HEALTHCARE INFORMATION SYSTEM FOR TUBERCULOSIS AND THE CREATION OF EVENT LOG FOR PROCESS MINING

Bukohwo Michael Esiefarienrhe

MaSIM, Computer Science Department,
North-West University, Mafikeng, South Africa
Michael.Esiefarienrhe@nwu.ac.za

Abdul Ncube

MaSIM, Computer Science Department,
North-West University, Mafikeng, South Africa
abdulncube1@gmail.com

Abstract

The implementation of healthcare Information System (HIS) has had a major impact on the interpretation of the vast amount of data available in the medical field. This has led to techniques relating to data science such as process mining (PM) which has emerged as a set of methods and tools aimed at discovering and analysing process execution. Researchers have capitalized on this to solve problems are relating to time delay in diagnosis and treatment process. However, there still areas in the healthcare that lacks the application of PM to better discover and analyse process executions. One such area is tuberculosis (TB), the disease which has been the cause of considerable morbidity and mortality worldwide. Early diagnosis and prompt initiation of treatment are essential for an effective TB control program. Delayed diagnosis and treatment of tuberculosis TB leads to greater morbidity and mortality and can also increase the rate of infection within the community. Moreover, with PM as a promising technique to combat the above-mentioned problems there are still issues with applying PM to the existing data in the HIS. Since event data is typically not stored in a process-oriented manner, an event log should be generated first. Literature shows that event log generation takes a substantial effort in PM projects. Our goal was to design and develop a software system capable of collecting and storing event log data for TB patients as this would bridge the identified gaps. The problem of lack of data with timestamps of events has been a major concern to TB specialist and process miners. The time at which each event occurs in the diagnosis and treatment of TB is crucial if the disease behaviour, diagnosis, and treatment pathways are to be understood. Therefore, this study designed and developed a software system called ELS using Design Science Research (DSR) methodology and using the JavaScript Node.js platform that can track the underlying processes in the diagnosis and treatment of TB and thus stored the data and timestamps in a secured databases and cast the data as events log for the purpose of adequate mining and conformity checks of the process of TB. The results obtained shows that the database developed could assist to unravel the diagnosis and treatment of TB from data analysis perspective.

***Keywords*: Event log for tuberculosis; Process mining for tuberculosis; Health Information Systems; diagnosis and treatment of tuberculosis; software for tuberculosis log creation.**

1. Introduction

Data timeliness, availability, safety and protection, privacy of information is instrumental to any data management efforts or policy regulating the dissemination and use of data as a resource. The positive impact of data mining, process mining (PM), big data, and even artificial intelligence (AI) all revolves around data available and timeliness. The statement “data is not available for research” is quite common and worrisome in the research community. The direction where research efforts should seriously be directed is making data available particularly in the healthcare domain. Health is wealth and health is life, says a popular adage. It is therefore essential to find a way to bridge the gap in data availability by developing a software system that

captured TB related data as patients undergoes consultation, diagnosis and/or treatment in a health facility. The data captured are stored in a format that is easily assessable and useable for PM of TB.

Tuberculosis (TB) is caused by a bacterium called Mycobacterium tuberculosis. The disease could affect the lungs and other parts of the body. Researchers are constantly seeking data to perform research in the healthcare sector and motivated by the fact that around 10 million people worldwide contract TB every year (Suárez et al.,2019). According to the World Health Organization (WHO), approximately one-quarter of the world's population is latently infected with Mycobacterium tuberculosis. Rapid diagnosis and targeted treatment are essential to prevent an unfavourable course of the disease as well as its transmission to other individuals (Suárez et al.,2019). There are global efforts focusing curbing the spread and fast-tracking the treatment of TB. One way such efforts can be successful is to encourage the sharing of TB related data among healthcare providers. Hence, the aim to design and develop software system that can capture and store event log data for TB patients whereby the system is designed with focus centred around the diagnosis and the treatment of TB.

The objective of PM is to derive knowledge from data stored in systems as events logs by using models to analyse patterns and underlying trends in the event logs of processes as shown in figure 1. Event logs can be viewed as sets of registers in which all transactional activities of a process are internally recorded by the process itself (Pereira, 2020; Suriadi, 2013). Events are recorded in event logs, which are lists of the activities that occurred throughout a business process. It enables analysts to exploit information about the actual performance of these processes for conformance analysis, analysis of decision pathways and the role of human involvement in the various phases that constitute the system. In an event log file, events are noted together with their characteristics.

Clinical procedures whether administrative, diagnosis or treatment are built on processes, and clinical guidelines developed by healthcare professionals are frequently characterized as processes that show how/why a disease behaves the way it does and to direct the subsequent course of treatment. Additionally, PM offers instruments to improve the view by emphasizing the most fascinating aspects of the process, emphasizing comparison techniques, and facilitating evaluation of applied health protocols (Carmona, 2018). PM approaches may not always be the most precise data-driven method of solving machine learning challenges, but their capacity to be understandable by humans suggests a new way to draw knowledge from the data-driven world. Through the use of interactive models, PM enables the fusion of the data and knowledge-driven worlds (Fernández-Llatas, 2013).

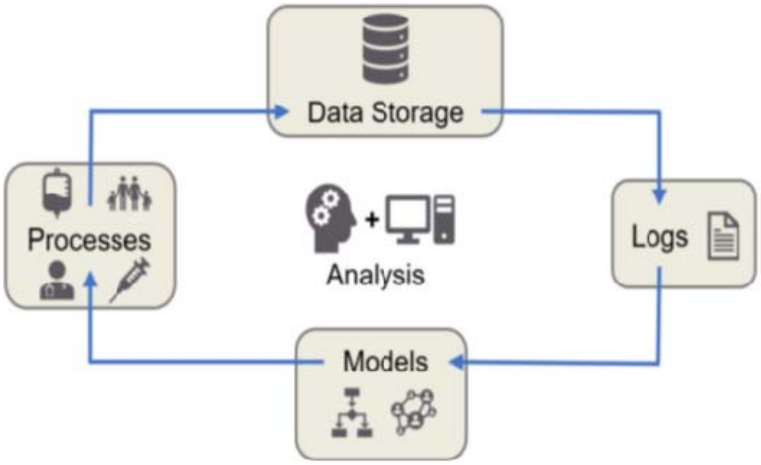


Figure 1: Overview of PM in healthcare (Pereira, 2020).

The growing interest in PM is fuelled by increasing data availability. PM technology uses event logs to automatically discover process models, validate compliance, identify bottlenecks and deviations, suggest improvements, and predict processing times (Bose, 2013). Event logs can be defined as lists of activities together with their attributes namely Case IDs, timestamps (start and end) of activities, and other characteristics of the peculiar to the process recorded by the health information system (HIS). An event log can be used to

record several linked business processes as well as one or more instances of a particular business process. Event logs when used as input to a system, is critical for the success of any PM effort (Bose, 2013).

Even though the success stories associated with PM are undoubtedly impressive, it can be challenging to replicate these best practices in many contexts due to the unavailability and where available, the quality of the event logs. The success of the PM activity depends critically on the quality of the data (in both form and logs) supplied to the modelling algorithms Suriadi, 2017). Application of PM in the healthcare domain normally encounter several major challenges which are mainly posed by:

I. *Missing Events:* When one or more events are absent from the trace despite having occurred, this one of the quality issues present.
II. *Incorrect Timestamps:* This quality problem arises when (some or all) events' recorded timestamps in the log do not match the actual time at which the events took place.

The above mentioned challenges are compelled by the complexity of the data involved, which includes characteristics of the patients' illness, behaviour, clinical history, psychology, etc (Gatta, 2018).

Hence, this research project seeks to design the event logs suitable to PM in the healthcare domain with TB as the area of interest and furthermore develop a software artifact (prototype) to capture event data about TB and produce event logs suitable for PM from the captured data. Therefore, the aim of this research is to design and develop software system that can capture and store event log data of TB patients that can be used for PM. In carrying out the above task, the following research questions (RQ) will be addressed:

RQ1: What is the state of research on the availability of event logs for processing mining of in the healthcare domain?
RQ2: How can UML tools be used to design event logging system (ELS) for capturing and storing TB patients' data for the purpose of PM of TB?
RQ3: How can a software be developed to implement the UML design in RQ2?

The research questions will be addressed using the following research objectives (RO):

RO1: Conduct a search of literature review of the current state of research on the availability of event logs for processing mining in the healthcare domain.
RO2: Use UML tools to design Event logging system (ELS) for capturing and storing TB patients' data for the purpose of PM of TB?
RO3: Develop a software to implement the UML design in RO2.

Innovation improves, lowers the cost of, and increases the effectiveness of healthcare. Healthcare is advanced by developments like new business models and technologies (Herzlinger, 2006). Additionally, healthcare systems around the world face unprecedented challenges, including the constant and rapid adaptation of clinical processes based on new scientific knowledge (Keeley, 2020). In this context, healthcare organizations such as hospitals recognize the need to improve both clinical processes (such as treatment pathways that describe the treatment of specific medical conditions over time) and organizational/administrative processes (such as billing practices) (Martin, 2020). These highlighted clinical processes can be improved through PM (PM), but for PM to be efficient there must be event logs in place and event logs of quality. This Research is fuelled by motivation to combat issues of Event logs mentioned above on the section of problem statement and the aim mentioned above is to be met by developing a Hospital Information System (HIS) that will capture several entries in the HISs are recorded (e.g., when a patient was registered or was subject to a clinical examination by a physician), during the execution of a process. Moreover, utilize the database records to create event logs outlining the series of actions taken, when they were taken, by whom, and for whom (for example, for which particular patient) (Suriadi, 2017). Lastly this research is to contribute new knowledge to automated Event logs for HIS by capturing the activities that constitute the pathway for TB treatment.

2. Related literature

2.1 Event Logs/data Challenges in the Healthcare Domain

In the healthcare domain several event logs/data challenges are compelled by the complexity of the involved data, which includes information about the disease, behaviours, clinical history, psychology, etc. of the patients.

Consequently, resulting in the lack of methods and tools for collecting data from existing and diverse EHRs which is the biggest problem that contributes to event logs being hard to gather, and cannot be analysed in Realtime. Existing tools for gathering data are extremely specific, are not well-integrated with PM pipelines and techniques, and require a significant pre-processing of data.

The issues listed above open room for the following challenges (i) the availability of data from the heterogeneous, fragmented, and sparsely connected Healthcare Information System (HIS) (Gatta, 2018); and (ii) the wide range of standards for data format and communication that are currently accessible. Building event logs that record all the phases of operating processes is essential for carrying out PM in the healthcare sector successfully. These event logs must be derived from the knowledge kept in Electronic Health Records (EHRs).

2.2 Availability of event logs suitable for PM in HIS

The data in event log serve as the foundation for PM. As a result, it follows that implementing PM techniques requires that the data have the right structure and content. HISs contain and provide a wide variety of transactional based event data and in this section, we explore whether the structure in which these data are captured is suitable for PM. This section is aimed at answering RQ1 and achieving RO1 as presented in section 1.0. The table 1 show the literature conducted on the availability of event logs in the healthcare domain.

Table 1: literatures on availability of event logs in HIS.

Authors	Purpose	Findings/Summary Points
Suriadi *et al.,* (2017)	Explains a group of data quality problems frequently discovered in PM event logs, extracted from experience of performing PM analysis.	Logs usually have many issues that need to be resolved before they can be used as input for analysis. May contain imprecise or inaccurate data and may not record certain important events (e.g., certain transaction types may not be recorded), or they contain data that needs to be interpreted carefully (e.g., timestamps).
Rojas *et al.,* (2016)	A survey of the literature on PM in the healthcare domain.	Some literatures on PM were reviewed and in their future work, they advised that it should focus on the process-aware implementation of HIS which have analytical capabilities to manage data and store it in a manner that enable applying PM to the data.
Gatta *et al.,* (2018)	Deriving a framework for Healthcare PM using event log generation and knowledge representation.	An emphasis is made that to effectively implement PM in healthcare, it is important to create an HIS that captures every step of the process in action. No system was developed or recommended that created such a system.
Oliveira *et al.,* (2020)	Outlines a novel approach for dealing with the intricacy of the occurrences in medical event logs.	This research offered an approach to deal with the complexity of activity-related event logs. Artificial labels are developed based on autoencoding to characterize these occurrences, and they can be employed in PM.
Kaymak *et al.,* (2012)	Even for clearly characterized clinical processes, it is argued that current PM methods fall short of finding appropriate process models.	The study concluded that physical systems described by dynamic signals, in addition to sequences of events, also contribute to the characteristics of medical processes. This means that we require a technique to convert the system's state-space description into a collection of associated events that define the process.
Homayounfar *et al.,* (2012)	Discussed the issues that would arise if healthcare information systems and PM were integrated.	Accurate data acquisition, data redundancy, and inaccurate and inadequate logging are the main challenges of PM in hospitals Information system.
Torres, Fried & Manjunath, (2018)	A medical ICU room's events and actions are automatically and unobtrusively monitored by the healthcare event and action logging (HEAL) system, which also reports on these events and activities.	ICU activities are monitored and identified by HEAL using a multimodal distributed camera network, which also makes estimates of sanitation-event qualifiers.
Timimi *et al.,* (2012)	Gives decision-makers and other stakeholders in TB control useful guidance for organizing changes to information systems, whether they are developing new ones or improving current ones.	Most TB data gathering and reporting still takes place on paper. As a result, it is labour-intensive, time-consuming, and wasteful to verify data quality, compile reports, update missing information, appropriate attribute treatment outcomes, and analyse and disseminate information which are additionally lengthy

The studies reviewed in table 1 has presented/highlighted the gap between the style/format of the data that is being stored in the healthcare facilities and the style/format of data that are needed for PM. Thus, an emphasis was made by one of the studies that future work should focus on implementing a process-aware HIS with the analytical capabilities to manage data and enable the application of PM correctly (Rojas *et al.,*2016). In summary it can be concluded that the availability of event logs suitable for PM in the healthcare domain are either not available or not accessible easily by third world countries where HIS administration is still basically done manually.

2.3 Systems Developed for TB

This section looks at the systems that have already being developed for TB and then discussed how the reviewed studies are related to this study. At the end, we highlighted the gap in literatures and the importance of developing event logging System to bridge the research gap identified. This section is also aimed at answering RQ1 and achieving RO1 as presented in section 1.0

Table 2: Literature on systems developed for TB literature

Authors	Purpose	Summary Points
Kim *et al.,* (2021)	They authors developed a deep learning algorithm for differential diagnosis.	The goal of this study was to create a deep-learning algorithm that uses colonoscopy pictures to differentiate between intestinal Bechet's disease (BD), Crohn's disease (CD), and intestinal tuberculosis (ITB). The research was on algorithm performance of identifying a particular type of TB and does not generate event log even for the TB identified as the algorithm executes.
Guo *et al.,* (2021)	To develop the grey self-memory system model to predict the incidence trend of infectious diseases.	The design of this study combines the prediction approach and empirical analysis to explore the prediction and early warning of infectious diseases using TB as a case study with nothing done on event logs.
Gambhir *et al.,* (2017)	Review of EPTB's common and unusual imaging characteristics.	The fast detection of TB and the monitoring of TB therapies are being evaluated and developed in this study. The study did not track the processes for a possible event log generation.
Konduri *et al.,* (2017)	To create the web-based eHealth system for e-TB Manager.	Users of the e-TB Manager comprised nurses, doctors, statisticians/data officers, and professionals/assistants in a cross-sectional, anonymous survey but did not consider how the underlying processes are stored as logs.
Riandari & Panjaitan, (2019)	Cantered on developing a system of compelled by experts' ideas to detect TB.	With the use of Bayes' Theorem, which is well known as the technique that is frequently used to diagnose illnesses in many circumstances., an expert system is an artificial intelligence program that mimics an expert by combining knowledge bases and information systems. No information given on the event log of the system.
Escombe *et al.,* (2009)	Tuberculosis Prevention using ultraviolet light and negative air ionization	Most airborne TB transmission that could be detected by guinea pig air sampling was avoided by upper-room UV lamps and negative air ionization. No information given on the event log of the system.
Li *et al.,* (2021)	A method that uses deep learning to produce quantitative CT reports for the diagnosis of pulmonary TB.	They gathered 501 CT imaging files from 223 patients with active PTB as well as an additional 501 datasets from populations who were healthy and served as negative samples. Professional radiologists painstakingly identified and labelled each dataset in the PTB. Following that, four cutting-edge 3D convolution neural network (CNN) models were trained and assessed using PTB CT images. No event logs of processes were generated. No information given on the event log of the system.
Muyoyeta *et al.,* (2017)	To define probable TB and to understand how each technique affects TB detection, they compared CXR and computer-aided diagnosis (CAD) to symptom screens.	To optimally rule out TB, screening strategies other than symptom screening should be considered. In particular, caution should be exercised for HIV-positive people who are at the highest risk of TB. No information given on the event log of the system.

From the studies reviewed on the developed system for TB in table 2, it can be concluded that most system developed were more focused on the diagnosing rather than on optimization of processes or data availability for better understanding of treatment etc. for TB. None of the system presently used for both diagnosis and treatment of TB were focussed on the creation of event log from the various activities that constitute their system processes.

Event logs in HIS suitable for PM do not yet exists or existed at few selected applications as seen in literature review. The challenges that lead to the unavailability of event logs in the healthcare can be summarized into four categories which are as follows i) medical processes are highly dynamic, ii) health care processes are highly complex, iii) health care processes are increasingly multi-disciplinary, and iv) health care processes are ad hoc. These characteristics are often considered to be the cause of complex models in healthcare. Thus, the unavailability of logs suitable for PM for TB can be narrowed down to lack of effort in developing systems that would help improve process analysis. The value of PM in healthcare is based on its capacity to enhance care delivery methods while maintaining high standards for patient care. Process modelling is an important stage in process improvement because it gives a model of the process that can be examined and improved. Therefore, this study aims to develop an event logging system as a solution to the mentioned challenges for the mining of TB treatment pathways.

3. Research methodology and design

Research is the search for knowledge in an objective and systematic ways to proffer solutions to problems. The term "research" therefore refers to a systematic method consisting of enunciating the problem, formulating a hypothesis, collecting the facts or data, analysing the facts, and reaching certain conclusions either in the form of solutions(s) towards the concerned problem or in certain generalisations for some theoretical formulation. Three research paradigms applicable to information systems and computer science are briefly discussed next and the one that is more related to the study is used with justification.

- The *rationalist paradigm*, which was common among theoretical computer scientists, defines the discipline as a branch of mathematics, treats programs on a par with mathematical objects, and seeks certain, a priori knowledge about their ‘correctness’ by means of deductive reasoning.
- The *technocratic paradigm*, promulgated mainly by software engineers, defines computer science as an engineering discipline, treats programs as mere data, and seeks probable, a posteriori knowledge about their reliability empirically using testing suites.
- The *scientific paradigm*, prevalent in artificial intelligence, defines computer science as a natural science, takes programs to be on a par with mental processes, and seeks a priori and a posteriori knowledge about them by combining formal deduction and scientific experimentation.

This study uses the technocratic paradigm as this is motivated by the fact that technocratic paradigm sees computer science as an engineering discipline. This ideology is aligned with the study aim to engineer a solution for the lack of event logging data in healthcare domain using TB as example. This study used the Design Science Research (DSR) research approach, for which several research methodologies have been developed among them is the Systems Development Research Methodology (SDRM). The first significant paper on SDRM was proposed by (Nunamaker Jr, 1990) and consist of five-step research process with relevant research issues at each step, as shown in table 3.

Research Step	Research Issues
Construct a Conceptual Framework	• State a meaningful research question • Investigate the system functionalities and requirements • Understand the system building processes/procedures
Develop a System Architecture	• Develop a unique architecture design for extensibility, modularity,

	etc. • Define functionalities of system components and interrelationships among them
Analyse & Design the System	• Design the database/knowledge base schema and processes to carry out system functions • Develop alternative solutions and choose one solution
Build the (Prototype) System	• Learn about the concepts, framework, and design through the system building process • Gain insight about the problems and the complexity of the system
Observe & Evaluate the System	• Observe the use of the system by case studies and field studies • Evaluate the system by laboratory experiments of field experiments • Develop new theories/models based on the observation and experimentation of the system's usage • Consolidate experiences learned

Table 3: Systems Development Research Methodology

Research design is meant to address how the research questions are answered and the objectives of the research attained by applying a chosen method. Simple said, research design is a blueprint detailing what and how to achieve research aim. An important aspect of research design includes research methodology. While conducting research, researchers may discover weaknesses in the design, changes in study design may lead to alternative methods being selected or may require further design adjustments to the newly selected method. Three research methods were used to achieve the objectives of this research namely:

- *Comprehensive Literature Review*: This method was used to identify existing problems and other related work in that has been done by other researchers on PM and TB software projects as to identify research gaps and future works. The result of this approach is given in section 2.0 to 2.2 above.
- *Design and Implementation:* This method entails the design of HIS for-TB data using Context diagram, Class diagram, use case diagram, ERD. Then, software that solves PM for TB disease and generate events logs to the database were written using the following research tools:
    - Visual Studio Code: also commonly referred to as VS Code, is a source-code editor by Microsoft. This was the editor used in the coding, debugging, and testing of the application.
    - MongoDB: is a source-available cross-platform document-oriented database platform. Classified as a NoSQL database program, MongoDB uses JSON-like documents with optional schemas. MongoDB was used as the database for tables, views, and schemas creation for this application.
    - Node.js: is an open-source server environment. Node.js is a back-end JavaScript runtime environment. Node.js runs on the V8 JavaScript Engine and executes JavaScript code outside a web browser.
    - JavaScript is a programming language that is one of the core technologies of the World Wide Web, alongside HTML and CSS. As of 2022, 98% of websites use JavaScript on the client side for webpage behaviour, often incorporating third-party libraries as CSS.
    - HTML: The Hypertext Markup Language or HTML is the standard markup language for documents designed to be displayed in a web browser. This was used as the structure for the application development.
- *Testing:* The event logging application for TB is tested and debugged until the system performs as expected.

System analysis and design is a step-by-step process for developing high-quality system. The analysis and design of the event logging system that generates logs suitable for PM was created with the aim to provide support to both the healthcare sector and researchers aiming to conduct the mining of healthcare processes of TB. This aspect of the research provides answers to RQ2 and RO2 as outlined in section 1.0. In modelling the system, a Use case diagram is used to show the interactions with the identified actors of the system and their roles, the interaction between actors and system components are modelled using the sequence diagram. Figure 2 shows an overview of the steps adopted in the system analysis and design phase of this research.

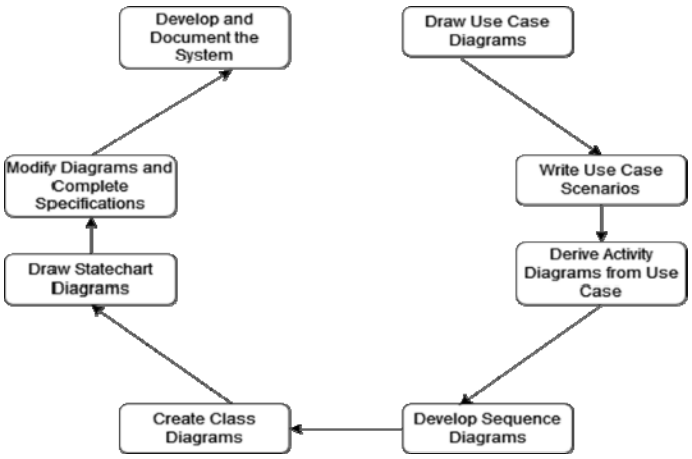


Figure 2: System overview of the System Analysis and Design

3.1 Functional and Non-functional requirement of ELS for TB

3.1.1 Functional

Functional requirements can be defined as requirements that outline what the system does, in particular the functions of the system (ReQtest, 2017). Table 4 shows the functional requirements of the proposed system.

Table 4: Functional requirements of the ELS for TB

Id	Functional requirements	Priority
FR1	The system shall provide user friendly interface for capturing TB patients' data	1
FR2	The system shall store data about healthcare practitioners and TB patients in a database in a form suitable for PM in individual tables in a nonredundant manner.	1
FR3	The system shall provide facility to retrieve tables related to TB cases and merge them into a single csv file with appropriate timestamps for event logging system (ELS) for TB	1
FR4	The system shall allow researchers and healthcare practitioners to download event logs files.	2
FR5	The system shall allow the healthcare practitioners to sign up if they are not yet registered.	2

3.1.2 Non-functional

Non-Functional requirements can be defined as requirements that outline or describe constraints imposed on the system functionalities (ReQtest, 2017). Table 5 shows the functional requirements of the proposed system.

Table 5: Non-functional requirements of the ELS for TB

Id	Non-Functional requirements	Priority
FR6	The system shall be easily and always accessible from the web	1
FR7	The system response time shall be very minimal	1
FR8	The system should operate in a security and conducive environment	2
FR9	The system shall consume minimal resources	3

3.2 System Modelling

System modelling refers to the process of abstracting different components of a system to form a model [28], this can be done graphically using methods like the Unified Modelling Language (UML). System models are used during requirements engineering to clarify how the system operates, additionally they are used to help explain the system requirements to other stakeholders (Silva Filho, 2017). The UML models used in the modelling of the ELS for TB are described and this section is aimed at answering RQ2 and achieving RO2 as presented in section 1.0

3.2.1 Context diagram

Context diagrams are used to depict the system's operational context; essentially the context models illustrate that there are other systems present in the proposed system's environment. This illustration is made without specifying the type of relationships between the systems and its environment (Jaco, 2019). Figure 3 illustrates the context model of the proposed system.

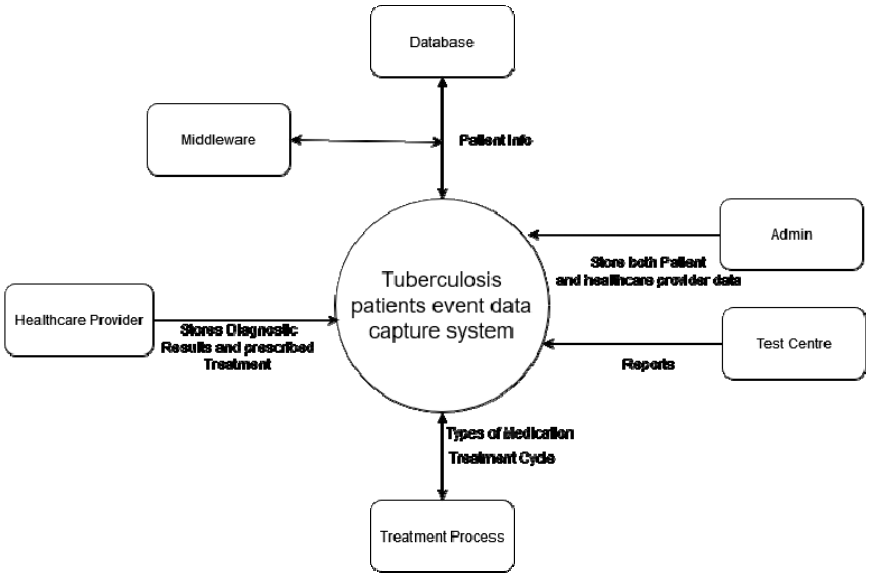


Figure 3: The Context diagram of the ELS for TB

3.2.2 Use case diagram

Figure 4 illustrate the Use case diagram Specifying the behavior of actors, functionalities offered and the interactions of the system with the actors, which are mainly the Patient administrator, and healthcare providers and the Patients which their role of providing information.

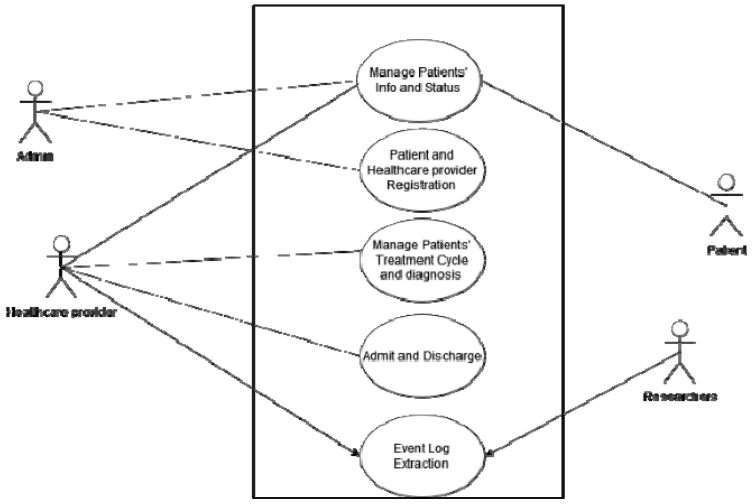


Figure 4: The Use case diagram of the ELS for TB showing functionalities of the system, the system actors, and their interaction with system functionalities.

Table 6: The roles of actor in the ELS for TB

Admin	Healthcare Provider	Researchers	Patient
*Registers both the healthcare providers and patients to the system. *Create or delete information	*Stores patients' health information into the system. *Have privileges to download logs from the system.	*Have privileges to download logs from the system.	*The person who visits the health facility for TB checkup, diagnosis, or treatment.

4 Use case description

Vital Use case in this research is described. The functionality is presented in the form of tables, use case descriptions of each actor functionality are presented in the form of a table.

Table 7: Registering the Healthcare providers

Use Case Name: Registering the Healthcare providers **Use Case ID:** Uc1	
Actors	Healthcare providers and Admin.
Description	The system administrator can register healthcare providers
Trigger	New healthcare provider working in the TB department.
Preconditions	The system is connected to the database.
Postconditions	The database contains data about the registered healthcare provider.
Normal Flow	-The actors open the system and register their personal details. -The actors populated the database with the actor's information - The system assign timestamp to each, data fragment stored in the system
Frequency of Use	Medium
Assumptions	None

Table 8: Extracting Event logs

Use Case Name: Extracting Event logs Use Case ID: Uc2	
Actors	Healthcare providers or Researcher.
Description	The healthcare providers can add data and download logs while researchers can only download logs.
Trigger	Event logs file request.
Preconditions	The system is connected to the database.
Postconditions	The database contains data and logs has been generated.
Normal Flow	-The actors open the system and register their personal details. -The actors populated the database with both the actors and patients' information - The system assign timestamp to each data fragment stored in the system -Actors download *CSV.file* with logs
Frequency of Use	Medium
Assumptions	None

3.2.3 Sequence diagram

The sequence diagram is used primarily to show the interactions between objects in the sequential order that those interactions occur, much like the class diagram (Bell, 2015). Figure 5 illustrates sequence diagram for the proposed system with brief details of how each actor interact with the system in a sequential manner. The diagram in figure 5 demonstrates the sequential events that take place in the healthcare when a patient with TB or one displaying symptoms of TB visits the healthcare, meets the desk officer, where details about the patient is recorded, then the patients proceed to do vital checks recorded by a Nurse, followed by a doctor's diagnosis and capturing of the results into the system, and lastly the patient enters the treatment cycle.

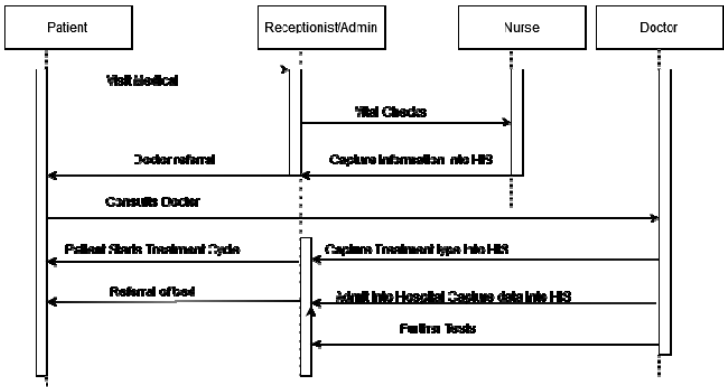


Figure 5: Showing sequence diagram for ELS for TB

3.3 System Design

3.3.1 ELS Architecture

System architecture refers to the overall structure of a system. The architecture of the ELS consists of 3 major components namely, the event logs creation which takes place at the backend and at the database server, capturing of patient's information by the Health Professionals and lastly the internet which provide the connectivity.

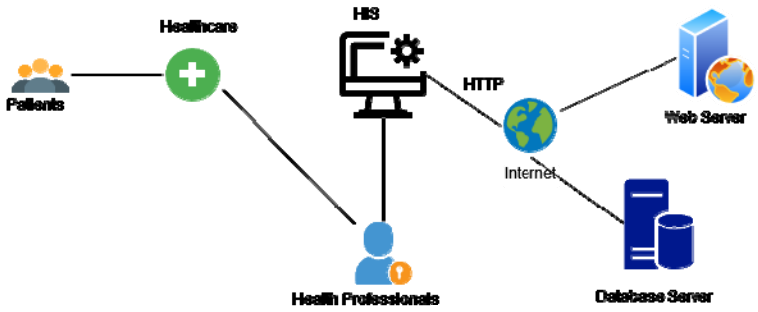


Figure 6: System Architecture

The components shown in figure 6 constitute the nucleus of the ELS in the overall implementation of the system: HIS – refers to the device being utilized to run the HIS in real time; Web server – Responsible for processing all the admin requests; Database server – Processes all the requests pertaining to the storage and retrieval of event logs and MongoDB environment was utilized; Health Professionals – the Medical doctor, nurses or any healthcare employee with the privileges to use the system through their registered Healthcare providers; Application Hosting Device-which requires a desktop, laptop, mobile devices etc connected to the internet.

3.3.2 Event Logging Algorithm Design

Capturing data: The data captured from the healthcare professionals are stored in document format of JSON, BSON or XML. The document is stored in the collections as schema less, hence any number of fields can be added to the database without keeping empty spaces for other document in the collection. There is no need of JOINs as in RDBMS due to embedded documents and arrays (Chauhan, 2017). The popular document databases are MongoDB, CouchDB, Terrastore, ThruDb.

Algorithm: Capturing data
Input: TB patients' data and Health Professionals data.
Output: TB Event logging data.

Steps:

1. Connect the Database to the Application
2. Sign up as the Healthcare provider
3. Login as the admin
4. Capture the information about TB patients' data (Patients ID, Vital Readings, treatment, Diagnosis etc.)

Extracting Event Logs from the Database: mongoexport is a command-line tool that produces a JSON or CSV export of data stored in a MongoDB instance (Tomar, 2019). A button was created for on the user interface that invoke the function that calls backend and save the CSV file on the user's machine.

Algorithm: Extracting Event Logs from the Database
Input: TB patients' data and Health Professionals data.
Output: TB Event logging data.

Steps:

1. Connect the Database to the Application
2. Login as the admin

3. Use either of the option to download either of the csv.file mainly (SignUp.csv, Patients.csv, Diagnosis.csv, etc.).
4. Use the Merge button to merge the table(s) if the study is focused on all the section of the database.

3.3.3 Database

A good database designs should at least include the following: Divides information into subject-based tables to reduce data redundancy, provides access with the information it requires to join the information in the tables together as needed, helps support and ensure the accuracy and integrity of information and accommodates data processing and reporting needs (Harrington, 2016).

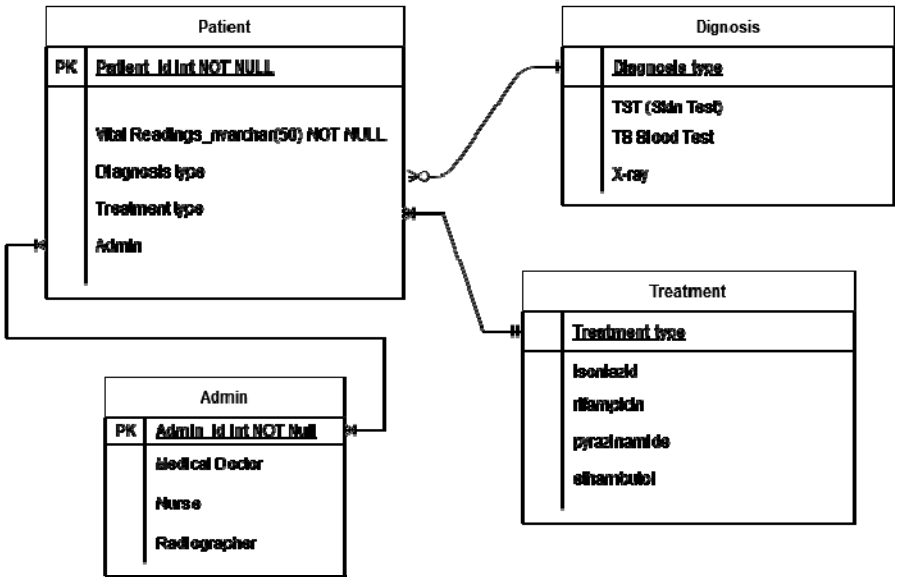


Figure 7: The Entity relationship diagram (ERD) of the ELS of TP

The database that supports the ELS stores the login and registration information of the administrator, and medical related information about the patient. MongoDB was used to implement the ERD shown in figure 7. MongoDB is a document-oriented database program classified as a NoSQL and uses JSON-like documents with optional schemas.

3.3.4 Network of the ELS for TB

The proposed system will utilize the client-server network architecture, which consists of two main stakeholders namely, the client who makes requests for specific resources or services and the server which responds and handles the requests made by the client (Oluwatosin, 2014). To facilitate this process the system will utilize internet connection to access the database hosted on cloud. This is made possible with the use of the database and back-end Servers. Figure 10 illustrates the client-server network architecture utilized by the ELS.

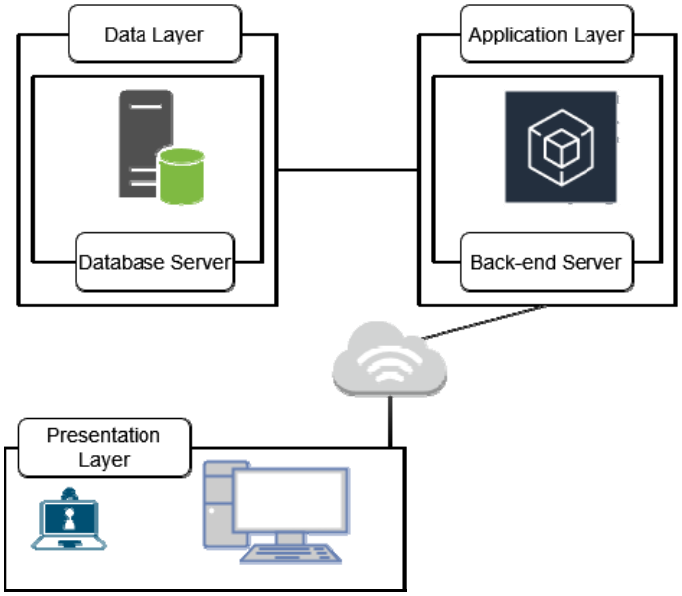


Figure 8: Network architecture of the ELS OF TP.

3.4 System Properties

The operating system and hardware properties of the computer used in developing the ELS of the TP shown in table 9.

Table 9: Hardware properties required the implement ELS for TB

Parameter	Values
Software	OS Name: Microsoft Windows 11 Pro, 64 bits Operating System.
	HD type: Full HD
	Display resolution: 1920 x 1080 pixels
Hardware	Processor: 11th Gen Intel(R) Core (TM) i5-11320H @ 3.20GHz, 2496 Mhz, 4 Core(s), 8 Logical Processor(s)
	RAM: 8.00 GB.
	Hard Disk: 256GB

3.5 ELS Operations

The steps showing how a user will make use of the system is shown in figure 9.

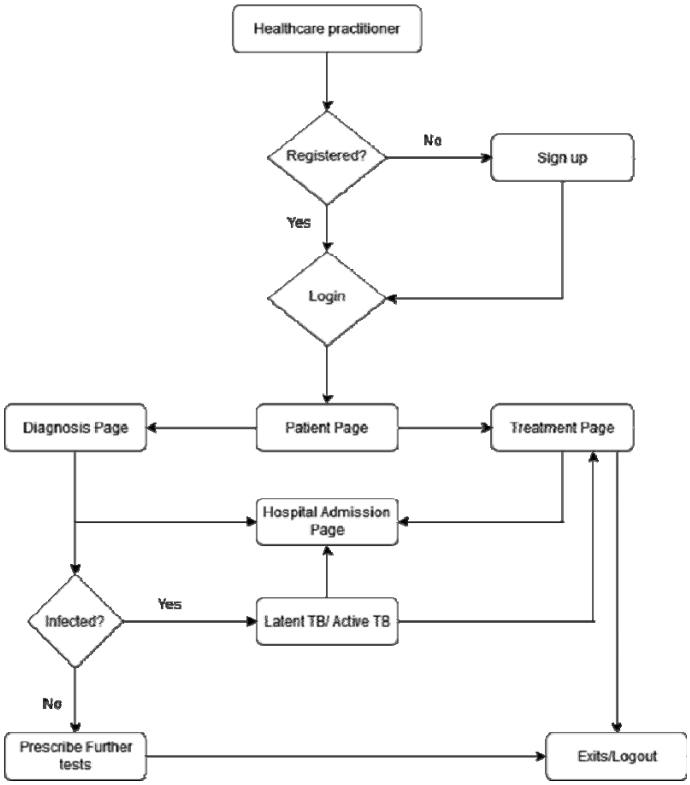


Figure 9: Flow diagram of the ELS for TP

3.6 The ELS for TP prototype Interfaces. This section is aimed at answering RQ3 and achieving RO3 as presented in section 1.0

1. Login page: on the login page, the user (health Practitioner) has two options as shown in figure 10: which are “Sign Up” or “Admin Login”. These buttons are utilised to create a new health practitioner in the ELS or for an existing practitioner to log into the ELS. To login, the user will use their unique ID provided when they were signing into the ELS.

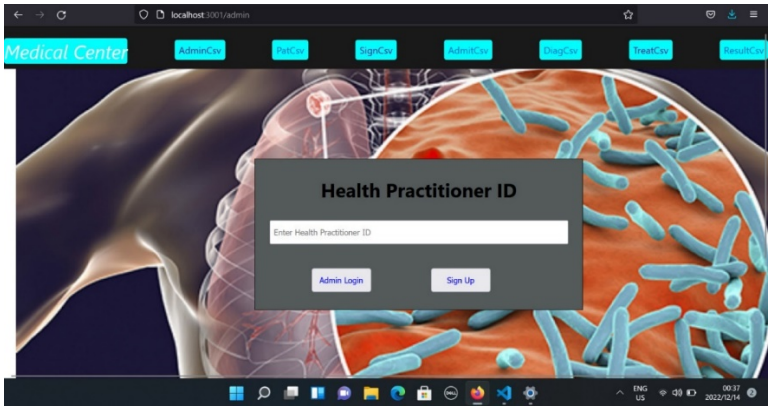


Figure 10: Login page of the ELS for TP

2. Registration details of Health practitioners: To register a practitioner, they should sign up and provide their Health Practitioner ID, Name, Surname and Job Description as shown in figure 11.

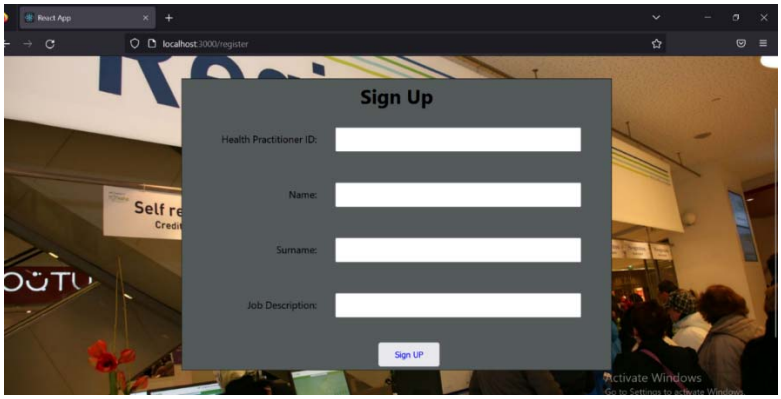


Figure 11: Sign up of health practitioners in the ELS for TP

3. Patient page: The focus of figure 12 is to capture data about patient before an in-depth diagnosis or treatment take place. Then the patient can go for diagnosis or treatment after vital health statistics taken by the doctor or health personnel have been saved into the database.

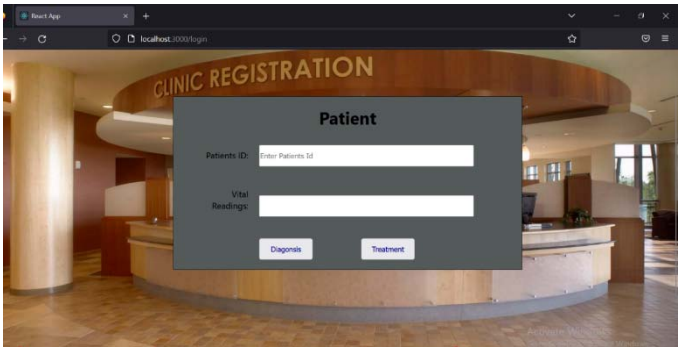


Figure 12: Showing patient page of the system.

4. Diagnosis pages: the diagnosis section is comprised of three pages which are mainly the types of diagnosis see figure 13 (which shows the type method used for testing TB), the diagnosis page see figure 14 (shows whether the patient is infected or not infected) and the results page see figure 15 (shows the type of TB the patient has).

Figure 13: Showing types of diagnosis page of the system.

Figure 14: Showing diagnosis page of the system.

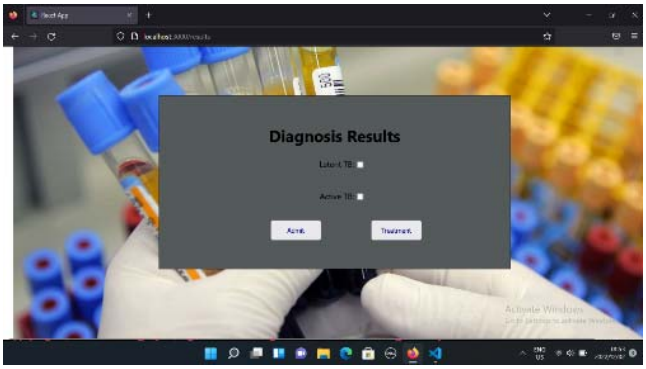


Figure 15: Showing diagnosis results page of the system.

5. Treatment page: This page provides the health professionals with the options to choose the prescribed treatment for the patient and save the information into the database. While the "Admit" button allows the patient to be admission into the health facility as shown in figure 16.

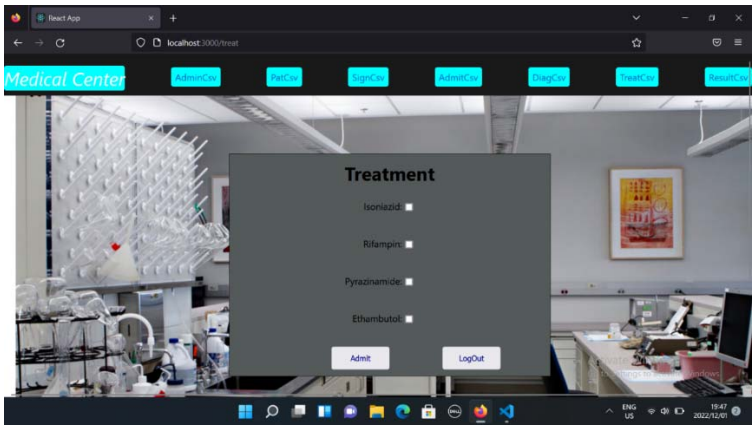


Figure 16: Showing treatment page of the system.

6. Admission page: when the patient is in critical conditions that needs healthcare professionals to monitor the situation closely, the patient is admitted. Figure 17 stores the information about the patients allocated hospital facility.

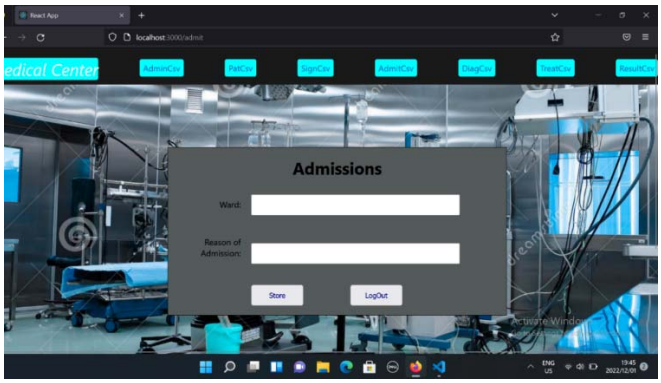


Figure 17: Showing admission page: of the system.

4. Results and Discussion

The results in figures 18 to 20 contains the various logs generated by the ELS in different parts of the TB diagnosis and treatment as highlighted in figures 10 to 17. These logs are generated, processed and stored as .csv files as shown in figures 19 and 20 for the purpose of making data about TB easily available for the purpose of research and analysis of the various path of treatment regine in the fight against the spread of TB. These data are paramount and pivotal to any meaniful breakthrough that can be recorded in understanding the behaviour of health personnel and the manner in which TB is diagnosed and treated including the various drugs reginement admistered to the patient in the entire treatment process. This section is also aimed at answering RQ3 and achieving RO3 as presented in section 1.0

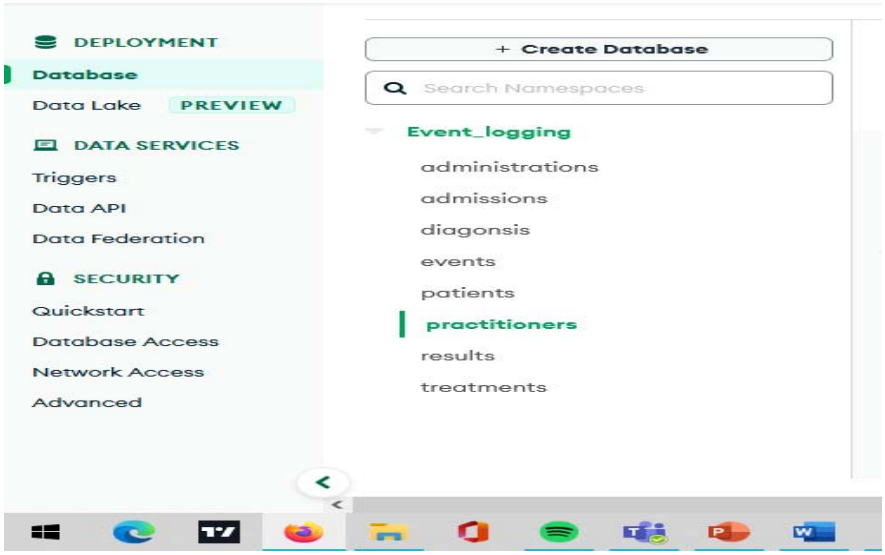


Figure 18: Database with different tables containing data for user interfaces.

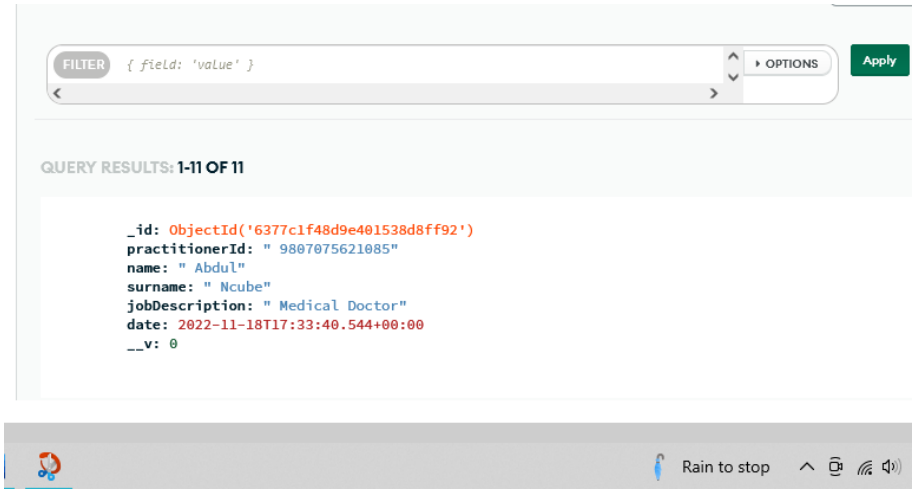
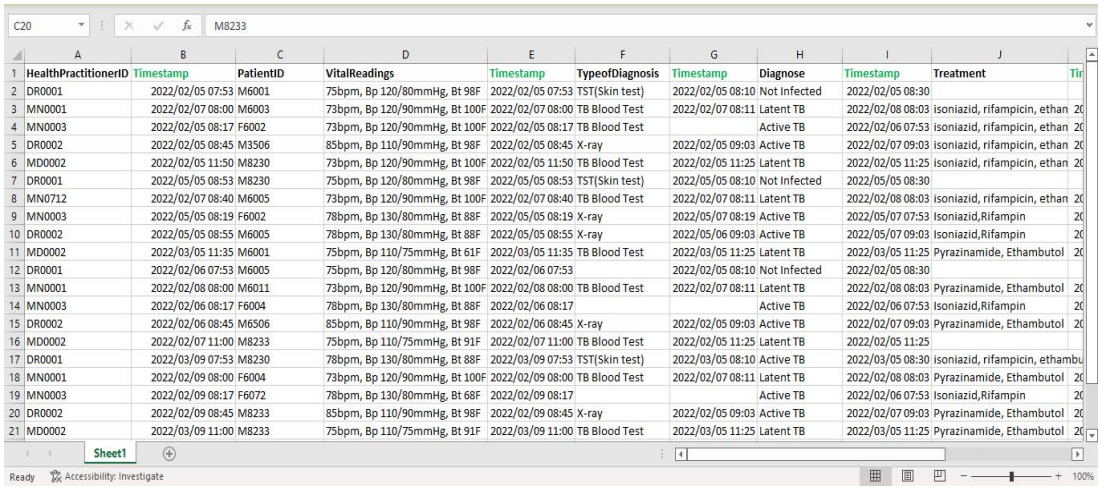


Figure 19: shows 1 of 11 rows in the practitioners table, together with the timestamp allocated by the database



	A	B	C	D	E	F	G	H	I	J
1	HealthPractitionerID	Timestamp	PatientID	VitalReadings	Timestamp	TypeofDiagnosis	Timestamp	Diagnose	Timestamp	Treatment
2	DR0001	2022/02/05 07:53	M6001	75bpm, Bp 120/80mmHg, Bt 98F	2022/02/05 07:53	TST(Skin test)	2022/02/05 08:10	Not Infected	2022/02/05 08:30	
3	MN0001	2022/02/07 08:00	M6003	73bpm, Bp 120/90mmHg, Bt 100F	2022/02/07 08:00	TB Blood Test	2022/02/07 08:11	Latent TB	2022/02/08 08:03	isoniazid, rifampicin, ethan
4	MN0003	2022/02/05 08:17	F6002	73bpm, Bp 120/90mmHg, Bt 100F	2022/02/05 08:17	TB Blood Test		Active TB	2022/02/06 07:53	isoniazid, rifampicin, ethan
5	DR0002	2022/02/05 08:45	M3506	85bpm, Bp 110/90mmHg, Bt 98F	2022/02/05 08:45	X-ray	2022/02/05 09:03	Active TB	2022/02/07 09:03	isoniazid, rifampicin, ethan
6	MD0002	2022/02/05 11:50	M8230	73bpm, Bp 120/90mmHg, Bt 100F	2022/02/05 11:50	TB Blood Test	2022/02/05 11:25	Latent TB	2022/02/05 11:25	isoniazid, rifampicin, ethan
7	DR0001	2022/05/05 08:53	M8230	75bpm, Bp 120/80mmHg, Bt 98F	2022/05/05 08:53	TST(Skin test)	2022/05/05 08:10	Not Infected	2022/05/05 08:30	
8	MN0712	2022/02/07 08:40	M6005	73bpm, Bp 120/90mmHg, Bt 100F	2022/02/07 08:40	TB Blood Test	2022/02/07 08:11	Latent TB	2022/02/08 08:03	isoniazid, rifampicin, ethan
9	MN0003	2022/05/05 08:19	F6002	78bpm, Bp 130/80mmHg, Bt 88F	2022/05/05 08:19	X-ray	2022/05/07 08:19	Active TB	2022/05/07 07:53	Isoniazid,Rifampin
10	DR0002	2022/05/05 08:55	M6005	78bpm, Bp 130/80mmHg, Bt 88F	2022/05/05 08:55	X-ray	2022/05/06 09:03	Active TB	2022/05/07 09:03	Isoniazid,Rifampin
11	MD0002	2022/03/05 11:35	M6001	75bpm, Bp 110/75mmHg, Bt 61F	2022/03/05 11:35	TB Blood Test	2022/03/05 11:25	Latent TB	2022/03/05 11:25	Pyrazinamide, Ethambutol
12	DR0001	2022/02/06 07:53	M6005	75bpm, Bp 120/80mmHg, Bt 98F	2022/02/06 07:53		2022/02/05 08:10	Not Infected	2022/02/05 08:30	
13	MN0001	2022/02/08 08:00	M6011	73bpm, Bp 120/90mmHg, Bt 100F	2022/02/08 08:00	TB Blood Test	2022/02/07 08:11	Latent TB	2022/02/08 08:03	Pyrazinamide, Ethambutol
14	MN0003	2022/02/06 08:17	F6004	78bpm, Bp 130/80mmHg, Bt 88F	2022/02/06 08:17			Active TB	2022/02/06 07:53	Isoniazid,Rifampin
15	DR0002	2022/02/06 08:45	M6506	85bpm, Bp 110/90mmHg, Bt 98F	2022/02/06 08:45	X-ray	2022/02/05 09:03	Active TB	2022/02/07 09:03	Pyrazinamide, Ethambutol
16	MD0002	2022/02/07 11:00	M8233	75bpm, Bp 110/75mmHg, Bt 91F	2022/02/07 11:00	TB Blood Test	2022/02/05 11:25	Latent TB	2022/02/05 11:25	
17	DR0001	2022/03/09 07:53	M8230	78bpm, Bp 130/80mmHg, Bt 88F	2022/03/09 07:53	TST(Skin test)	2022/03/05 08:10	Active TB	2022/03/05 08:30	isoniazid, rifampicin, ethambu
18	MN0001	2022/02/09 08:00	F6004	73bpm, Bp 120/90mmHg, Bt 100F	2022/02/09 08:00	TB Blood Test	2022/02/07 08:11	Latent TB	2022/02/08 08:03	Pyrazinamide, Ethambutol
19	MN0003	2022/02/09 08:17	F6072	78bpm, Bp 130/80mmHg, Bt 68F	2022/02/09 08:17			Active TB	2022/02/06 07:53	Isoniazid,Rifampin
20	DR0002	2022/02/09 08:45	M8233	85bpm, Bp 110/90mmHg, Bt 98F	2022/02/09 08:45	X-ray	2022/02/05 09:03	Active TB	2022/02/07 09:03	Pyrazinamide, Ethambutol
21	MD0002	2022/03/09 11:00	M8233	75bpm, Bp 110/75mmHg, Bt 91F	2022/03/09 11:00	TB Blood Test	2022/03/05 11:25	Latent TB	2022/03/05 11:25	Pyrazinamide, Ethambutol

Figure 20: the *Csv.file* downloaded from the ELS database.

4.1 Evaluation of ELS with existing system

The criteria for the evaluation of the system consists of assessing the similarities and dissimilarities between the existing system and the one developed in this study, this is illustrated in table 10.

Table 10: Comparative of developed ELS and existing similar system in literature

	Developed ELS	**Related System from Literature (Shah, 2008)**
Aim	Capture and store event log data of TB patients during their interaction with the healthcare Practitioners/facility	Logging user-interface events generated in the Microsoft Internet Explorer during a user's interaction with the pages of a website.
Storage	Timestamps are recorded in a cloud database (MongoDB) each time data is recorded	The event logs are dispatched through HTTP to the server where they are stored as text files.
PM purpose	event data is to be captured for purpose of performing PM to acquire more knowledge about the TB treatment.	event data is to be captured for usability evaluation of a website.

From the data given in table 10, it is evident that there is no system that have been developed with the goal of providing event logs for TB that will assist to understand the processes it's diagnosis and treatment which places our ELS at the top position. This conclusion is substantiated by the dissimilarities between the existing system and our ELS as the related system from literature discussed in this table is the only one in literature that is most related our ELS. Lastly, ELS is evaluated for useability, confidentiality, and accuracy of timestamps and response time.

- User ability: the naming of buttons, web pages headings, dialog boxes, menus and text fields defines the functionalities which makes the user interface of ELS user-friendly.
- Confidentiality: since the systems main purpose is to create event logs suitable for PM, the system records patients IDs only for the purpose of uniqueness and ethical related reasons.
- Accuracy of timestamps: this is the most crucial aspect of the ELS. It assigns system timestamp to every data fragment captured into the system. The timestamps are important for PM.
- Response time: refers to the speed with which the system capture data into the database and the speed at which the system can retrieve the Csv.file from the database. The was achieved using an application from Apache called Jmeter (https://jmeter.apache.org/). It takes about 389 milliseconds and 95000 milliseconds to store and download respectively the Csv.file on the average.

5. Conclusion

The need to develop software to bridge the gap in data available for the study of TB diagnosis and treatment pathways has necessitated the design ELS. The study aim, research questions and research objectives were given in section 1 which succinctly stated, seeks to develop a software for tracking various activities that occurs from when a person visits a health center for TB checkup, diagnosis, and treatment. The data generated throughout this process is tracked by the ELS. Section 2 reviewed articles related to process mining and tuberculosis in other to establish the research gap that necessitated the paper. Section 3 discusses the research paradigm, methodology and methods used in developing the ELS. Section 4 discusses the various UML diagrams utilized in developing our ELS. Section 5 shows the results of the developed system in form of dialog boxes, menu and functions the system execute. It is hope that when the ELS is put into use, data related to TB can be captured at source which will facilitate the study of various pathways for TB regiment, reduce lack of conformity and performance bottlenecks plaquing TB free society.

Future work

In future, we shall populate our ELS with data from TB treatment centres in Mahikeng, South Africa (as most of the TB testing and treatment centres are manual and records are kept on physical files) and then use PM software such as Disco, ProM, IBM Process Mining, Celonis, UiPath Business Automation Platform, Aris Process Mining etc to generate Process models and then study different performance and conformance metrices on actual TB data.

Conflicts of interest- The authors have no conflicts of interest to declare.

Reference

[1] Bell, D. (2006). UM basics: The sequence diagram. Available at: https://www.semanticscholar.org/paper/UML%E2%80%99s-Sequence-Diagram-Bell/44cbfe0da0e7f6f57c384d3c1605ba9952051c38
[2] Bose, R.J.C.; Mans, R.S.; Van der Aalst, W.M. (2013) *Wanna improve process mining results*? DOI: 10.1109/CIDM.2013.6597227
[3] Carmona, J., Van Dongen, B., Solti, A., and Weidlich, M. (2018). *Conformance checking*, Switzerland. Springer.
[4] Chauhan, D.; Bansal, K. (2017). Using the advantages of NOSQL: a case study on MongoDB, International Journal on Recent and Innovation Trends in Computing and Communication, **5**(2), pp. 90-93.
[5] De Oliveira, H.; Augusto, V.; Jouaneton, B.; Lamarsalle, L.; Prodel, M.; Xie, X. (2020). Automatic and explainable labeling of medical event logs with autoencoding, IEEE Journal of Biomedical and Health Informatics, **24**(11), pp. 3076-3084.
[6] ReQtest (2012). Functional vs non functional requirements (2012), Available at: https://reqtest.com/requirements-blog/functional-vs-non-functional-requirements/
[7] Escombe, A.R.; Moore, D.A.J.; Gilman, R.H.; Navincopa, M.; Ticona, E.; Mitchell, B.; Noakes, C.; Martínez, C.; Sheen, P.; Ramirez, R. (2009). Upper-room ultraviolet light and negative air ionization to prevent tuberculosis transmission, PLoS medicine, **6**(3), pp. e1000043.
[8] Fernández-Llatas, C.; Meneu, T.; Traver, V.; Benedi, J.-M. (2013). Applying evidence-based medicine in telehealth: an interactive pattern recognition approximation, International journal of environmental research and public health, **10**(11), pp. 5671-5682.
[9] Gambhir, S.; Ravina, M.; Rangan, K.; Dixit, M.; Barai, S.; and Bomanji, J. (2017). Imaging in extrapulmonary tuberculosis. International Journal of Infectious Diseases, **56**, pp. 237-247.
[10] Gatta, R.; Vallati, M.; Lenkowicz, J.; Casá, C.; Cellini, F.; Damiani, A.; Valentini, V. (2018). *A framework for event log generation and knowledge representation for process mining in healthcare*, DOI: 10.1109/ICTAI.2018.00103
[11] Guo, X.; Shen, H.; Liu, S.; Xie, N.; Yang, Y.; Jin, J. (2021). Predicting the trend of infectious diseases using grey self-memory system model: a case study of the incidence of tuberculosis, Public Health, **201**, pp. 108-114
[12] Harrington, J.L. (2016). *Relational database design and implementation*, Morgan Kaufmann: Elsevier.
[13] Herzlinger, R.E. (2006). Why innovation in health care is so hard, Harvard business review, **84**(5), pp. 58
[14] Homayounfar, P. (2012). Process mining challenges in hospital information systems, Available at: https://www.researchgate.net/publication/261119671_Process_mining_challenges_in_hospital_information_systems
[15] Jacob, N.M. (2019). A Review of Mobile Application Development in the Agile Software Development Environment', Global Journal of Computer Science and Technology, **19**(1).
[16] Kaymak, U.; Mans, R.; Van de Steeg, T.; Dierks, M. (2012). *On process mining in health care*, DOI: 10.1109/ICSMC.2012.6378009
[17] Keeley, C.; Jimenez, J.; Jackson, H.; Boudourakis, L.; Salway, R.J.; Cineas, N.; Villanueva, Y.; Bell, D.; Wallach, A.B.; Boyle Schwartz, D. (2020). Staffing Up For The Surge: Expanding The New York City Public Hospital Workforce During The COVID-19 Pandemic: Article describes how New York City's public health care system rapidly expanded capacity across 11 acute-care hospitals and three new field hospitals to meet the challenges of the COVID-19 Pandemic, Health Affairs, **39**(8), pp. 1426-1430.
[18] Kim, J.M.; Kang, J.G.; Kim, S.; Cheon, J.H. (2021). Deep-learning system for real-time differentiation between Crohn's disease, intestinal Behçet's disease, and intestinal tuberculosis, Journal of Gastroenterology and Hepatology, **36**(8), pp. 2141-2148.
[19] Konduri, N.; Bastos, L.G.V.; Sawyer, K.; Reciolino, L.F.A. (2017). User experience analysis of an eHealth system for tuberculosis in resource-constrained settings: a nine-country comparison, International journal of medical informatics, **102**, pp. 118-129.
[20] Li, X.; Zhou Y.; Du, P.; Lang, G.; Xu, M.; Wu, W. (2021). A deep learning system that generates quantitative CT reports for diagnosing pulmonary Tuberculosis, Applied Intelligence, **51**(6), pp. 4082-4093.
[21] Martin, N.; De Weerdt, J.; Fernández-Llatas, C.; Gal, A.; Gatta, R.; Ibáñez, G.; Johnson, O.; Mannhardt, F.; Marco-Ruiz, L.; Mertens, S. (2020). Recommendations for enhancing the usability and understandability of process mining in healthcare, Artificial Intelligence in Medicine, **109**, pp. 101962.
[22] Muyoyeta, M.; Kasese, N.C.; Milimo, D.; Mushanga, I.; Ndhlovu, M.; Kapata, N.; Moyo-Chilufya, M.; Ayles, H. (2017). Digital CXR with computer aided diagnosis versus symptom screen to define presumptive tuberculosis among household contacts and impact on tuberculosis diagnosis, BMC infectious diseases, **17**(1), pp. 1-8.

[23] Nunamaker Jr JF, Chen M, Purdin TD. (1990). Systems development in information systems research, Journal of management information systems. **7**(3), pp. 89-106.
[24] Oluwatosin, H.S. (2014). Client-server model, IOSR Journal of Computer Engineering, **16**(1), pp. 67-71.
[25] Pereira, G.B.; Santos, E.A.P.; Maceno, M.M.C. (2020). Process mining project methodology in healthcare: a case study in a tertiary hospital, Network Modeling Analysis in Health Informatics and Bioinformatics, **9**(1), pp. 1-14.
[26] Ridandari, F.; Panjaitan, A.C. (2019). Expert System to Diagnose Extra Lung Tuberculosis Using Bayes Theorem: Expert System to Diagnose Extra Lung Tuberculosis Using Bayes Theorem, Jurnal Mantik, **3**(3), pp. 34-39.
[27] Rojas, E.; Munoz-Gama, J.; Sepúlveda, M.; Capurro, D. (2016). Process mining in healthcare: A literature review, Journal of biomedical informatics, **61**, pp. 224-236.
[28] Shah, I.; Toaimy, L.; Jawed, M.(2008). RWELS: A remote web event logging system, Journal of King Saud University-Computer and Information Sciences, **20**, pp. 1-11.
[29] Silva Filho, M.C.; Oliveira, R.L.; Monteiro, C.C.; Inácio, P.R.; Freire, M.M. (2017). CloudSim Plus: A cloud computing simulation framework pursuing software engineering principles for improved modularity, extensibility and correctness, DOI: 10.23919/INM.2017.7987304.
[30] Suárez, I.; Fünger, S.M.; Kröger, S.; Rademacher, J.; Fätkenheuer, G.; Rybniker, J. (2018). The diagnosis and treatment of tuberculosis, Deutsches Aerzteblatt International, **116**, pp. 729-35.
[31] Suriadi, S.; Wynn, M.T.; Ouyang, C.; ter Hofstede, A.H.; Dijk, N.J.V. (2013). *Understanding process behaviours in a large insurance company in Australia: A case study*, DOI: 10.1007/978-3-642-38709-8_29.
[32] Suriadi, S.; Andrews, R.; ter Hofstede, A.H.; Wynn, M.T. (2017). Event log imperfection patterns for process mining: Towards a systematic approach to cleaning event logs, Information systems, **64**, pp. 132-150.
[33] Tomar, D.; Bhati, J.P.; Tomar, P.; Kaur, G. (2018). *Migration of healthcare relational database to NoSQL cloud database for healthcare analytics and management, Healthcare Data Analytics and Management*, DOI: 10.1016/B978-0-12-815368-0.00002-6
[34] Torres, C.; Fried, J.C.; Manjunath, B. (2018). Healthcare event and activity logging, IEEE Journal of Translational Engineering in Health and Medicine, **6**, pp. 1-12.
[35] Timimi, H.; Falzon, D.; Glaziou, P.; Sismanidis, C.; Floyd, K. (2012). WHO guidance on electronic systems to manage data for tuberculosis care and control, Journal of the American Medical Informatics Association, **19**(6), pp. 939-941.

Authors Profile

Prof. Esiefarienrhe Michael Bukohwo is an Associate Professor of Computer Science and Information systems from the North-West University, South Africa. He holds a BSc (Computer Science), University of Benin, MSc (Computer Science), ATBU, PhD (Computer Science) Modibbo Adama University of Technology, Yola. He is an Oracle certified Professional, member of IEEE, ACM, IITPSA and CPN. He has been teaching undergraduate and postgraduate students for over 20 years. He has successfully graduated over 12 PhD and 35 MSc students in Computer Science and Information systems. His area of research is Business intelligence systems design, risk analysis for business enterprises and business project management. He has published over 50 articles in high impact journals and has presented papers in over 20 international peer review conferences.

Mr. Abdul Ncube holds a BSc Hons in Computer Science from the North-West University, South Africa and presently enrolled for MSc in Computer Science at the North-West University. His research interest is Information systems for Healthcare. He is also a Technology consultant where he is responsible for taking a strategic look on how technology can help improve organizations, which includes architecting and realizing exceptional value from technology. He works mainly with Technology, Media and Telecommunications (TMT) Industries as well as the Healthcare Sector.