Smart Web Application for Data Sorting and Filtering

Shahad Aljoufi, Malak Alnory

[1,2]College of Engineering, Effat University, AnNazlah Al Yamaniyyah, Jeddah 22332, Saudi Arabia
[2]malnory@effatuniversity.edu.sa

Abstract:

Data management is an important element for an organization. By having a good data management system, an organization can sustain its overall business management process. This work presents the development ofsmart web applicationfor data management in terms of sorting and filtering for Saudi Binladin Group (SBG). The functional requirement of the developed web application included creating user account, editing user account, uploading excel sheet, sorting and filtering data, editing reject table, and checking golden table. This web application will be used by two users, which are human resource (HR) admin and data quality team (DQT). In this work, the web application was designed using bootstrap, HTML, CSS, Javascript and C# programming. The developed platform was successfully tested for each use case under several scenarios. The test outcome showed that web application system functioned well was successful for each tested scenarios. Therefore, the developed web application system will enhance thedata management for each sector in SBG organization.

1. INTRODUCTION

Business data management are viewed as important process for any organization and it offers organizations a chance to advance in the manner that will contribute to the overall sustainability of an organization [1].Despite the fact that their value is not represented in the accounting report, data and information are one of the most significant resources of the association. Company and business associations are collectively genuine in their view of information and data as their main business assets, requiring sound management as a benefit, as they face growing competition between business contenders [2].

If properly handled and processed, information and data would become the most precious asset of the organization, assisting it to stay efficient and flexible, proactively meeting client requirements, and keeping expenses under control [3]. For instance, information on the profitability of a specific client may be used in various company situations. Furthermore, public and private organizations of all dimensions are endeavouring to supervise data as a venture asset, to be distributed and utilized across a wide range of consumer applications and frameworks, business processes and customers across the organization [4]. Organization finds that they need to establish values, methods, and processes for managing and sorting their valuable data. Furthermore, organizations feel that it is equally essential to build a privileged authority framework and an entrepreneurship base to assist manage their data [5].

In terms of commercial and information technology governance exercise, data management is described as the procedures, strategies, norms, organisation and technology needed to handle and guarantee the accessibility, reliability, performance, accuracy, auditability and data safety of an organisation [6]. Likewise, data management aims to guarantee that information serves company requirements, safeguard and handle information as a prized

company asset, and reduce data management expenses [7]. In some cases, this is some part of the more extensive business data management of executive activities, which is a procedure to improve the accuracy and reliability of data resources in order to better serve the specialized, operational and business objectives of the organization [8].

Several previous works have reported on the practise of data management and governance. Wolfert et al.[9] reviewed the practice of data management in the smart farming and agriculture industry and found that data management had a positive impact on the farming industry and it enhanced the prospect for long term sustainability. Siddiqa et al.[10] presented a review work on data management and concluded that there are many available big data techniques that are currently being utilized by various organization to manage their data in a comprehensive manner. Alhassan et al. [11] presented a work that compared data management procedures and found that data management operations provide a broad perspective for the organization in terms of overall performance and business direction. Hondula et al. [12] examined the role of data management in the government management of natural hazards in the United States of America (USA) and found that data management provided an accurate view of the severity of environmental hazards and made it possible for the authority to take data-based corrective action. Paskaleva et al.[13] analyzed the application of data governance for smart cities and found that data management has assisted in cultivating sustainability features for the smart cities. Pardee et al. [14] work presented the application of data management in the healthcare sector and found that the creation of a data management system has helped to improve the overall organization of the healthcare sector. Motenko et al. [15] have developed a new data management system for biological applications where the system is designed to manage and sort laboratory mouse data. The findings of this study have shown that

the system has simplified the listing and analysis process for laboratory mouse data.

Organized and coordinated reports are now very important for every organization. The Saudi Binladin Group (SBG) is a multinational construction company in Saudi Arabia, with more than 100,000 employees working in 8 different divisions.SBG has different sectors that contain a huge number of employees. However, the SBG organization faces difficulties in terms of data management where it wastes as precious time to request records from each division. In addition, there are human errors and errors while editing records.As a result, there is more risk involved while editing the logs. In addition, there is a lack of tracking and updating of staff records that leads to a huge mess between the SBG and the divisions.Therefore, this work was done to develop a single smart web applicationfor data management in terms of sorting and filtering for SBG.

2. SYSTEM ANALYSIS

2.1 User Functional Requirement

The six main functional requirements of this study are Create users account, Edit user account, Upload Excel sheet, Sorting and filtration, Edit rejected table, and Check Golden table. The precondition, post condition, and exception paths of the six functional requirements are outlined in Table 1. There is the need to be "quality-conscious" when developing functional requirements for an application. It is important to reduce the occurrence of errors in business processes, in dataentry etc. And, analysts need to ensure requirements and rules that impact data quality are explicitly stated, not implied or assumed. This will help to prevent data quality issues.

Table 1. Functional Requirements

Use Case Name	Precondition	Post condition	Exception Paths
Create users account	The HR admin has accessed the profile main page to add an account.	The user account has been added to the web application.	The User might already have an account.
Edit user account	The HR Admin has accessed his profile main page.	The correction information has changed.	None
Upload Excel sheet	The user has accessed to the excel sheet table page.	The request has been sent.	The column name in excel sheet not the same as the field name written in the programming logic.
Sorting and filtration	The system will show the rejected data in a list.	The correction information has been transferred to the Golden table. The rejected information has been transferred to the rejected table.	If the field is undefined
Edit rejected table	The user access the Rejected table main page.	The database has been updated.	If the field is empty or already edited with wrong data regarding to the rule, the data will be returning back in rejected table and showing field error.
Check Golden table	The user has accessed the Golden table page.	The records automatically updated in database.	The user may abandon the operation at any time.

2.2 User Characteristics

Users should use internet while using the web application and be able to use withno or minimal support. The data entry team should be able to understand how to use the web application withminimal support or documentation and be well verifying while editing employees datain records. The users must use the web application and adding values in fields. The users should know how to back up data avoiding losing employees data and information.

2.3 Non Functional Requirement

The web application is on a server with high speed Internet capability. The physical machine used is determined by oracle. The software developed tool is Oracle Application Express for connection between the Web Application and the database. The speed of the web application connection will depend on the hardware used rather than characteristics of this system. The HR Admin will run on the user's PC and will contain an Access database. Access is installed on the computer with Windows operating system. The description of the non-functional of this study is tabulated in Table 2.

Table 2. Non Functional Requirements

Criteria	Description.
User-friendly Interface	The system must have a user friendly interface.
Availability	The system must be available at all times and to everyone.
Maintainability	The system must be in continues improvement, always maintaining its level.
Reliability	The system's attribute that consistently performs according to its specifications.
Scalability	The ability of the system to handle a growing and increasing about of work in a capable manner.
Capacity Planning	Estimating the space of the device hardware, software, and connection infrastructure resources that can grow to meet future needs.
Back-Up	A duplication of important data in case of a hard drives failure, user error, disaster or accident, and then restores the data.

2.4 User Cases

Figure 1 demonstrates the use case diagram of this study. The user (HR) will access the web application through the secure IP address

to upload Excelsheets. Next, The Data Quality Team (DQT) will be editing and deleting error fields in the table. The HR Admin has the ability to add new DQT account and create accounts,and view or edit all the records for their sector in the web application.
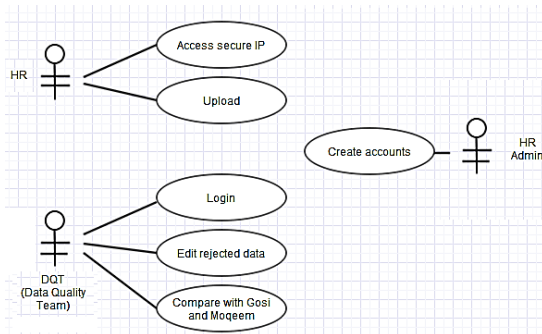

Figure 1. Use Case Diagram

3. SYSTEM DESIGN

3.1 System Architecture

Data Quality Engine (DQT) will begin its process after receiving records that is been sent afterthe ETL process. The data will be landing in the "Landing Area database" then it will besend to Data Quality Engine process which is browsing the data in three steps; Dataprofile, Rules repository and enrichment and Data engine. After this process the rejected files will be transferred to "Rejected table" or to "Golden table" to be reviewed by theData entry team and analysis team through the web application, and then the records will beupdated in one database. Figure 2 shows the system architecture.

Figure 2. The data quality engine high level architecture

3.2 Data flow Diagrams

A Data Flow Diagram in Figure 3 shows the connections between the distinctive techniques in a framework.It shows how the information streams and how it is moved in a certain framework. It likewisedemonstrates how an information is changed over to a yield come about through the diversemethods in the framework and frequently started by the on-screen actors.
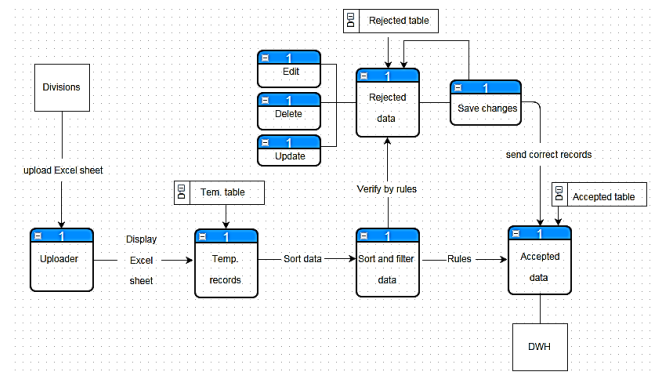

Figure 3. Data flow Diagrams

3.3 Context Diagram

The context diagram shown in Figure 4 defines the boundary between the system, or part of a system, and its environment, showing the entities that interact with it.
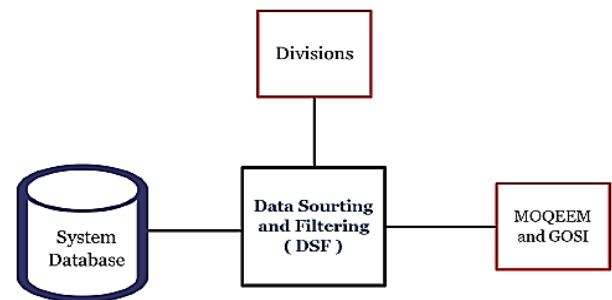

Figure 4. Context diagram

4. IMPLEMENTATION AND TESTING

There is five database tables had been created (Golden table, Rejected table, Temporary table, Error table) and build in database tables (Users). Golden table is the accepted records that are going to show through the interface with 14 fields, and Rejected table includes records that not apply on the rules.

The build in database tables, Users table which include Username, Password, and Role for the following users, which are the HR employee, HR admin and Data entry team. For the Error table, is the table where all the errors are recorded when move data from temporary table to other tables (Golden tables and rejected table).

4.1 Creating User Interface

Interface is designed using modern day techniques like CSS media queries, to detect the different screen sizes and resolution. The main CSS library used is bootstrap which designed by twitter incorporation to make it easy for any web application to adapt different screen sizes. Bootstrap, can be added in solution explorer in visual studios using nugget package manager. While designing interface a special care, has been taken for ease of users to understand the usage. Each function has a different page, which can be accessed through a top navigation bar. There is a master page, which keep the design uniform in all across the web application advance Modal used for uploading excel.

4.2 Interface Description

In the implementation phase, the web application was done using HTML, CSS, Javascript and C# programming. The bootstrap is used for the design, and bootstrap is open source and accessible on Github. Figure 5 to Figure 9 illustrates the platforms of homepage, uploading excel sheet, temporary records, tejected records, and golden records respectively.
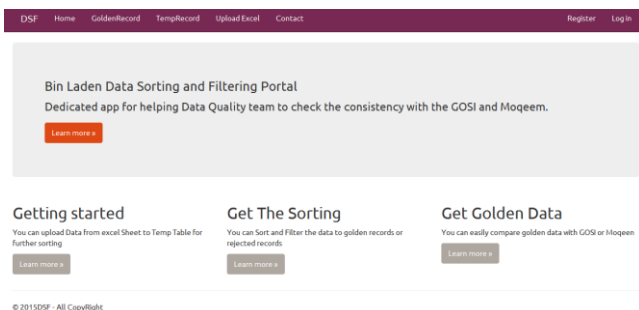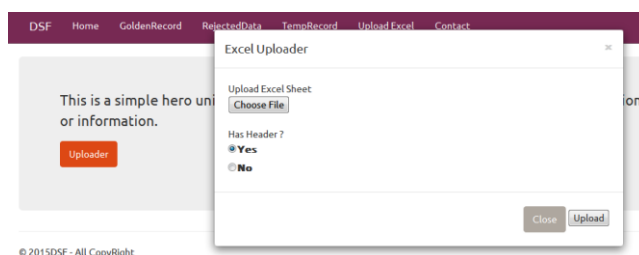
Figure 5. Homepage

Figure 6. Uploading Excel Sheet

Figure 7. Temporary records

Figure 8. Rejected records

Figure 9. Golden records

4.3 Unit Testing

The break points is placed in different parts of the code, and after running the debugger each part of the code were run between the break points and wherever there was an error can be identity it by observing that which breakpoint is not working probably.Table 3 tabulates the unit testing result.

Table 3. Unit testing

Use Case	Scenario	Expected System Behaviour	Actual System Behaviour	Result
Log In	Enter username = "Shahad" And password = "mypass"	The system should allow this DQT o login successfully.	The system allows DQT to login successfully.	Success
	Enter username = "Shahad123" And password = "my pass"	Invalid username or password.	Invalid username or password.	Success
Upload Excel Sheet	Choose excel sheet from computer to upload	Table should display to HR employee. The system should allow HR employee to upload excel sheets.	Table is displayed to HR employee. The system allow HR employee to upload excel sheets.	Success
	Choose Word doc. from computer to upload	Upload right file type.	Upload right file type.	Success
Sort and Filter	Click "Sort and Filter" button for previously displayed table.	The system should allow DQT to get the result of error fields. Tables should in rejected and golden.	The system allows DQT to get the result of error fields. Tables shown in rejected and golden.	Success
	Click "Sort and Filter" button without table shown.	No data is found	No data is found	Success
Edit rejected records	1) Select rejected records tab. 2) Click "Edit" field. 3) Update page. 4) Save changes.	If the data entered is correct, the row should show to DQT in golden table otherwise it will remain in rejected table.	If the data entered is correct, the row show to DQT in golden table otherwise it remain in rejected table.	Success

	1) Select rejected records tab. 2) Click "Edit" field. 3) Update page.	Data remain in rejected records.	Data remain in rejected records.	Success
Check Golden records	Select Golden records tab after saving the changes.	The system should show for DQT the correct records.	The system shows for DQT the correct records.	Success
	Select Golden records tab without editing records.	No data is found	No data is found	Success

5. OVERALL DISCUSSION

This study has developed a single smart web application for data management. The application consisted of six main functional requirements to for data management, namely create users account, edit user account, upload Excel sheet, sorting and filtration, edit rejected table, and checking golden table. The data sorting and filtering (DSF) comprised of several elements namely system database, division, as well as MOQEEM and GOSI. Furthermore, five database tables had been created, which are golden table, rejected table, temporary table, error table and build in database tables (Users).The main CSS library used was bootstrap to create the user interface. Furthermore, the web application was implemented and tested under different scenarios across each case namely Log In, Upload Excel Sheet, Sort and Filter, Edit rejected records, and Check Golden records. The outcome of the testing phase has shown that the developed web application has function accordingly and good data management process was achieved. The result of this work has shown that the developed system has improved data management, which is in line with the work of Wolfert, et al. [9] and Siddiqa et al. [10] where it has been confirmed that the data management system contributes to improving the efficiency of the organization's data assessment, sorting and filtering process.

6. CONCLUSION

This study has presented the development of a smart web applicationfor data management. The application was tested and the results have shown that it is reliable in managing, sorting and filtering the required data. In addition, the web application will overcome data sorting errors within a limited time period and prevent incorrect or missing information from being inserted. Finally, this application will be the main solution for the HR in each division facing employee data reporting problems.

7. REFERENCES

[1] Pels, P., Boog, I., Henrike Florusbosch, J., Kripe, Z., Minter, T., Postma, M., ... & von Poser, A. (2018). Data management in anthropology: the next phase in ethics governance?. *Social Anthropology*, *26*(3), 391-413.

[2] Koltay, T. (2016). Data governance, data literacy and the management of data quality. *IFLA journal*, *42*(4), 303-312.

[3] Sylvia, M. L. (2018). *Clinical analytics and data management for the DNP*. Springer Publishing Company.

[4] Winter, J., Davidson, E., Boyce, C., & Fan, V. (2019, July). Emergence, Convergence, and Differentiation of Organizational Forms of Health Data Governance. In *Academy of Management Proceedings* (Vol. 2019, No. 1, p. 16403). Briarcliff Manor, NY 10510: Academy of Management.

[5] Mahanti, R. (2019). Data Quality: Dimensions, Measurement, Strategy, Management, and Governance. ASQ Quality Press.

[6] Hill, D. G. (2016). Data protection: Governance, risk management, and compliance. CRC Press.

[7] DeStefano, R. J., Tao, L., &Gai, K. (2016, June). Improving data governance in large organizations through ontology and linked data. In *2016 IEEE 3rd International Conference on Cyber Security and Cloud Computing (CSCloud)* (pp. 279-284). IEEE.

[8] Espinosa, J. A., &Armour, F. (2016). The big data analytics gold rush: a research framework for coordination and governance. In *2016 49th Hawaii International Conference on System Sciences (HICSS)* (pp. 1112-1121). IEEE.

[9] Wolfert, S., Ge, L., Verdouw, C., &Bogaardt, M. J. (2017). Big data in smart farming–a review. *Agricultural Systems*, *153*, 69-80.

[10] Siddiqa, A., Hashem, I. A. T., Yaqoob, I., Marjani, M., Shamshirband, S., Gani, A., &Nasaruddin, F. (2016). A survey of big data management: Taxonomy and state-of-the-art. *Journal of Network and Computer Applications*, *71*, 151-166.

[11] Alhassan, I., Sammon, D., & Daly, M. (2018). Data governance activities: A comparison between scientific and practice-oriented literature. *Journal of Enterprise Information Management*, *31*(2), 300-316.

[12] Hondula, D. M., Kuras, E. R., Longo, J., & Johnston, E. W. (2018). Toward precision governance: infusing data into public management of environmental hazards. *Public Management Review*, *20*(5), 746-765.

[13] Carretero, A. G., Freitas, A., Cruz-Correia, R., &Piattini, M. (2016). A case study on assessing the organizational maturity of data management, data quality management and data governance by means of MAMD. In *ICIQ* (pp. 75-84).

[14] Paskaleva, K., Evans, J., Martin, C., Linjordet, T., Yang, D., &Karvonen, A. (2017, December). Data governance in the sustainable smart city. In *Informatics* (Vol. 4, No. 4, p. 41). Multidisciplinary Digital Publishing Institute.

[15] Pardee, R. E., Bachman, D., Hornbrook, M. C., Cleveland, C. R., Mathur, P., Ng, D., ... & Wong, C. C. (2016). The Learning Healthcare (Data) System: Virtual Data Warehouse Data Capture Revisited. *Journal of Patient-Centered Research and Reviews*, *3*(3), 226-227.

[16] Motenko, H., Neuhauser, S. B., O'keefe, M., & Richardson, J. E. (2015). MouseMine: a new data warehouse for MGI. *Mammalian Genome*, *26*(7-8), 325-330.