-time visualization, analysis, and data optimization.

As illustrated in Figure 1, the proposed framework consists of five main components. The data captured by wearable sensors are ingested into the pipeline via an event hub. The use of an event hub, which is scalable and easy to manage, is crucial for incorporating data captured by multiple and different types of sensors into the system. Upon ingestion into the framework, the data are simultaneously directed to both the streaming and batch layers. The batch layer primarily includes a storage unit, which continuously accumulates data volume to provide historical patient data for reference. It is used to develop a reliable machine learning models for health data optimization through feature extraction during the preprocessing phase. In contrast, the streaming layer is responsible for delivering real-time patient data with minimal latency to the therapist dashboard. Prior to display, the streamed data undergoes preprocessing for direct visualization and analysis. The machine learning models from the batch layer will be used to analyze the data in the streaming layer, and a postprocessing step is necessary to reduce false positives and extract the significant information to display in a dynamic visualization mode. In addition, the system includes a service layer that serves as an intermediary between the historical data and the streaming layer. The machine learning models generated from the historical data are forwarded to the streaming layer through this service layer. The system also incorporates a webserver interface between the streaming layer and the application layer, providing a seamless integration of the two layers. The application layer, located on the user side, presents a dashboard for interaction of the user with the various components of the monitoring system. This layer enables the user to visualize real-time data before and after the optimization process through a web application. This feature provides the user with a comprehensive understanding of the health status of the patient being monitored.

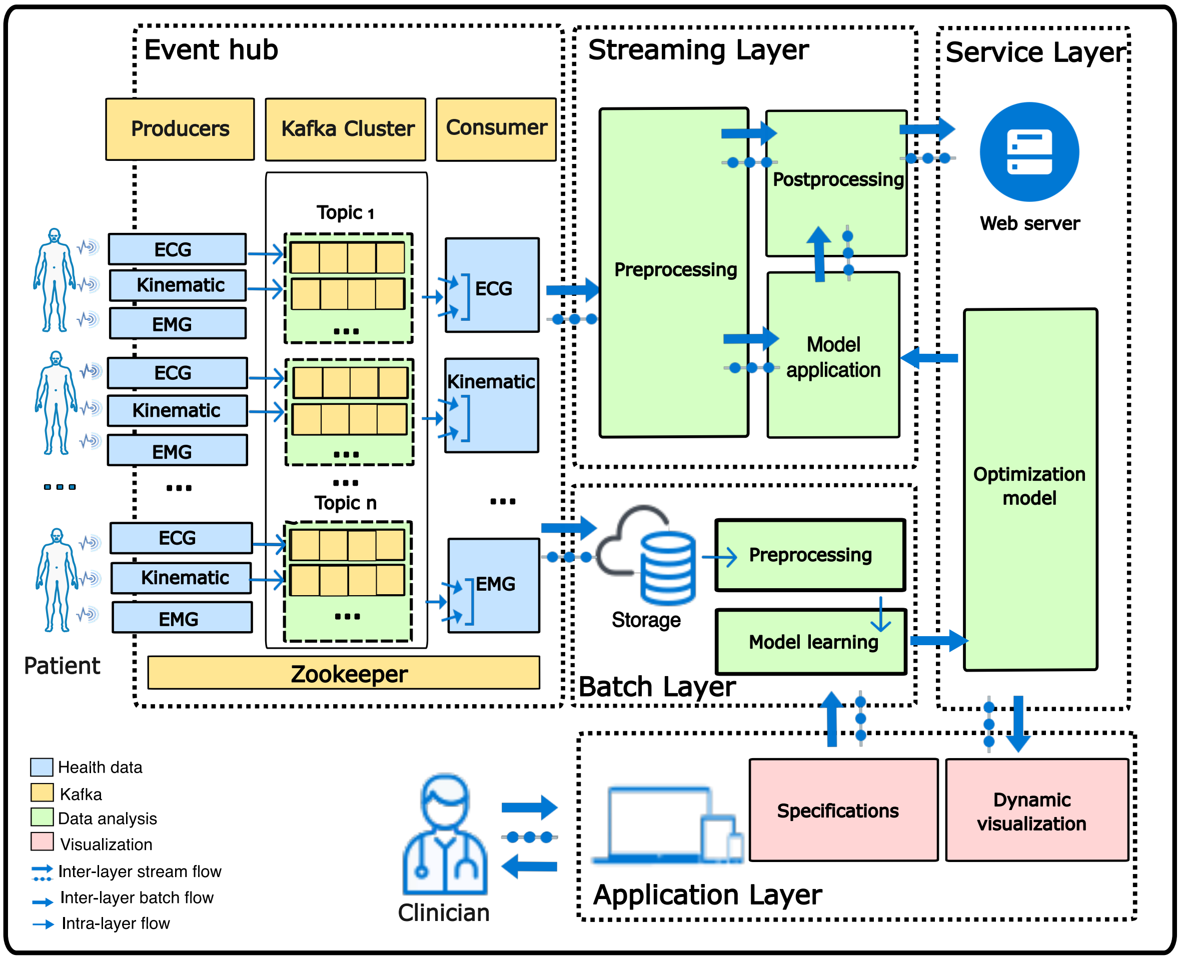

Figure 1. Architecture of the real time visual analytics framework for remote health monitoring and the structure of Kafka pipeline

The proposed system facilitates the interaction through customizable display modes and adjustable specifications of relevant information. For instance, in the context of post-stroke rehabilitation monitoring, the user can select the motor function data to visualize, and the specific optimized alerts sent to the therapist. This could enhance the evaluation of the patient's performance and health status, leading to a more personalized therapy approach. The multiple selective alerts enable real-time feedback to the clinician, providing crucial support in the case of emergency situations and facilitating the adjustment of rehabilitation protocol as necessary. This feature brings an important added value for the effectiveness and efficiency of the real-time patient health monitoring framework.

4. APPLICATION OF THE PROPOSED FRAMEWORK TO POST STROKE IN-HOME REHABILITATION

The objective of the development of this framework is to propose a real-time visual analytics system to assist physiotherapists in monitoring post-stroke patients when they practice rehabilitation eservices at home. The system described in this paper is focused on real-time visualization of medical data streams. The development of the framework was carried out following the principle of user-centered design (UCD) to ensure its effective usability. This process aims to analyze the application context and user requirements, design and develop the prototypes and evaluate the, effectiveness of the system. Thus, by repeating the activities flow until the usability is redeemed by the rehabilitation experts who are the target end users of the system. The UCD process was performed by generating three iterative cycles: analysis, design, prototyping, testing, and design refinement (Figure 2). This section provides a detailed description of the research activities and the methods used in each iteration.

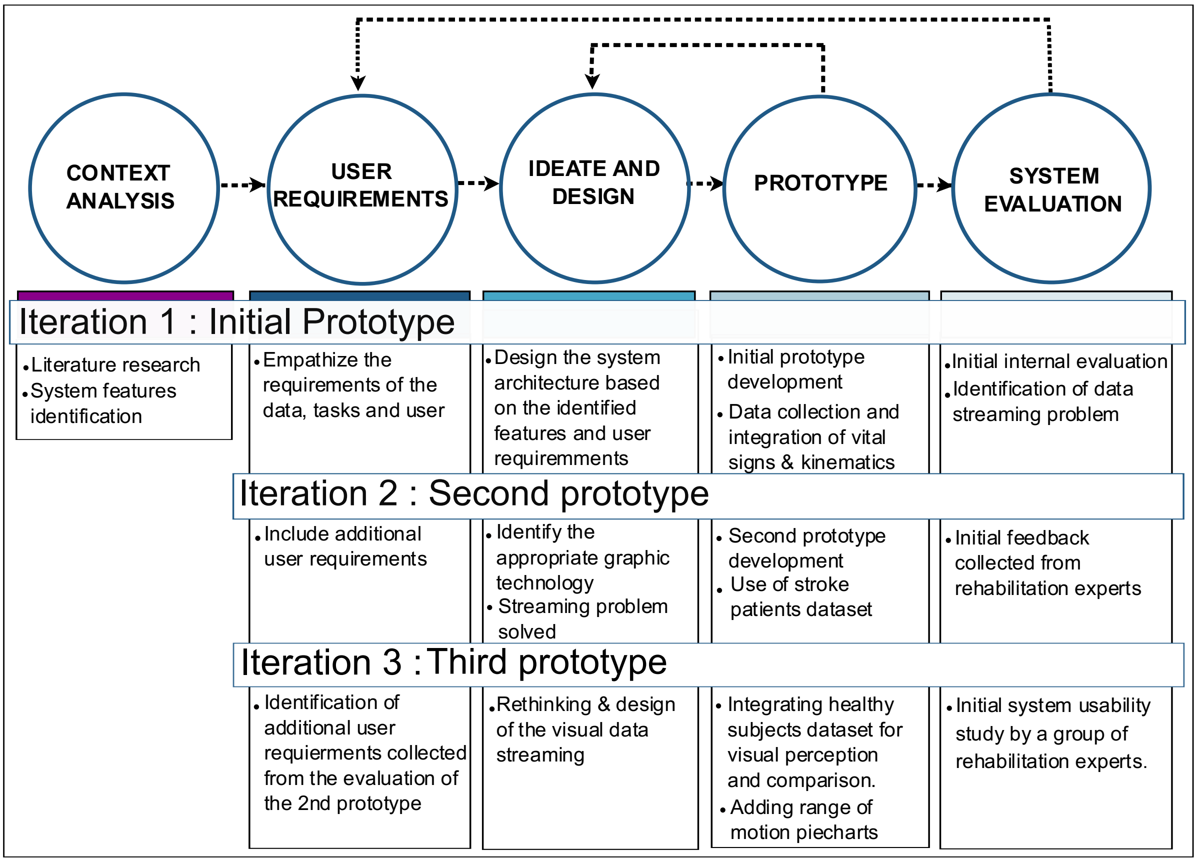

Figure 2. The workflow of the development process of the visual analytics framework

First Iteration

After analyzing the context of the proposed framework, it was critical to identify the requirements of the rehabilitation experts as they are the main end-users of the system.

In the preliminary iteration of the system design, emphasis was made on the identification of the requirements of the rehabilitation experts. This task was carried out by the research team members who have expertise in neuromotor rehabilitation practices, and they were in regular interaction with physiotherapists and neurologists in the hospital. Thus, for a more exhaustive understanding of the user requirements, the design triangle: Data–Users–Tasks was identified [30] and adopted as guideline:

Data requirements: The classification of patient-generated health data (PGHD) obtained during post-stroke rehabilitation sessions is crucial to provide meaningful insights into the patient's well-being. To this end, the PGHD are categorized into four key outcome measures: quality of life measures, activity measures, balance measures, and motor function measures [31]. The quality-of-life measures provide a holistic evaluation of the patient's physiological and psychological well-being, including vital signs and emotions. The activity measures evaluate the patient's ability to perform rehabilitation exercises, while the balance measures inform about the patient's ability to maintain stability. Finally, the motor function measures monitor physical capabilities such as range of motion and muscle strength. In light of the aforementioned categorization, the health data generated during post-stroke rehabilitation sessions is primarily numeric, time-oriented, and multivariate in nature, making it challenging to manage through conventional systems due to its alignment with the requirements of Big Data such as high volume, rapid velocity, diverse variety, complex variability and time-sensitive nature.

Tasks and user requirements: The proposed system is intended for rehabilitation specialists such as rehabilitation physicians, physiotherapists, and human movement scientists. The main objective is to enable remote monitoring of stroke patients during rehabilitation sessions. To achieve this objective, the system must provide users with access to real-time and historical health data through intuitive graphic representations and an interactive dashboard. This will enable rehabilitation experts to assess patients' physical abilities to perform the prescribed exercises and monitor their health status during rehabilitation practices. Additionally, the system should enable the analysis, interpretation, and utilization of collected data to support clinical decision-making and detect potential abnormalities. Furthermore, in order to optimize the therapy protocol, the users expressed the need for comparing real-time data with historical datasets from previous sessions for the same patient or with comparison with other patients. To fulfill these requirements, the system's dashboard and data representation tools must be user-friendly, enable simple and straightforward manipulation of data to extract the desired information. The goal is to ensure that the system is accessible and easy to understand, allowing users to effectively analyze and utilize the data.

After the identification of the user requirements, the design phase was initiated. It was focused on the selection of open-source technologies to implement a unified framework for post-stroke rehabilitation monitoring. To meet the specified needs, Apache Kafka was selected as the central event hub, with Python libraries for data processing and Dash was used for graphical visualization. Apache Kafka has an open and distributed architecture. It can handle real-time data ingestion from multiple producers, scale up to handle high-volume data, offer low latency, and ensure robust reliability [32]. These features made it well suited for this application. In a hypothetical scenario where there are no network delays, the processing is efficient, and the hardware is properly configured, a Kafka system with five brokers and 12GB of memory could sustain around 55,000 events per second [33]. If each patient wore five sensors generating 50 events per second, the system could theoretically monitor up to 220 patients:

$$220\ patients = \frac{55.000\ events\ per\ second}{50\ events\ per\ second * 5\ sensors\ per\ patient}$$

However, it's important to keep in mind that this is a simple approximate estimation but the actual performance depends on various factors such as the complexity and size of the events, the processing load on the system, CPU resources, hard disk I/O speed, network bandwidth, message size and frequency, and the configuration parameters [34].

The development of the first prototype of the system was carried out using Python 3.8.2, Kafka 2.13-2.8.0, and Dash 2.0.0. The Dash framework was chosen for data visualization due to its flexibility and easy connectivity with Kafka using Python. Dash is built on the Flask web providing a simple and straightforward interface for creating rich visualizations and dashboards based on complex data and allows users to interact with the data through various controls and filters. As shown in Figure 1, the data collected from each type of sensor for a single patient was represented by a producer. The data collected from each type of sensor for a specific patient were transmitted by a producer to the corresponding partitions within a unique topic representing the patient. This organizational structure facilitated the subsequent processing and analysis of data. The data was then consumed based on their type but not on the patient ID, as they were organized into consumer groups to facilitate the analysis in the following step. The processed data were then transmitted to the Dash application via a web server for real-time display.
The evaluation of the system's functioning is important for the preliminary prototyping stage. In order to assess the system's performance, two datasets related to rehabilitation were selected for analysis. The first dataset consists of hand kinematics data of 22 healthy subjects while performing daily activities [35]. The second dataset includes vital sign recordings of 30 healthy individuals, including ECG, blood pressure, and respiration rate [36]. To mimic real-time data flow, the datasets were continuously looped and ingested into the prototype. The user dashboard provides a visual representation of the processed data, as illustrated in Figure 3. This approach allowed for the thorough evaluation of the system's capabilities, providing valuable insights into the system's performance, and enabling further improvements.
An initial usability evaluation of the system was conducted by the research team to assess its performance. The evaluation involved integrating four data streams into the prototype pipeline and assessing their real-time display.

Results of the evaluation of the first prototype:
During the running of the framework, the system initially displayed the data properly with an acceptable latency for a short time but after few minutes, and as the number of streams increased, the live graph display was slowed down reaching a latency of three seconds and later the flow of data was completely suspended at the application layer of the prototype. This issue had a significant impact on the aspect of real-time data visualization which had to be addressed in the second iteration. The results of this evaluation provided important insights into the limitations of the system and areas for improvement, highlighting the

need for further optimization of the prototype to meet the requirements of real-world applications and the needs of the target user. The objective was to obtain smooth, scalable, and low-latency data visualization experience.

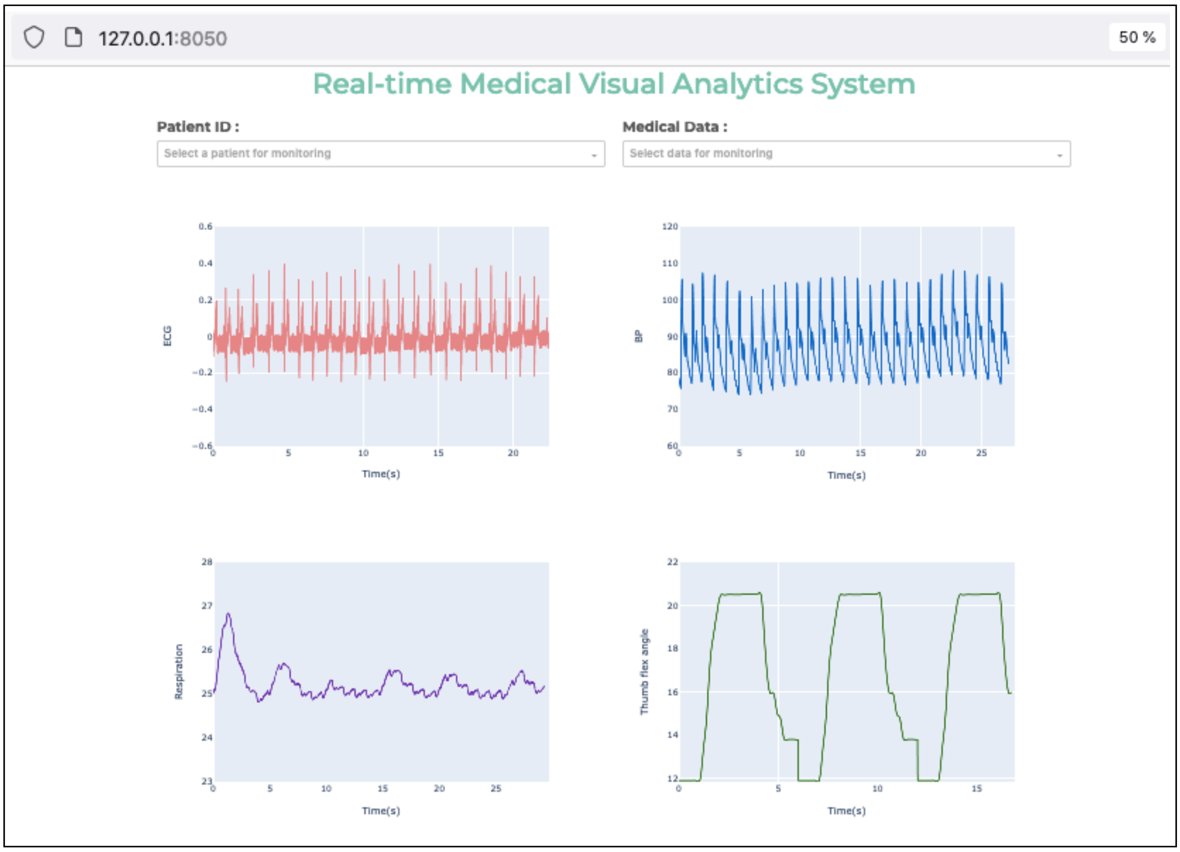

Figure 3. First prototype: visualization of the multimodal health data streams in the dashboard

Second Iteration

In order to address the limitations of the first prototype, a second development iteration was necessary. Then, an evaluation of the system's performance was carried out using real datasets of post-stroke patients practicing rehabilitation exercises.

To fulfil the new identified requirements, the technology used for graphical display using Dash was reviewed in order to avoid the blockage of data stream in time and ensure smooth display of the graphical representations. In the first prototype, the graphs were based on SVG (Scalable Vector Graphics) and multiprocessing. The multiprocessing in Python guarantees the parallelism of multiple flows as processes do not share the same memory space. However, multithreading could be a better choice for the current system because the threads are lighter and less likely to cause overload. Additionally, the SVG is an easy option for rendering high-quality vector graphics, but its performance is limited. As an alternative solution, the WebGL provides a JavaScript API that allows to create GPU-accelerated graphics. The second prototype was developed by replacing the old tools and inbuilding the combination of Multithreading method and WebGL technology. As a result, the visualization of the data flow in the implemented system is currently smooth and continuous in real-time. It supports a significant amount of data streams, including nine types of heavy streams that are updated every 100ms. Additionally, the system maintains a very acceptable latency that does not exceed one second. In order to assess the performance of the prototype, it was essential to conduct tests using real data that simulate a typical rehabilitation session for post-stroke patients. Various datasets from real stroke patients have been reported in the literature.

The recent dataset of post-stroke upper limb kinematics of daily living tasks (UZH) was deemed to be the most suitable in the existing datasets of stroke rehabilitation due to its corresponding number of subjects (20 stroke patients and five healthy individuals) and the variety of the covered upper limb rehabilitation activities (30 exercises). This data was collected by Averta et al. as part of U-limb, which is a large and multi-modal database that was released to help the research contribution towards the assistive rehabilitation of the upper extremity of post stroke survivors [37]. To acquire the upper limb kinematic data, inertial wearable motion capture sensors of the type of Xsens MVN Awinda were used. The sensors were attached to various locations on the subject's body, including above the sternum, shoulder blade, upper arm, forearm, and back of the hand (Figure 4). The subjects undergoing the rehabilitation exercise were instructed to perform 30 daily living activities with both sides of their upper limb, performing each activity three times [37]. The kinematic measures used to evaluate the performance of the second prototype in this study included shoulder flexion/extension, shoulder abduction/adduction, shoulder internal/external rotation, elbow flexion/extension, elbow pronation/supination, wrist flexion/extension, wrist abduction/adduction and wrist pronation/supination.

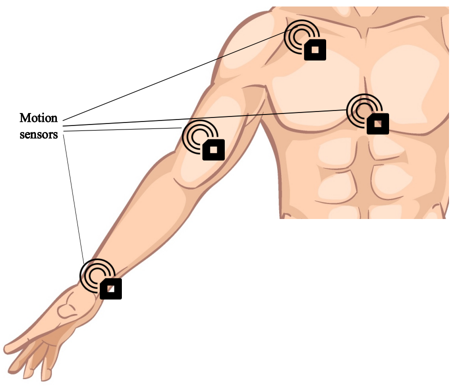

Figure 4. Motion sensors attached to the patient's body for data capture [38]**.**

These measures provided a comprehensive evaluation of the system's ability to accurately stream and visualize the upper limb movements of stroke patients during rehabilitation activities.

The second prototype features a user interface that displays live-streamed data from the patient alongside the standard deviation (SD) of data from healthy subjects used as a reference for the targeted performance of the patient during rehabilitation exercises. Figure 5 illustrates an example of a graph of shoulder flexion/extension for a stroke patient performing the task of reaching and grasping a glass of water, drinking for three seconds, and returning the glass to its initial position. This prototype was initially evaluated by a panel of four rehabilitation experts to collect end users' feedbacks.

Results of the evaluation of the second prototype: The experts identified two main requirements for the system improvement. The first issue pertained to the difficulty of comparing the performance of the stroke patient and healthy subject with the naked eye while the patient data graph was updating in real-time. The experts recommended implementing visual aids to illustrate the comparison and facilitate more accurate analysis of the patient's performance. The second requirement highlighted by the panel of experts was the need for control over the pace of data display. This would enable the user to slow down or pause the data stream to obtain thorough analysis and relevant information extraction. This could also allow the option to speed up the stream to keep pace with real-time updates. This level of control over the pace of data display would provide greater flexibility and usefulness for the rehabilitation therapist.

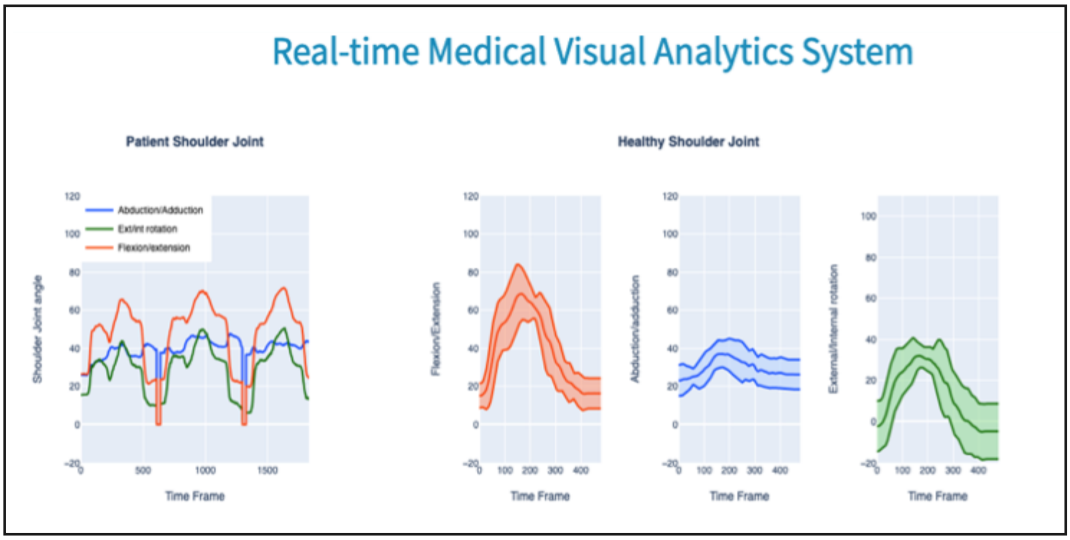

Figure 5. Second prototype: visualization of real-time data flow of the shoulder joint angle of stroke patient Vs healthy subjects

Third Iteration

Based on the feedback received from the evaluation of the second prototype, additional requirements were integrated to improve its functionality. To address the challenge of comparing the performance of patients with healthy subjects, the real-time data of the patients were superimposed with the pre-recorded data and the SD of healthy subjects. This could provide better visual comparison of the patient's performance during rehabilitation activities. The range of motion of the patient and the healthy subjects were calculated and displayed through pie charts in the dashboard. The user interface also displayed important information about the patient, including their ID, age, gender, the time since the brain stroke took place, and an indication about which limb is affected by the stroke. A description of the current exercise instructions was also included in the dashboard. Figure 6 shows an example of the user interface, displaying live stream data of a patient performing an activity of daily living with upper limb data of the monitored patient mapped to healthy subjects' data.

Results of the evaluation of the third prototype:
In order to evaluate the third version of the prototype, a group of 16 rehabilitation practitioners used the dashboard of the framework to monitor the streamed data used in the previous iterations. The subjects were requested to map the displayed data on the kinematics of the patient's limb and interpret his/her progress with reference to healthy subjects. The feedback of experience was collected by means of questionnaire about the appreciation of the visualization system by the users. The analysis of the collected data revealed that the rehabilitation practitioners considered that the system allowed them to interpret the visualized data. They also declared that the system is beneficial for monitoring patients receiving rehabilitation therapy in home settings. Additionally, they expressed their desire to use the final version of the system in their protocols for patient care.

5. CONCLUSION AND RESEARCH PERSPECTIVES

The objective of this research was to develop and evaluate a unified real-time visual analytics framework for remote patient monitoring. The development process of the proposed framework was carried out in three iterations and following the principle of User Centered Design (UCD). The system was applied for monitoring post-stroke survivors receiving neuromotor rehabilitation treatment. In order to comply with the principle of involving end-users since the design stage, each iteration was evaluated, and the collected feedbacks were analyzed and used for the refinement of the framework. The final prototype of the developed framework allowed a group of target users to monitor patient data during the practice of rehabilitation exercises. Real-time monitoring is not to be limited to the visualization of the current health data of patients. It should also allow predicting the evolution of patients' health during and after the rehabilitation sessions. This aspect will be addressed in the future of this research by the development of an intelligent module based on machine learning for health prediction.

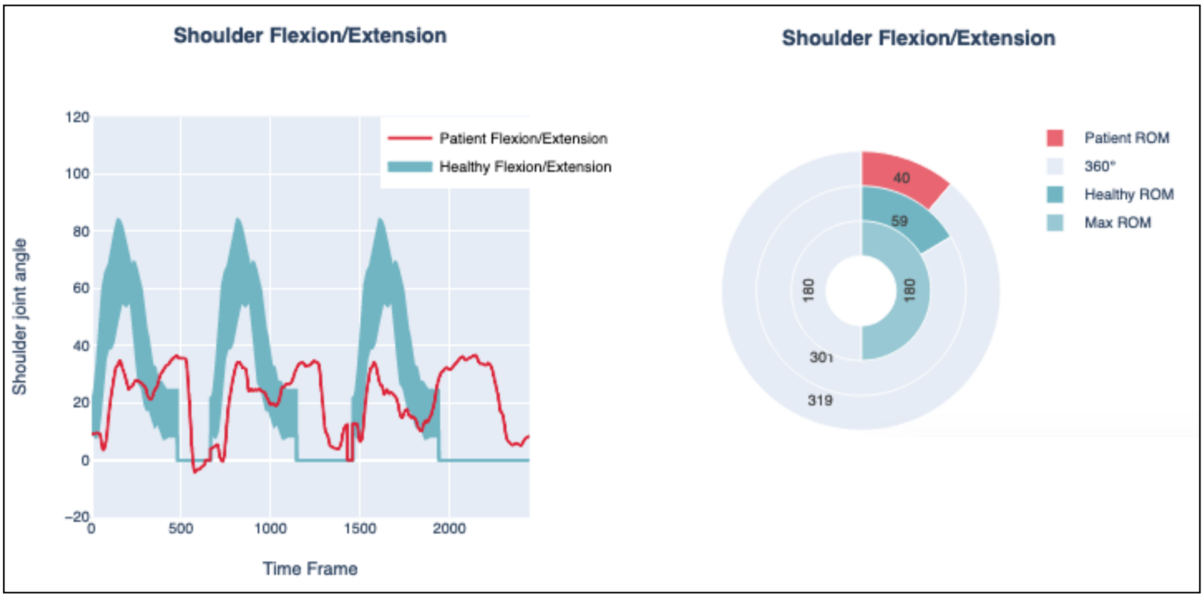

Figure 6. Third prototype: Shoulder's Flexion/Extension and Range of motion related to the current exercise for both patient (red) and the average of healthy subjects

https://www.doi.org/10.24132/CSRN.3301.61

REFERENCES

[1] J. J. Caban and D. Gotz, "Visual analytics in healthcare - opportunities and research challenges," *J. Am. Med. Informatics Assoc.*, vol. 22, no. 2, pp. 260–262, 2015, doi: 10.1093/jamia/ocv006.

[2] N. Kamal, "Big Data and Visual Analytics in Health and Medicine: From Pipe Dream to Reality," *J. Heal. Med. Informatics*, vol. 05, no. 05, pp. 998–1000, 2014, doi: 10.4172/2157-7420.1000e125.

[3] J. W. Tukey, "Exploratory Data Analysis," in *The Concise Encyclopedia of Statistics*, New York, NY: Springer New York, 2008, pp. 192–194.

[4] D. Keim, J. Kohlhammer, G. Ellis, and F. Mansmann, *Mastering The Information Age – Solving Problems with Visual Analytics*. Eurographics Association, 2010.

[5] P. C. W. John Dill, Rae Earnshaw, David Kasik, John Vince, *Expanding the Frontiers of Visual Analytics and Visualization*. Springer Science & Business Media, 2012.

[6] "Apache Kafka." https://kafka.apache.org/.

[7] "Dash." https://dash.plotly.com/dash-html-components.

[8] S. Park and A. D. Flaxman, "Understanding data use and preference of data visualization for public health professionals : A qualitative study," *Public Health Nurs.*, no. December 2020, pp. 1–11, 2021, doi: 10.1111/phn.12863.

[9] G. Harerimana, S. Member, and B. Jang, "Health Big Data Analytics : A Technology Survey," *IEEE Access*, vol. 6, pp. 65661–65678, 2018, doi: 10.1109/ACCESS.2018.2878254.

[10] K. Sawhney and D. F. Sittig, "Information Overload and Missed Test Results in EHR-based Settings," vol. 173, no. 8, 2014, doi: 10.1001/2013.jamainternmed.61.Information.

[11] P. West, M. Van Kleek, R. Giordano, M. J. Weal, and N. Shadbolt, "Common Barriers to the Use of Patient-Generated Data Across Clinical Settings," *CHI Conf. Hum. Factors Comput. Syst.*, pp. 1–13, 2018.

[12] M. Musterman *et al.*, "Big Data in Health Care : What Is So Different About Was ist so anders am Neuroenhancement ?," *Data Inf. Manag.*, vol. 2, no. 3, pp. 175–197, 2018, doi: 10.2478/dim-2018-0014.

[13] S. Kliment, F. Information, and C. Technologies, "Big Data Analytics in Medicine and Healthcare Abstract :," *J. Integr. Bioinform.*, pp. 1–5, 2018, doi: 10.1515/jib-2017-0030.

[14] K. K. Khedo *et al.*, "Health Data Analytics: Current Perspectives, Challenges, and Future Directions," in *IoT and ICT for Healthcare Applications*, N. Gupta and S. Paiva, Eds. Cham: Springer International Publishing, 2020, pp. 117–151.

[15] A. Seals *et al.*, "'Are They Doing Better In The Clinic Or At Home?': Understanding Clinicians' Needs When Visualizing Wearable Sensor Data Used In Remote Gait Assessments For People With Multiple Sclerosis," 2022, doi: 10.1145/3491102.3501989.

[16] D. J. Feller *et al.*, "A visual analytics approach for pattern-recognition in patient-generated data," *J. Am. Med. Informatics Assoc.*, vol. 25, no. 10, pp. 1366–1374, 2018, doi: 10.1093/jamia/ocy054.

[17] A. L. Vu, "Visual Analytics: Identifying Informative Temporal Signatures in Continuous Cardiac Monitoring Alarms from a Large Hospital System," Diss. The Ohio State University, 2017.

[18] A. Dagliati *et al.*, "A dashboard-based system for supporting diabetes care," *J. Am. Med. Informatics Assoc.*, vol. 25, no. 5, pp. 538–547, 2018, doi: 10.1093/jamia/ocx159.

[19] E. V Hobson *et al.*, "The TiM system : developing a novel telehealth service to improve access to specialist care in motor neurone disease using user-centered design," *Amyotroph. Lateral Scler. Front. Degener.*, vol. 19, no. 5–6, pp. 351–361, 2018, doi: 10.1080/21678421.2018.1440408.

[20] L. Griffin *et al.*, "Creating an mHealth App for Colorectal Cancer Screening : User-Centered Design Approach," *JMIR Hum. factors*, vol. 6, 2019, doi: 10.2196/12700.

[21] A. Raghu, D. Praveen, D. Peiris, L. Tarassenko, and G. Clifford, "Engineering a mobile health tool for resource-poor settings to assess and manage cardiovascular disease risk : SMARThealth study," *BMC Med. Inform. Decis. Mak.*, pp. 1–15, 2015, doi: 10.1186/s12911-015-0148-4.

[22] U. Backonja, S. C. Haynes, K. K. Kim, and K. K. Kim, "Data Visualizations to Support Health Practitioners ' Provision of Personalized Care for Patients With Cancer and Multiple Chronic Conditions : User-

Centered Design Study," *JMIR Hum. factors*, vol. 5, pp. 1–20, 2018, doi: 10.2196/11826.

[23] C. L. Petersen *et al.*, "Using Natural Language Processing and Sentiment Analysis to Augment Traditional User-Centered Design : Development and Usability Study," *JMIR mHealth uHealth*, vol. 8, pp. 1–13, 2020, doi: 10.2196/16862.

[24] L. David, G. Bouyer, and S. Otmane, "Towards a low-cost interactive system for motor self-rehabilitation after stroke," *Int. J. Virtual Real.*, vol. 17, no. 2, pp. 40–45, 2017, doi: 10.20870/IJVR.2017.17.2.2890.

[25] T. Caporaso, S. Grazioso, D. Panariello, G. Di Gironimo, and A. Lanzotti, "Understanding the Human Motor Control for User-Centered Design of Custom Wearable Systems: Case Studies in Sports, Industry, Rehabilitation," in *Design Tools and Methods in Industrial Engineering*, 2020, pp. 753–764.

[26] C. L. Osborne, S. B. Juengst, and E. E. Smith, "Identifying user-centered content , design , and features for mobile health apps to support long-term assessment , behavioral intervention , and transitions of care in neurological rehabilitation : An exploratory study," *Br. J. Occup. Ther.*, 2020, doi: 10.1177/0308022620954115.

[27] M. Wentink *et al.*, "How to improve eRehabilitation programs in stroke care ? A focus group study to identify requirements of end-users," *BMC Med. Inform. Decis. Mak.*, vol. 3, pp. 1–11, 2019.

[28] M. Boumrah, S. Garbaya, and A. Radgui, "Real-time visual analytics for in-home medical rehabilitation of stroke patient — systematic review," *Med. Biol. Eng. Comput.*, pp. 889–906, 2022, doi: 10.1007/s11517-021-02493-w.

[29] J. Marz, Nathan ; Warren, *Big data : principles and best practices of scalable real-time data systems*. Simon and Schuster, Shelter Island, NY : Manning, 2015. - 308 p, 2015.

[30] S. Miksch and W. Aigner, "A matter of time: Applying a data-users-tasks design triangle to visual analytics of time-oriented data," *Comput. Graph.*, vol. 38, no. 1, pp. 286–290, 2014, doi: 10.1016/j.cag.2013.11.002.

[31] G. L. Dimaguila, K. Gray, and M. Merolli, "Person-generated health data in simulated rehabilitation using kinect for stroke: Literature review," *JMIR Rehabil. Assist. Technol.*, vol. 20, no. 5, 2018, doi: 10.2196/rehab.9123.

[32] B. R. Hiraman, M. C. Viresh, and C. K. Abhijeet, "A Study of Apache Kafka in Big Data Stream Processing," *2018 Int. Conf. Information, Commun. Eng. Technol. ICICET 2018*, pp. 1–3, 2018, doi: 10.1109/ICICET.2018.8533771.

[33] H. Jafarpour, R. Desai, and D. Guy, "KSQL: Streaming SQL engine for Apache Kafka," *Adv. Database Technol. - EDBT*, vol. 2019-March, pp. 524–533, 2019, doi: 10.5441/002/edbt.2019.48.

[34] Apache Software Foundation, "Kafka 3.4 Documentation." https://kafka.apache.org/documentation/#performance.

[35] N. J. Jarque-Bou, M. Vergara, J. L. Sancho-Bru, V. Gracia-Ibáñez, and A. Roda-Sales, "A calibrated database of kinematics and EMG of the forearm and hand during activities of daily living," *Sci. Data*, vol. 6, no. 1, pp. 1–11, 2019, doi: 10.1038/s41597-019-0285-1.

[36] S. Schellenberger *et al.*, "A dataset of clinically recorded radar vital signs with synchronised reference sensor signals," *Sci. Data*, vol. 7, no. 1, pp. 1–11, 2020, doi: 10.1038/s41597-020-00629-5.

[37] G. Averta *et al.*, "U-Limb: A multi-modal, multi-center database on arm motion control in healthy and post-stroke conditions," *Gigascience*, vol. 10, no. 6, p. giab043, Jun. 2021, doi: 10.1093/gigascience/giab043.

[38] A. Schwarz *et al.*, "Characterization of stroke-related upper limb motor impairments across various upper limb activities by use of kinematic core set measures," *J. Neuroeng. Rehabil.*, vol. 19, no. 1, pp. 1–18, 2022, doi: 10.1186/s12984-021-00979-0.