E-Health Telecommunication Systems and Networks, 2023, 12, 15-33
https://www.scirp.org/journal/etsn
ISSN Online: 2167-9525
ISSN Print: 2167-9517

# Automating the Competence Matrix of a Quality Control Laboratory with the Relational Database Management Framework—A Case Study with the Analytical Chemistry Unit of a National Regulatory Agency

**Marcel Okezue*, Eric Dietz**

Computer and Information Technology, Purdue University, West Lafayette, USA
Email: *okezue@purdue.edu

**How to cite this paper:** Okezue, M. and Dietz, E. (2023) Automating the Competence Matrix of a Quality Control Laboratory with the Relational Database Management Framework—A Case Study with the Analytical Chemistry Unit of a National Regulatory Agency. *E-Health Telecommunication Systems and Networks*, **12**, 15-33.
https://doi.org/10.4236/etsn.2023.122002

**Received:** June 1, 2023
**Accepted:** June 27, 2023
**Published:** June 30, 2023



## Abstract

This study analyzed the concept of time efficiency in the data management process associated with the personnel training and competence assessments in one of the quality control (QC) laboratories of Nigeria's Foods and Drugs Authority (NAFDAC). The laboratory administrators were burdened with a lot of mental and paper-based record keeping because the personnel training's data were managed manually, hence not efficiently processed. The Excel spreadsheet provided by a Purdue doctoral dissertation as a remedial to this challenge was found to be deficient in handling operations in database tables, and therefore did not appropriately address the inefficiencies. **Purpose:** This study aimed to reduce the time it essentially takes to generate, obtain, manipulate, exchange, and securely store data that are associated with personnel competence training and assessments. **Method:** The study developed a software system that was integrated with a relational database management system (RDBMS) to improve manual/Excel-based data management procedures. To validate the efficiency of the software the mean operational times in using the Excel-based format were compared with that of the "New" software system. The data were obtained by performing four predefined core tasks for five hypothetical subjects using Excel and the "New" system (the model system) respectively. **Results:** It was verified that the average time to accomplish the specified tasks using the "New" system (37.08 seconds) was significantly ($p = 0.00191$, $\alpha = 0.05$) lower than the time measurements for the Excel system (77.39 seconds) in the ANACHEM laboratory. The RDBMS-based "New" system provided operational (time) efficiency in the personnel training and com-

petence assessment process in the QC laboratory and reduced human errors.

## Keywords

Data Management, RDBMS, NAFDAC, Personnel Assessment, SOP, Proficiency

## 1. Introduction

Many national regulatory agencies (NRAs) monitor the manufacturing, testing, and distribution of medicines to conform to international standards as part of their core functions [1]. These NRAs' Quality Control (QC) laboratories are facilities that conduct tests on medicines and generate reports used to make administrative decisions on pharmaceutical products that are utilized within countries. The oversight functions of an NRA depend on the competencies of the laboratory personnel who are exposed to several pieces of training that equip them with the skills requisite for their various job roles. These functions are driven by the competence matrix, which appropriately aligns and ties personnel skillsets, training, and competence assessment statuses to job roles, in form of a data management mechanism [2]. Institutions have challenges maintaining a competency matrix to map core functions for their personnel especially when they depend on manual systems which are associated with various inefficiencies [3] [4].

The National Agency for Foods and Drugs Administration and Control (NAFDAC), Nigeria's NRA has several laboratories that carry out QC functions for drug products used within the country [5]. The analytical chemistry (ANACHEM) unit is one of the key departments in NAFDAC laboratories saddled with the QC core function. The unit comprises several cadres of scientists who undergo different pieces of training to equip them with the skills needed for their job roles. A competence matrix provides comprehensive information on the training and competency assessment status of all personnel. It also links job roles to personnel's current levels of proficiencies attained after satisfactory evaluations of designated competency assessments. It, therefore, forms a configuration that supports personnel data management on training and competence. However, human errors were inherent in managing the personnel training records manually and led to inefficiency in tracing personnel proficiency statuses in the drug laboratory, in addition to process pace issues [6]. In general, the skills required for laboratory QC analysts are broadly categorized as functional, managerial, and behavioral [2]. Some studies have reported variations in the proposed timelines required for QC analysts to acquire these competencies [6] [7] [8]. This is understandable given that a personnel's previous work experience, occupation, skill, educational background, etc., can affect the ease with which they can successfully pass a competency assessment.

However, using a manual process to perform the required data management

operations in NAFDAC's ANACHEM is a complex task which involves determining the sequence of personnel training needs, pooling training records, and keeping timelines of training activities. These requirements pose challenges related to mental operations, human errors and delays in conducting events. Consequently, the personnel competence training and assessments are conducted slowly and inefficiently [6]. This study used relational database management system (RDBMS) to improve the time inefficiencies associated with ANACHEM data management process. The RDBMS described in this study can be adapted and customized for other potential users, similar to Sohne's recommendations that a system he created for automating revenue collection can be adapted by other communities with low infrastructures [8].

The problems with paper-based records have been extensively documented in some research reviewed. Redundancy of hospital records and inadequate information from data structure were identified as major challenges in a study that elucidated the benefits of computerizing the documentation procedures in some health-related facilities [9] [10]. To address the challenges identified with the manual operations, a prior effort developed an Excel-based system to manage personnel competence mapping for the ANACHEM unit [11]. Although the Excel-based competence matrix supported creating a list of the pieces of training and competence assessment tests and generated the duration (hours) required for each piece, it did not adequately mitigate those challenges. Like every other spreadsheet package, Excel does not essentially handle database management operations efficiently. For example, Melissa in his article pointed out some of the shortcomings in using Excel to include performance slowness, no support for data integrity, and defect in handling relational database management operations [12]. Furthermore, Gupta delineated the deficiencies of using such an Excel-based process in handling database tables [13]. The flaws observed by Gupta imply that Excel tables are not logically linked, and data query operations are difficult. Nevertheless, according to Azam, Powell, *et al.*, and Rahman *et al.*, data queries among other operations, are essential in effective management of the information for personnel training and competence assessments [14] [15] [16].

Therefore, this paper provides an automation model that improves on the sluggishness and human error problem type that characterized the manual and Excel-based data management processes in the personnel competence matrix. The model aligns with the World Health Organization (WHO) recommendation to regulatory agencies in relation to automating their manual procedures for training and competence assessment data, to achieve international standard practices [17]. NRAs use the WHO guidelines to accomplish an essential strategy to deliver internationally recognized drug testing, inspection, and assessment of Good Manufacturing Practices [5]. Other research automated or computerized the business processes of health agencies to increase profits and reduce costs. A study provided details about the computerization of an erstwhile manual process and evaluated various benefits within a hospital [18]. Another attempt at automating erstwhile manual systems were earlier published in a low-income re-

source country [New Guinea]. The country computerized the processes in its health management systems as a means of improving healthcare delivery services [19].

### Purpose of Study

This study addressed the time wastage associated with the manual or Excel-based data management processes in personnel skills training and acquisition in the ANACHEM laboratory. It also determined how the existing Excel-based system affects the time demand in administering personnel data management in the competence matrix. It compared the existing system to the automated-based model to determine if the time for the Excel-based operation demand was substantially high. Therefore, it generated the study research question: to what extent did computerization of data management operations reduce the time demand for administering the personnel competence training and assessments in the ANACHEM laboratory?

Aside determining time demands for conducting personnel competence training and assessment, other deliverables for this study were that the computerized competence matrix system should be able to: register and keep track of personnel details, record the sequence of exposures to SOPs and OJTs for all cadres of personnel. It should also track date stamps for the start and completion of individual SOP per personnel, provide the time-point for assessing personnel competence, and access personnel proficiency status on SOPs, OJTs, and competence assessments.

## 2. Fundamental Principles

The NAFDAC's ANACHEM laboratory competence matrix is a framework that maps each personnel's skills and job roles to a definite proficiency level. Each proficiency level is associated with defined weight/credit (in hours) for standard operating procedures (SOPs) training, on-the-job tasks (OJTs) and competence assessment tests. Table 1 illustrates a matrix that maps personnel proficiency levels to their SOP trainings.

A status (quantitatively denotes a weight) is a Boolean declaration that states whether the staff is competent or not with respect to a particular SOP when a data query is executed. The competency matrix specified different minimum proficiency levels for all categories of personnel [6]. The various job roles, proficiency levels, and the number of hours needed to attain the recommended levels

Table 1. Matrix representation of personnel vs. number of available SOPs.

| SOP / Employee | SOP_1 | SOP_2 | SOP_n |
|---|---|---|---|
| Staff_1 | status | status | status |
| Staff_2 | status | status | status |
| Staff_n | status | status | status |

are listed in Table 2. The number of weights each employee (or staff) accumulates determines the achieved proficiency level in that department. The new (automated) competence matrix developed by this study utilized personnel demographics, job roles and corresponding minimum proficiency levels to create database tables, and subsequently track and update the matrix. The study categorized the new matrix into two broad functionalities: staff registration and SOP administration (or staff training).

There are four levels of competency assessments for the ANACHEM unit. The weight allocated to an increasing level of assessment is higher than the previous lower competency tests. For example, competency level 1 assessment for determining accuracy in gravimetric measurements weighs lower than Level 2, etc. The rationale is that higher levels of skills and SOP training are required as personnel progresses in the QC activities over time. Some examples of weight hours associated with some SOP and competency assessments are listed in Table 3.

## The Automated-Based Model Structure

RDBMS was used as a framework to automate personnel records management as a structured database system in this study. This championed the computerization of the data management operations to enhance the operational pace and efficiency. Computer-based data management is a concept that indicates a database management prototype to integrate DBMS for efficient handling of data storage, retrieval, manipulation, and security. Computerization of some processes reported where the use of an RDBMS provided an interface between "users and application" and the database. Such studies highlighted using RDBMS as a tool in administrative functions for managing data storage, access, and performance both in the health sector, and other related fields [20]-[28]. It was identified from these reviews that computerizing processes with well-structured RDBMS was crucial in data management operations. Unfortunately, ANACHEM's Excel-based spreadsheet did not apply any relational database management structure. However, this study addressed this limitation by extensively making use of RDBMS structure in the "New" system that it developed.

Table 2. Hours of training requirements for attaining the minimum proficiency levels for each role in the analytical chemistry unit.

| Job Role | Minimum Proficiency Level | Hours of Training Needed (Weight) to Achieve Proficiency Level |
|---|---|---|
| Guest | Observer | 0 - 30 |
| Temp Staff (Interns, etc.) | Beginner | 30 - 100 |
| Laboratory Attendants | Beginner | 30 - 100 |
| Analysts | Skilled | 300 - 400 |
| Head of Unit | Expert | >400 |
| Head of Lab | Learner | 100 - 300 |

**Table 3.** Rationale for weights allocated to some SOPs and competency assessments in determining cumulative hours of training gained by a QC lab personnel.

| SOP (title) | Hours of Training allotted to SOP (Weight) | Justification |
|---|---|---|
| Training of New Staff in Analytical Chemistry Laboratory | 10 | Required for all personnel joining the unit |
| Purified Water Handling and Management | 5 | Only relevant to a specific activity |
| Determination of loss of Drying (LOD) | 5 | Routine SOP for all analysts and temporary staff. |
| Cleaning Laboratory Glassware and Pipettes | 15 | Key SOP for lab attendants |
| Level 1 Competency (Gravimetric Measurements) | 70 | Personnel must have obtained training on at least 4 relevant SOPs before they qualify to take this assessment. Also, it is a Go/No-go point assessment. |
| Level 2 Competency | 100 | Combines 4 assessments for different procedures-pH, LOD, Karl-Fischer, and Melting Point. |

ANACHEM's Excel spreadsheet for competence matrix did not adequately address the sluggishness of its personnel data management system. Some studies documented using automated-based system to explore time reduction as a means of improving the management of an erstwhile slower system. Daphine used visual studio 2005 to improve the efficiencies of the management system in a school library. Computerizing the manual process allowed the Liberians to add users and books into the system, simplified search from the system database, and managed borrowing from the library resources more efficiently [29]. Other studies demonstrated how RDBM framework improved the efficiencies of handling data; Albalkhi used the framework to increase the time efficiency of querying large heterogeneous datasets in his study core design [30]. Another study elaborated on how access control was applied to secure medical records. This study implemented the mechanism using a relational database system. Employing ANSI standard SQL statements the project created a set of access-control rules encoded as relational table rows. The mechanism determines via a single SQL query if a particular user is authorized to perform a specific operation on a definite data object [31].

Essentially, one of the main objectives of this research is to use RDBMS-based platform to improve the turnaround time to perform the task of tracking personnel competency in the Nigerian FDA lab, this is similar to the efficiencies described by some researchers who used RDBM to improve time efficiencies of operations within healthcare and other systems [32] [33] [34].

## 3. Methods

### 3.1. Research Design

The methods for implementing the manual and Excel-based competence matrix were reviewed as the backbone of this study. A model system was integrated with an RDBMS structure to automate the data management of the personnel training and competence assessment operation in the laboratory. This structure enables all required fields and attributes for the personnel demographics, job details, training, and competence statuses to be effectively incorporated into well-connected database tables and schema. This study designed two process methods (staff registration and staff assessment modules) using SQL language to mobilize the RDBMS structure. Specifically, MySQL database software was employed to develop the "New" system because it is an open-source database tool that centers on the specification of standard SQL [35] [36].

### 3.2. Model System Design

A model system was designed as data management software to simplify the manual/Excel-based process-induced difficulty in managing the personnel competence matrix in the laboratory. Python programming language and structured query language (SQL) were employed to design the "New" system that adopted the RDBMS framework. In creating the "New" system two methodical steps were covered, namely design criteria and testing criteria. At the design criteria stage, the process flow for the manual/Excel procedures and "New" system were evaluated and validated to determine the elements required to create the backend design of the RDBMS-structured system. In designing the frontend, the backend design of the framework was integrated into the system and the process flow validated. Flowcharts, algorithms, and source codes generations were computational problem-solving tools engaged both at the frontend and backend stages. At the testing criteria level, all flowcharts and program source codes developed to generate results (inference/decision) from interoperability of the designs were evaluated and validated. The designs are illustrated in **Figure 1**.

Key data objects employed to design the desired relational database structure include tables for personnel's demographics, departments, job roles, and SOPs/OJTs. RDBMS framework was utilized to interconnect the tables for designing the "New" system, which is the automated personnel competence training and assessments data management format of the competence matrix. This study setup and configured MySQL software in a Microsoft Windows system to enhance the tasks integration. In operating the "New" system that this study developed, the ANACHEM administrator takes a decision on two tasks. Either to conduct a data management operation (on personnel enrollment using the Staff Registration module) or employee SOP/OJT or competence assessment (using the Staff Training/Assessment module). **Figure 2** illustrates the process flow part for the "New" system.

However, either option will lead the operational control to flow down through

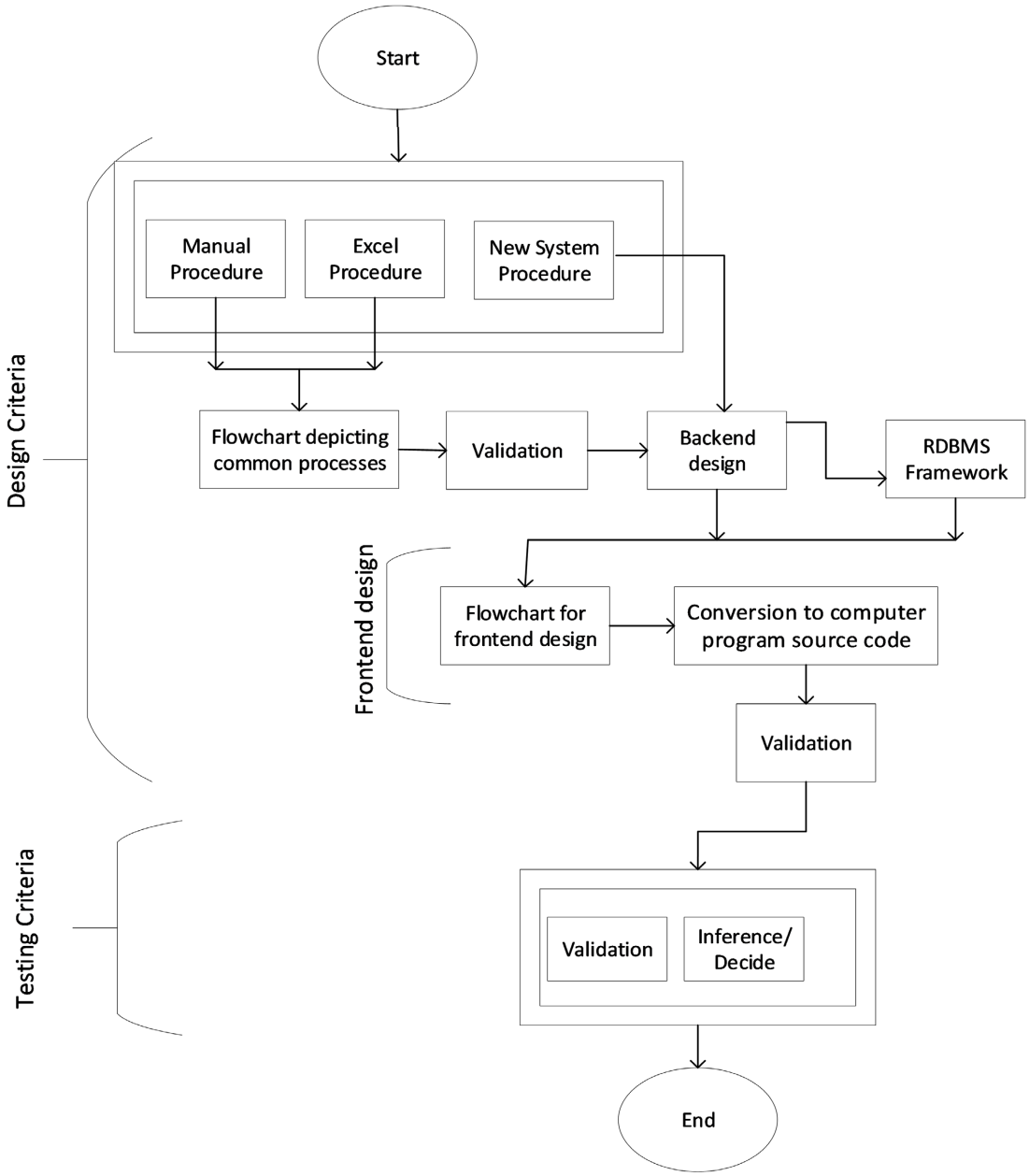



**Figure 1.** Flowchart of the design procedure for the "New" system.

the point where the process outcome is committed into the DBMS repository.

### 3.3. Data Collection and Analysis

The population in this study was the data management tasks that a laboratory administrator (head of lab.) conducts for personnel competence training and assessment operations in ANACHEM. The model or "New" system handles the stages of those tasks as modules. They include 5 stages: personnel registration (new employees or transfers into ANACHEM unit), first SOP allotment (SOP #1 to personnel), complete SOP (to confirm the previously assigned SOP is done), administer subsequent SOPs (SOP # 2, #3, #4, etc., and competence level #1),

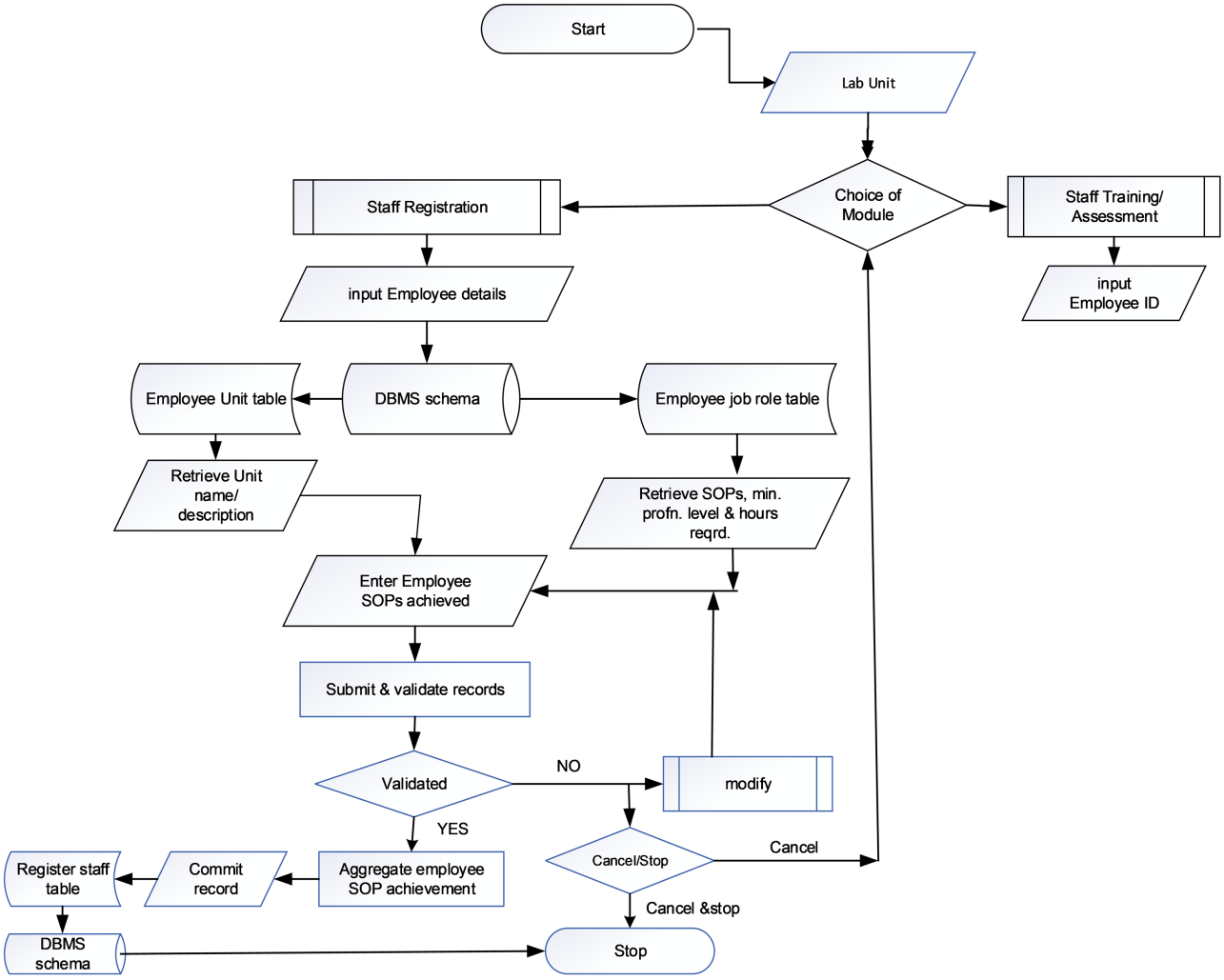



**Figure 2.** The "New" system process flow is determined by a task option (Staff Training or Assessment) module selected by an administrator.

and personnel assessment (evaluates the cumulative hours of SOP modules at competence level #1 and determine/report current proficiency level). In collecting the core data, the times (T1, T2, T3, and T4) for performing 4 tasks were computed at the stipulated 5 stages of the operation.

The data samples used in this research were time measurements obtained from 40 simulated tasks that laboratory administrator conducts for competence training and assessment operations in ANACHEM. These laboratory administrative tasks encompassed the five stages of operation as highlighted above in this section. The tasks were run within four consecutive days to obtain the times t1, t2, t3, and t4 spent in performing tasks T1, T2, T3, and T4. Essentially, these 4 tasks were run on 5 hypothetical subjects (P1, P2, P3, P4, and P5) using the Excel and the New system respectively. The sampling was performed by simulating the laboratory administrator randomly. The rationale for the sampling method and size includes:

- The need to accommodate the absence of the human subject approach that

this research employed.

- The sampling size of 40 is significantly adequate for a two-sample (or 20 matched pairs) problem from a randomized comparative experiment.

The "New" system (which is the automated software) used hypothetical subjects to generate data to create an assumed competence matrix alternative to the Excel system. Thus, comparative experimental design was a good fit for determining the relative performances of the alternatives. Using each option, data was obtained for performing the predefined core tasks for 5 hypothetical subjects. RStudio was used to evaluate the test statistics for the core data [37].

This study used the "two-sample problem" approach to answer the research question by comparing the means of the response variable in two populations (operations performed with Excel system, μExcel, and the "New" system, μNew), with an assumption that the two groups were independent. Measurement used is by "matched pair" approach. The two populations were compared by obtaining the difference in the measurement for each pair of the samples. Therefore, the following null and alternative hypotheses were formulated:

Null Hypothesis: The mean in the amount of time spent within the population of tasks performed for completing the scheduled activities in ANACHEM did not differ using the Excel system and the "New" system.

Alternative Hypothesis: The difference in the mean of time spent within the population of tasks performed for completing the scheduled activities in ANACHEM using the Excel system and the "New" system was greater than zero.

A t-test statistical test was performed to determine which of the hypotheses to reject.

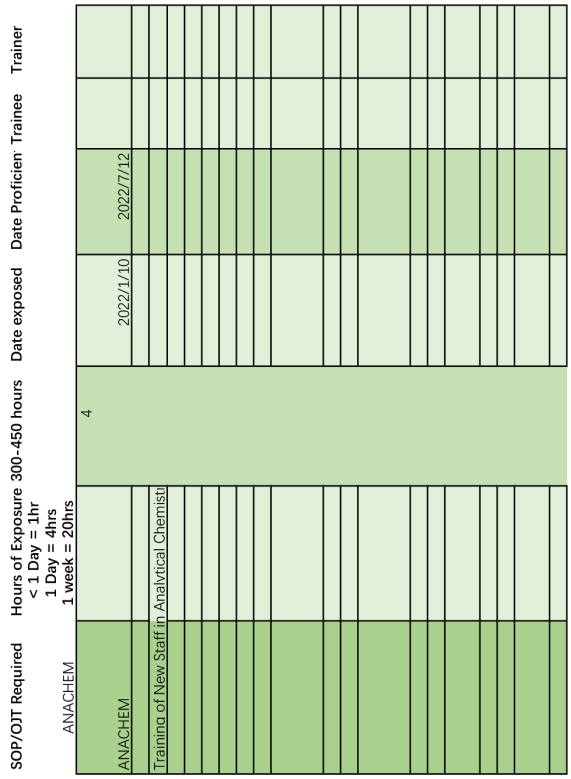
## 4. Results

### 4.1. A Review of the Old System

The old system was a manual register that contained employee name, job role, and date employed. It also contained a list of schedules of SOPs/OJTs and competence assessment tests for all cadres of staff. The schedules were arranged in sequence administered in the laboratory. One of the lists (or sheets) contained in the Excel competence matrix independently monitored duration, in terms of the hours required for each SOP/OJT training, see **Figure 3**. This list also aggregated the man-hours required for analyst to attain each competence level. As earlier observed the tables contained in Excel competence matrix are neither physically nor logically linked. Consequently, queries that were crucial for achieving efficient personnel data management in the competence matrix were difficult to attain as Excel spreadsheets were inherently not structured for database management.

### 4.2. Advantages of SQL over the Excel System

Unlike the Excel spreadsheet setup, the SQL used for the "New" system linked the tables in the competence matrix while defining their relationships in a database. This configuration provided a medium where tables for storing specific



| STEPS | ACTIVITIES + SOPS / ON-THE-JOB TRAINING REQUIREMENTS | | | | | | | SOP/OJT Required | Hours of Exposure < 1 Day = 1hr 1 Day = 4hrs 1 week = 20hrs | 300-450 hours | Date exposed | Date Proficien | Trainee | Trainer |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | ANACHEM | | 4 | | | | |
| 1. Test Samples | | | | | | | | ANACHEM | | | 2022/1/10 | 2022/7/12 | | |
| Task A: Collect sample | (a) Identify sample; inspect all sample parameters | (b) Remove sample with proper technique | (c) Label sample properly | (d) store according to specification | | | | Training of New Staff in Analytical Chemist | | | | | | |
| Task B: Determine appropriate testing | (a) Use familiarity with tests performed in the | (b) Look up in product SOP/ | | | | | | | | | | | | |
| Task C: Prepare reagents | (a) Use a balance | (b) Use a pH meter | ( c ) fill out appropriate documentation | (d) Use of conductivity meter | ( e ) use of filter/degas | (f) Review post-preparation | | | | | | | | |
| Task D: Review testing procedures | (a) Read procedure for clarity | (b) Acquire all necessary materials and equipment | | | | | | | | | | | | |
| Task E: Execute testing | (a) Set up labeled Glassware | (b) Perform pipetting | ( c ) Dilute schemes with many variables | | | | | | | | | | | |
| Task F: Analyze data | (a) Correct and interpret results | (b) Look for aberrant results | (c) Determine assay validity | | | | | | | | | | | |

**Figure 3.** Ensample of ANACHEM Excel competence matrix spreadsheet.

objects, for example, employees, SOPs, and Job Roles, were linked to easily communicate with each other. Therefore, SQL was a fit database programming language choice to effectively deploy RDBMS. The RDBMS framework contains features that support the strategy required to address the deficiencies that characterize the manual or Excel-based procedures in handling the tables associated with the personnel competence matrix.

### 4.3. Time Efficiency of the New System

The time for performing the 40 simulated tasks using the Excel and the "New" system to generate the core data is summarized in Table 4. The data set provides the simulated times it took a laboratory administrator to perform 4 different types of tasks involved in competence training and assessment using the 2 systems for 5 hypothetical subjects in the laboratory.

This study used R software as a statistical tool to compare the means of the response variable in two populations (operations performed with Excel system and with "New" system). The values for the mean difference for a sample ( $\bar{d}$ ), the standard deviation of the differences for the sample ( $S_d$ ), and the number of observations ($n$), were computed as $\bar{d} = 2.0155$, $S_d = 2.736482$, and $n = 20$. The standard error $SE_d$ and test statistic $t_s$ were computed as $SE_d = 0.6118959$ and

$$t_s = \frac{2.0155}{\frac{2.736482}{\sqrt{20}}} = 3.293861$$

The critical value $t_c$ followed $t(n-1)$ distribution and the degree of freedom,

Table 4. Time spent in performing the 4 core tasks within the population for T_New ($T_N$) and T_Excel ($T_E$).

| | Variable Definition | P1 $T_N$ (sec) | P2 $T_N$ (sec) | P3 $T_N$ (sec) | P4 $T_N$ (sec) | P5 $T_N$ (sec) | P1 $T_E$ (sec) | P2 $T_E$ (sec) | P3 $T_E$ (sec) | P4 $T_E$ (sec) | P5 $T_E$ (sec) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| T1 | T1 will be the time taken by the laboratory administrator to register personnel in ANACHEM, using their demographics, locate the first SOP training, and administer it | 0.59 | 1.15 | 1.18 | 1.12 | 1.11 | 5.51 | 3.58 | 4.16 | 3.33 | 3.51 |
| T2 | Time T2 is the time taken to record the return of the first SOP, administer the second, third, fourth, and fifth SOPs, and administer the competence level 1 assessment | 5.33 | 6.32 | 5.55 | 6.08 | 6.49 | 17.27 | 10.55 | 8.38 | 7.36 | 7.32 |
| T3 | T3 is the time taken to evaluate the cumulative hours of training acquired by the personnel at competence level 1 | 0.24 | 0.21 | 0.23 | 0.21 | 0.20 | 0.43 | 0.34 | 0.32 | 0.27 | 0.34 |
| T4 | T4 is the time taken to assess the current proficiency level (hours of training and training modules successfully completed) for the personnel after attaining competency level 1 | 0.25 | 0.21 | 0.18 | 0.19 | 0.24 | 1.03 | 1.01 | 1.19 | 1.02 | 0.47 |

$df = n-1 = 19$. Therefore, the values of critical value $t_c$ and $p$-value were computed as $t_c = 2.093024$ and $p$-value = 0.001908888. The significant $p$-value obtained ($p$ = 0.00191, $\alpha$ = 0.05), provided enough evidence to reject the null hypothesis. Therefore, the alternate hypothesis was upheld: the mean difference for the entire population was greater than zero. This means that the time taken to accomplish the tasks using the "New" system (37.08 seconds) is significantly lower than the time measurements for the Excel system (77.39 seconds).

Therefore, there was substantial evidence that the average time of completing the scheduled activities in ANACHEM was reduced when performed using the "New" system. This implied that the difference between the average time taken to complete an operation using the Excel System and using the "New" System was found to be greater than zero. A confidence interval (CI) for mean difference ($\mu_d$) was computed as estimate ± margin of error:

$$\text{margin of error, } ME_d = t_c * SE_d = 2.093024 * 0.6118959 = 1.280713$$

$$\text{CI} = \text{estimate} \pm ME_d = 2.0155 \pm 1.280713 = (0.7347872, 3.296213).$$

The CI for mean difference estimated the mean increase to be about 2.01, with a margin of error of ~1.28 points. Therefore, with 95% confidence the mean increase $\bar{d}$ in using the Excel System instead of the "New" (computerized data management) system was between 0.7 and 3.3: $0.735 \le \bar{d} \le 3.300$. The outcome verified that the Excel System performed the operation with longer time than using the "New" System. The time efficiency of the "New" system in performing data management operations of the competence matrix in ANACHEM laboratory is summarized in Figure 4 below.

| Task | xcel System (sec) | New System (sec) |
|---|---|---|
| T1 | 20.09 | 5.15 |
| T2 | 50.88 | 29.77 |
| T3 | 1.7 | 1.09 |
| T4 | 4.72 | 1.07 |

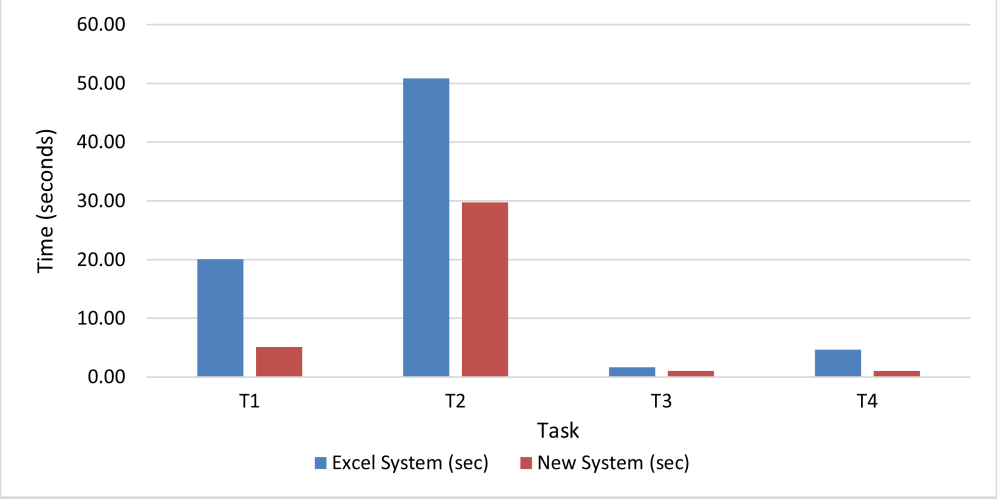



Figure 4. Time efficiency of the "New" system in performing personnel data management in ANACHEM laboratory.

The computerized competence matrix ("New" system) developed using RDBMS framework is able to register and keep track of personnel details and record the sequence of exposures to SOPs and OJTs. It also tracks date stamps for the start and completion of individual SOPs and therefore provides the time-point for assessing competence, see Figure 5. These features make it possible to access personnel proficiency status on SOPs, OJTs, and competence assessments on demand.

## 5. Discussion

A Purdue doctoral project provided the sequence of activities and timelines required by newly recruited personnel to attain different maturity levels of competence in analytical testing in an NRA QC laboratory. This project created linkages for the various activities in the unit, mapping the various recommended pieces of training. Danziger and Kraemer reported that in organizations where decision-making is highly data dependent, personnel efficiency was increased by using automated data management processes over paper records [38]. A similar study evaluated the impact of computerized information in healthcare system

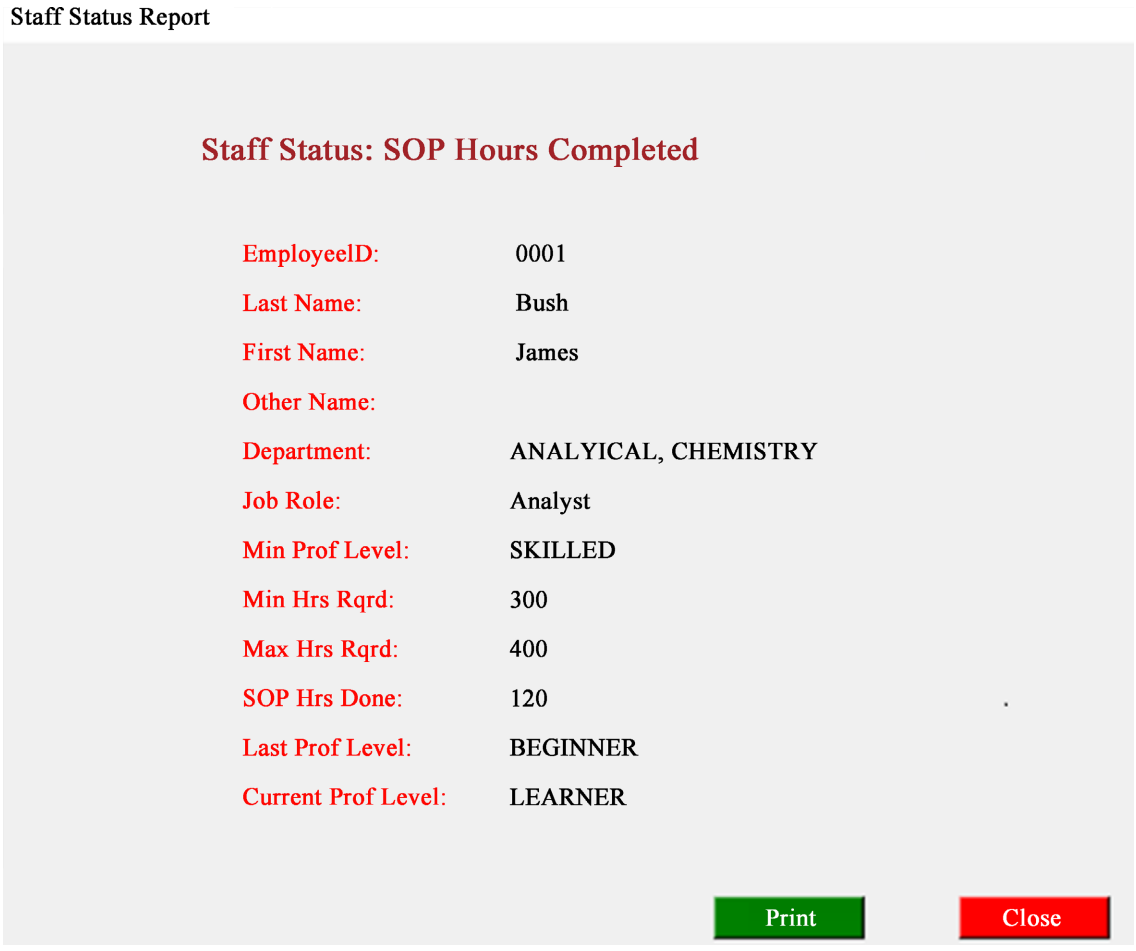



Figure 5. The Staff Status Report is a key deliverable of the computerized competence matrix with the RDBMS framework.

and recommended distinctive roles and tasks for technical personnel for effective operation [39]. This study identified various deficiencies with the manual record-keeping system in the NRA QC laboratory to include human errors, loss of data, difficulties with archival processes, and long processing and retrieval times. It also delineated the limitations of an Excel spreadsheet system that was used to mitigate the observed deficiencies. The study provided a more efficient computerized system that provides evidence that automation substantially addressed the limitations experienced in the personnel data management in NAFDAC ANACHEM. Statistical test results showed a significant decrease ($p$ = 0.00191, $\alpha$ = 0.05)) in the average time demand required to perform 4 scheduled tasks regarding data management operation in ANACHEM using the "New" system compared to using the Excel system. This study result confirmed that the computerization of personnel data management operations using RDBMS reduced the time demand for administering the personnel competence training and assessments in ANACHEM.

The "New" system mitigated the limitations of the old system and was able to: register and keep track of personnel details; record the sequence of SOPs and OJTs for all cadres of personnel; track date stamps for start and completion of individual SOP/personnel; provide the time-point for assessing personnel competence; and access personnel proficiency status on SOP & OJT and competence assessments.

In line with the WHO guidance for NRAs, personnel competence training and its assessment is a process that relies on good data management (GDM) [17]. In this study, the principles of GDM were useful to transform ANACHEM's manual personnel data management and scale up the Excel spreadsheet developed by a Purdue dissertation project [6]. This study reviewed the current competence matrix in the ANACHEM laboratory and automated its personnel data management processes. The "New" system can map the sequence and date stamps of personnel training and competence assessments. It also provides a comprehensive record of personnel training history. Therefore, the "New" system developed by this project provides a more reliable platform in line with WHO recommendation for GDM in ANACHEM laboratory.

It is envisaged that implementing a computerized system will enhance the laboratory's quality management system. Data privacy strategy in manually driven processes cannot support long-term maintenance. Filing either raw or processed data inside folders and storing them in cabinets does not achieve the intended data secrecy. It gives false security assurance as keeping documents in file cabinets can engender security threats to the object [40]. Other associated challenges include increased cost of man-hours, human errors associated manual operations, and costs accrued to data loss. These shortcomings were addressed by the RDBMS-based competence matrix system developed in this study.

This automated "New" system supported substantially lower time in executing the same operation in comparison to the Excel system in the ANACHEM labor-

atory. It is recommended that this "New" system can be customized to handle other quality management processes.

## 6. Conclusion

This study ascertained that there was a significant decrease in the average time (37.08 seconds Vs 77.39 seconds, $p = 0.00191$, $\alpha = 0.05$) for completing 4 scheduled sets of activities in ANACHEM laboratory when RDBMS-based "New" system was used to automate the Excel system. Unlike the Excel system, the "New" system software supports real-time and on demand access to personnel proficiency levels thereby enhancing time efficiency of the laboratory. Finally, computerization of the data management operations achieved with this study will support ANECHEM laboratory's desire for a more efficient process for administering competence, training, and assessment.

## Acknowledgements

We acknowledge the supports of Stephen Byrn and Mercy Okezue from Purdue University IPPH Department and the Nigerian FDA Analytical Chemistry laboratory.

## Availability of Data and Materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Authors' Contributions

Marcel Okezue: Conceptualization, Methodology, Data curation, Validation. Eric Dietz: Supervision, Funding Acquisition.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

[1] Pierre-Louis, L. (2014) Pharmaceuticals Regulation—An Overview. https://www.sciencedirect.com/topics/pharmacology-toxicology-and-pharmaceutical-science/pharmaceuticals-regulation

[2] Selvaraj, G. and Hari, S. (2018) A Study on Competency Mapping among Operators. *Indian Journal of Applied Research*, **8**, 45-47.

[3] Getha-Taylor, H., Hummert, R., Nalbandian, J. and Silvia, C. (2013) Competency Model Design and Assessment: Findings and Future Directions. *Journal of Public Affairs Education*, **19**, 141-171. https://doi.org/10.1080/15236803.2013.12001724

[4] Gonsalvez, C.J. and Calvert, F.L. (2014) Competency-Based Models of Supervision: Principles and Applications, Promises and Challenges. *Australian Psychologist*, **49**, 200-208. https://doi.org/10.1111/ap.12055

[5] National Agency for Food and Drug Administration and Control (2004) NAFDAC

Mandate.
https://www.nafdac.gov.ng/about-nafdac/nafdac-organisation/nafdac-functions/

[6] Okezue, M.A. (2022) Developing Multipronged Models to Enhance Effectiveness of the Management System of Quality Control Laboratories. Additional Focus on Synthesis and Characterization for 5 New Salts of Bedaquiline. Master's Thesis, Purdue University Graduate School, West Lafayette.
https://hammer.purdue.edu/articles/thesis/DEVELOPING_MULTIPRONGED_MODELS_TO_ENHANCE_EFFECTIVENESS_OF_THE_MANAGEMENT_SYSTEM_OF_QUALITY_CONTROL_LABORATORIES_ADDITIONAL_FOCUS_ON_SYNTHESIS_AND_CHARACTERIZATION_FOR_5_NEW_SALTS_OF_BEDAQUILINE/19620642/1

[7] Guerrero, C. (2021) Competencies of Medical Laboratory Personnel: Basis for Developing a Health Training Program. ScienceOpen Posters.
https://www.scienceopen.com/hosted-document?doi=10.14293/S2199-1006.1.SOR-.PPVQKVL.v1
https://doi.org/10.14293/S2199-1006.1.SOR-.PPVQKVL.v1

[8] Sohne, G. (2008) Community Revenue Collection System. *Journal Community Informatics*, **4**, 1-9.
https://www.semanticscholar.org/paper/Community-Revenue-Collection-System-Sohne/da2af9db9a802e746cd269d65d03e9b7929bbe07

[9] Kim, H., Dykes, P.C., Thomas, D., Winfield, L.A. and Rocha, R.A. (2011) A Closer Look at Nursing Documentation on Paper Forms: Preparation for Computerizing a Nursing Documentation System. *Computers in Biology and Medicine*, **41**, 182-189.
http://www.proquest.com/docview/1033165112/abstract/43D684298CF419BPQ/1
https://doi.org/10.1016/j.compbiomed.2010.08.006

[10] AL-Rawajfah, O.M., Aloush, S. and Hewitt, J.B. (2015) Use of Electronic Health-Related Datasets in Nursing and Health-Related Research. *Western Journal of Nursing Research*, **37**, 952-983. https://doi.org/10.1177/0193945914558426

[11] Okezue, M.A., Adeyeye, M.C., Byrn, S.J., Abiola, V.O. and Clase, K.L. (2020) Impact of ISO/IEC 17025 Laboratory Accreditation in Sub-Saharan Africa: A Case Study. *BMC Health Services Research*, **20**, Article No. 1065.
https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-020-05934-8
https://doi.org/10.1186/s12913-020-05934-8

[12] Melissa, B. (2021) How to Switch from Excel to SQL.
https://mode.com/blog/learn-sql-using-excel

[13] Gupta, Y. (2020) Excel V/S MySQL. Data Science Simplified.
https://medium.com/dssimplified/excel-v-s-mysql-de32b78e7988

[14] Azam, A., Alam, K.A. and Umair, A. (2019) Spreadsheet Smells: A Systematic Mapping Study. 2019 *International Conference on Frontiers of Information Technology* (*FIT*), Islamabad, 16-18 December 2019, 345-3455.
https://doi.org/10.1109/FIT47737.2019.00071

[15] Powell, S.G., Baker, K.R. and Lawson, B. (2009) Errors in Operational Spreadsheets: A Review of the State of the Art. 2009 42*nd Hawaii International Conference on System Sciences*, Waikoloa, 5-8 January 2009, 1-8.

[16] Rahman, S., Bendre, M., Liu, Y., Zhu, S., Su, Z., Karahalios, K., *et al.* (2021) NOAH: Interactive Spreadsheet Exploration with Dynamic Hierarchical Overviews. *Proceedings of the VLDB Endowment*, **14**, 970-983.
https://doi.org/10.14778/3447689.3447701

[17] World Health Organization (2016) Guidance on Good Data and Record Manage-

ment Practices. WHO Technical Report Series No. 996. https://www.gmp-compliance.org/files/guidemgr/WHO_TRS_996_annex05.pdf

[18] Soegoto, E.S., Wardana, M.E. and Oktafiani, D. (2021) Implementation of Management Information System at Health Agencies. *Journal of Physics: Conference Series*, **1764**, Article ID: 012191. https://doi.org/10.1088/1742-6596/1764/1/012191

[19] Newbrander, W.C. and Thomason, J.A. (1988) Computerizing a National Health System in Papua New Guinea. *Health Policy and Planning*, **3**, 255-259. https://doi.org/10.1093/heapol/3.3.255

[20] Rathnayake, W.D.W.T. (2021) Pharmacy Management System for the Central Pharmacy—Pokunuwita. https://dl.ucsc.cmb.ac.lk/jspui/handle/123456789/4312

[21] Sarker, N.C. (2018) Online Inventory and Order Management System for Pharmacy. http://dspace.uiu.ac.bd/handle/52243/469

[22] Mostafa, S.A., AbuSalim, S.W. and Saringat, M.Z. (2020) A Comparative Study of Data Management Systems. *Journal of Soft Computing and Data Mining*, **1**, 10-16. https://doi.org/10.30880/jscdm.2020.01.01.002

[23] Batra, S. and Sachdeva, S. (2021) Pre-Processing Highly Sparse and Frequently Evolving Standardized Electronic Health Records for Mining. Handbook of Research on Disease Prediction through Data Analytics and Machine Learning. IGI Global. https://doi.org/10.4018/978-1-7998-2742-9.ch002

[24] Tashkandi, A., Wiese, I. and Wiese, L. (2018) Efficient In-Database Patient Similarity Analysis for Personalized Medical Decision Support Systems. *Big Data Research*, **13**, 52-64. https://www.sciencedirect.com/science/article/pii/S2214579617303088 https://doi.org/10.1016/j.bdr.2018.05.001

[25] Li, H., Zhao, Z., Cheng, Y., Lu, W., Du, X. and Pan, A. (2019) Efficient Time-Interval Data Extraction in MVCC-Based RDBMS. *World Wide Web*, **22**, 2633-2653. https://doi.org/10.1007/s11280-018-0552-7

[26] Paul, A.K., Wang, B., Rutman, N., Spitz, C. and Butt, A.R. (2020) Efficient Metadata Indexing for HPC Storage Systems. 2020 20*th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing* (*CCGRID*), Melbourne, 11-14 May 2020, 162-171. https://doi.org/10.1109/CCGrid49817.2020.00-77

[27] Gai, L. (2020) Leveraging Neighborhood Summaries for Efficient RDF Queries on RDBMS. arXiv: 2001.08329. http://arxiv.org/abs/2001.08329

[28] Gard, M., Hasterok, D. and Halpin, J.A. (2019) Global Whole-Rock Geochemical Database Compilation. *Earth System Science Data*, **11**, 1553-1566. https://doi.org/10.5194/essd-11-1553-2019

[29] Daphine, A. (2013) An Automated Library Management System: A Case Study of Kashenyi Vocational Senior Secondary School. https://ir.kiu.ac.ug/handle/20.500.12306/13387

[30] Albalkhi, A. (2012) Repetitive Querying of Large Random Heterogeneous Datasets in RDBMS Using Materialized Views. Master's Thesis, University of Windsor, Windsor. https://scholar.uwindsor.ca/etd/311

[31] Azeez, N., Dasylva, A. and Omisore, O. (2018) RDBMS Query for Authenticating Access to Patient Information. *Computer Science & Telecommunications*, **54**, 44-56. https://www.researchgate.net/publication/341946386

[32] Coetzee, L.M., Tepper, M.E., Perelson, L., Glencross, D.K. and Cassim, N. (2020) Timely Delivery of Laboratory Efficiency Information, Part II: Assessing the Impact of a Turn-Around Time Dashboard at a High-Volume Laboratory. African Journal of Laboratory Medicine, 9, a948. https://hdl.handle.net/10520/EJC-1d3b5bf122

https://doi.org/10.4102/ajlm.v9i2.948

[33] Mayer, D.A., Rasmussen, L.V., Roark, C.D., Kahn, M.G., Schilling, L.M. and Wiley, L.K. (2021) ReviewR: A Light-Weight and Extensible Tool for Manual Review of Clinical Records. *JAMIA Open*, **5**, ooac071.

[34] Sánchez-de-Madariaga, R., Muñoz, A., Castro, A.L., Moreno, O. and Pascual, M. (2018) Executing Complexity-Increasing Queries in Relational (MySQL) and NoSQL (MongoDB and EXist) Size-Growing ISO/EN 13606 Standardized EHR Databases. *Journal of Visualized Experiments*, **1**, e57439. https://doi.org/10.3791/57439

[35] Ravikiran, A. (2021) Understanding the Difference between SQL and MySQL. https://www.simplilearn.com/tutorials/sql-tutorial/difference-between-sql-and-mysql

[36] Reese, G., Yarger, R., King, T. and Williams, H.E. (2002) Managing & Using MySQL: Open Source SQL Databases for Managing Information & Web Sites. O'Reilly Media, Sebastopol.

[37] R Core Team (2018) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

[38] Danziger, J.N. and Kraemer, K.L. (1985) Computerized Data-Based Systems and Productivity among Professional Workers: The Case of Detectives. *Public Administration Review*, **45**, 196-209. https://doi.org/10.2307/3110149

[39] Hubaish, A.M.A., Addeeb, B.M., Al-Serouri, A., Mughalles, S. and Ghaleb, Y. (2022) The Impact of Computerized Health Information System on Medical & Administrative Decisions, Sana'a Hospitals, Yemen (2017-2020). *E-Health Telecommunication Systems and Networks*, **11**, 47-66. http://www.scirp.org/Journal/Paperabs.aspx?paperid=117867 https://doi.org/10.4236/etsn.2022.112004

[40] Leiwo, J. and Heikkuri, S. (1998) An Analysis of Ethics as Foundation of Information Security in Distributed Systems. *Proceedings of the Thirty-First Hawaii International Conference on System Sciences*, Kohala Coast, 9 January 1998, 213-222.