

J-Selaras: New Algorithm for Automated Data Integration Tools

Mustafa Man^{1*}

Faculty of Ocean Engineering Technology and Informatics
Universiti Malaysia Terengganu (UMT)
21030 Kuala Nerus, Terengganu, Malaysia

Mohd. Kamir Yusof³

Faculty of Informatics and Computing
Universiti Sultan Zainal Abidin (UniSZA)
Besut Campus, 22200 Besut, Terengganu, Malaysia

Wan Aezwani Wan Abu Bakar^{2*}

Pusat Asasi Sains dan Perubatan,
Universiti Sultan Zainal Abidin (UniSZA)
Gong Badak Campus, 21030 Kuala Nerus, Terengganu,
Malaysia

Norisah Abdul Ghani⁴, Mohd Adza Arshad⁵

JKR Centre of Excellence for Engineering and Technology
(CREaTE)
Public Work of Department (JKR) Malaysia, Jalan Kemus,
Simpang Ampat, 78000 Alor Gajah, Melaka, Malaysia

Raja Normawati Raja Ayob⁶, Kamarul Azhar Mahmood⁷, Faizul Azwan Ariffin⁸, Mohamad Dizi Che Kadir⁹, Lily
Mariya Rosli¹⁰, Nurhafiza Binti Abu Yaziz¹¹

Cawangan Kontrak dan Ukur Bahan (CKUB)

Public Work of Department (JKR) Malaysia, Tingkat 16, Menara Tun Ismail Mohamed Ali, No. 25, Jalan Raja Laut 50350 Kuala Lumpur, Malaysia

Abstract—Data integration is a popular technique or method today for data converting and sharing within new application with different database format and location. The interaction of data from one application system to another application system requires an intermediary software or middleware that allows the data to be transferred or read systematically and easily. The development of dynamic algorithms allows data in various formats, whether structured or unstructured, to be transferred to various types of databases smoothly. A case study was conducted for the Bill of Quantity (BQ) data in the known Excel format generated through CostX software in a single sheet Excel file. It was transferred to a single workbook with multiple sheets with formulation generated automatically. Thus, an algorithm was developed and tested through the development of the J-Selaras System. This algorithm can remove the noisy data or data symbols that are not related in the excel single sheet (CostX) file and automatically transfer to multiple excel sheets with macros formulation quickly. The implementation results indicate a significant contribution where it reduces in execution time of BQ processes and manpower resources used.

Keywords—Automated data integration algorithm; bill of quantity (BQ); CostX; single and multiple sheets; J-Selaras tool

I. INTRODUCTION

This research case study focuses on Malaysian government agency, named as Jabatan Kerja Raya (JKR) which is responsible for the BQ for construction in a tendering process. The collaboration is done to solve the manual process in converting The CostX single sheet BQ into MS Excel multiple sheets with macro function.

The Malaysian Public Works Department or simplified as Jabatan Kerja Raya (JKR) Malaysia is the Malaysia's federal

government department under the roof of Ministry of Works Malaysia (MOW), that is accountable for construction and maintenance activities on public facility and infrastructure [1]. The JKR Strategic Plan 2021-2025 was issued, and its outlines are five strategic themes. The initiatives in the strategic themes are to realize the government effort in achieving the Shared Prosperity Vision 2030. In the alignment to the first theme in the JKR Strategic Framework, the demands for a project delivery in terms of the completion within the stipulated time, within the budget and the stated quality becomes more prominent in fulfilling customer satisfaction.

There is the need to expedite the preparation of the contract document after signing of the Letter of Acceptance (LoA) by the contractor before the project is being awarded. Both the Government and the contractor must fulfil their obligations according to the contract's provisions. Typically, the contractor must provide the contract conditions, the tender drawings, the Bills of Quantities (BQ), and the other related documents throughout the process of tendering. These materials will be bound into the Contract Document after the LoA is issued. The contract document must be prepared within four (4) months upon signing of the LoA by the Contractor [2]. A Bill of Quantities (BQ) [3] is a document created by a quantity surveyor or cost consultant that contains information on certain sections. The sections may include Form of Tender, Information such as the scope of work, Requirements on certain material's quantity, Pricing schedule, Provisional sums, and Day works and the labor costs in a construction project. The BQ is an essential part of the tender. Without a suitable BQ, a tender is incomplete. The contractors can use this information and data to quote rates for their specific work.

*Corresponding Author

The purpose of BQ is to make the tendering process more uniform. Other objectives are to establish a transparent and exact method for valuing the project, to provide a thorough description of the work and its rates, as well as the overall cost. A solid BQ helps quantity surveyors in ensuring that individual contractors have filed valid tenders that comply with the specifications.

The CostX [4] is the offline database system using PostgreSQL. CostX is a Construction Management Software that is designed to serve a construction activity especially for Enterprises or Small Medium Enterprise company. The limitation of CostX is subject to the offline system and it provides the details of each itemize sections for construction in only single sheet. This research aims at solving the manual process of transferring BQ from CostX system into Excel format by providing the automated converting tool for BQ to ease the price generation during tendering activities. The delay in preparing the contract document is due to the lengthy time taken in the process of the Rationalizing of Rates for Bills of Quantities (RoBQ) which are the part of the contract document. The RoBQ needs to be carefully and thoroughly examined, the reasonableness of the tenderer price rates, and the reasonably adjusted price rates will become the BQ contract rates. The BQ contract rates will determine the value of the variation work if there are any variation during the construction period. If any, the value of the variation work will increase the cost of the project delivery and will affect the initial budget approved. Currently, in JKR, the RoBQ process is not yet adopting an information and Communication Technology (ICT), and therefore the process of the RoBQ will be time consuming.

Data integration [5] is the data combination from different sources to help data managers and executives analyze it and make smarter business decisions. In this project, we deploy integration of formatting in CostX system via PostgreSQL database [6] and MS Excel format into a structured MySQL database in generating standardized BQ pricing with a verification and validation from an accountable party for the user in construction company.

II. PROPOSED METHODOLOGY

The process of tendering process comprises of six steps as illustrated in Fig. 1. Step 1 consists of the filling up the BQ using CostX with PostgreSQL database [6] system and converted to Excel macro format that is stored in Cloud database. In step 2, the softcopy of BQ estimation in Excel format is given to the successful tenderers and shortlisted tenderers that is obtained from the tender appraiser officer. Then in step 3, the BQ in Excel format from the estimation department that is awarded to the successful tenderers is uploaded to the J-Selaras system. In step 4, J-Selaras system will create a pricing comparison table. Then in step 5, each price rates for each BQ will be analyzed using the cut-off concepts such as mean, max, average, and standard deviation figures. The adjustment price prior to the cut-off formula will be reviewed by the appointed reviewer and the verified officer in charge in step 6. The J-Selaras system will check for arithmetical error and the adjusted price is matched with the

actual price in the contract tender as stated in the Letter of Acceptance (SST).



Fig. 1. Steps Involved in Data Transfer Process from CostX to Excel Format and the BQ Tender Price Adjustment Rate Determination Process.

III. J-SELARAS DEVELOPMENT MODEL

The J-Selaras converting tool is developed based on filtering algorithm as depicted in Fig. 2.

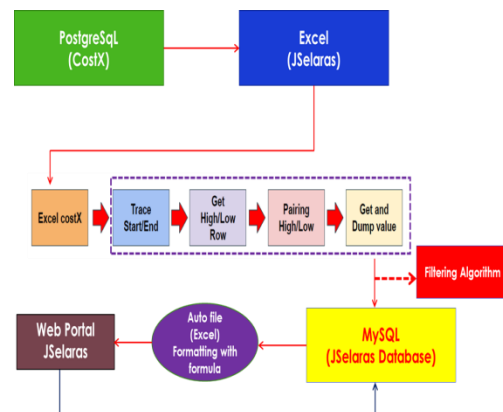


Fig. 2. J-Selaras Model.

The detail specification of BQ is first filled up into CostX system that is built with PostgreSQL database. Then, from CostX system, the BQ specification is downloaded in excel format and converted into J-Selaras system through filtering algorithm by first, tracing the Start and End (Start/End) keywords. Secondly, the cells that cover from Start to the End will be highlighted to get the High row to Low row. Thirdly, those cells involved in High and Low rows will be paired before getting all values involved in all the cells. Refer to next section for details on Filtering algorithm.

A. Filtering Algorithm

Fig. 3 shows the steps in filtering algorithm where the conversion is done from CostX file format into Excel format.

```

Pre-requisite: ∀ s {s: each sheets in Excel}
Start
1. Trace the Start and End Keyword
2. Focus the value from the highest row to the lowest row
3. Pair the value from the highest row to the lowest row
4. Get the value and dump into the MySQL database of J-Selaras system
End
    
```

Fig. 3. Filtering Algorithm.

B. BQ Pricing Standardization

Prior to being processed through filtering algorithm, the Bill of Quantity (BQ) is now converted and saved into MySQL database in J-Selaras system in Cloud environment. The list of BQ price is not yet standardized. The generated price in BQ needs to follow the standardization price that is set by the construction company. All prices of the materials set in BQ is standardized according to certain formula such as Z-Score [7]. The formula involved in Z-Score are Standard Deviation (SD), Min, Max, and cut-off indicates the pseudocode for Z-Score formula used in BQ pricing. The pseudocode for a BQ applying the Z-Score formulation is depicted in Fig. 4.

```

Given the BQ price of items, then
Start
1. Compute the rMin and rMax
2. Analyse the price with cut-off
3. J-Selaras propose the BQ price
3.1 If the proposed BQ price > 115% and < +10%
3.1.1 J-Selaras accept the proposed price
3.2 Else
3.2.1 J-Selaras accept the cut-off price
End
    
```

Fig. 4. Pseudocode in Z-Score Formulation.

Z-score [8-9] is based on the calculation of mean and standard deviation of an attribute. It is a measurement of difference between individual value and the mean population, and then divided by standard deviation of population. The computed, Z-score (Z) provides each feature with a zero mean and a unit variance. The foundation of Z-score is where the mathematical Gaussian curve or ‘Bell Shaped’ curve is applied to the data under study [10-11]. The Z-score, Z as in [12] is expressed as follow:

$$Z\text{- Score} = \frac{x_i - \bar{x}}{s} \tag{1}$$

Where x_i is an individual value, \bar{x} is mean of samples and s is standard deviation of samples. The Z-score technique was proposed in the study to analyze the comparison of each rate in the BQ between all the short-listed tenderer, the successful tenderer and the department’s cost estimation [13]. Meanwhile, the proposed rates in the BQ to be agreed by the successful tenderer will be automatically generated by the system based on the cut-off formula i.e., the rate of an item description is derived from the average rates of the total number of all the short-listed tenderer including the rate of the successful tenderer and the rate of the department’s estimate [14]. The cut-off formula and the Z-score technique are like the tender evaluation system format. The cut-off principle means the lowest acceptable rate to be certified. The tenderer will not be able to complete the project if the rate is too low i.e., the rate is lower than and below the cut-off.

In accordance with statistical methods, the setting of the cut-off rate is based on 'Mean' and 'Standard Deviation' of the shortlisted tenderer rate including the Department’s Estimates (AJ) rates, after rate 'freak' is removed. The range of the cut-off rate is in between not exceeded above +10% (<+10%) and is not exceeded below -15% (<-15%). After all the tenderer price BQ have been uploaded in the J-Selaras, the system will automatically do the analysis based on the setting of the cut-off and the z-score formula in the J-Selaras. Simultaneously, the J-Selaras suggests the contractor rate if the rate is within the cut-off range or suggests the cut-off rate if the contractor rate is out of the range.

In finalizing BQ for the contract document, the appropriate choice of the reasonable rate can be ultimately determined whether to accept the rate of a successful tenderer or the rate generated by the system or the rate of past projects or the current market review rate [15]. The reporting of the analysis is in the tabulation as shown in the Fig. 5. The added value, the J-Selaras system can be as the data cost collection center for price rate and as a reference especially in preparing department’s estimate for new project.

Nama Projek		Analisa Model J-Selaras				
		Amount (RM)	AJ Rate (RM)	JS Rate (RM)	rMin	rMax
H	Vibrated Reinforced Concrete Grade C30/37 or other approved equivalent concrete with minimum cement content 300kg/m3 and maximum free water cement ratio 0.55 as described in	5,100.00	210.00	335.00	336.00	340.00
J	Column stump (All provisional)	22,440.00	210.00	335.00	336.00	340.00
K	Ground beam	46,580.00	210.00	335.00	336.00	340.00
M	Ground floor slab exceeding 100mm and not exceeding 150mm thick	3,400.00	210.00	335.00	336.00	340.00
	Concrete topping not exceeding 100mm thick					

Fig. 5. Illustration of the Comparison and Analysis Table.

IV. EXPERIMENTATION RESULT

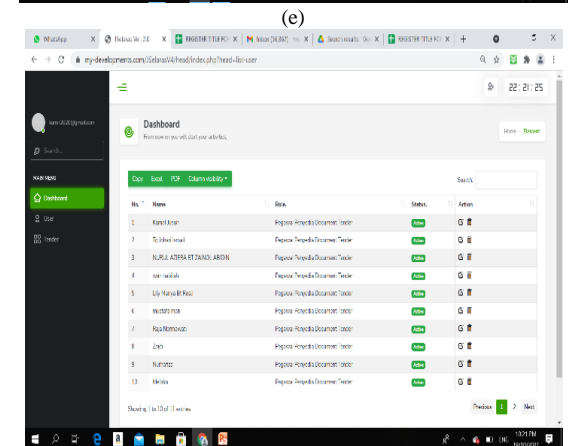
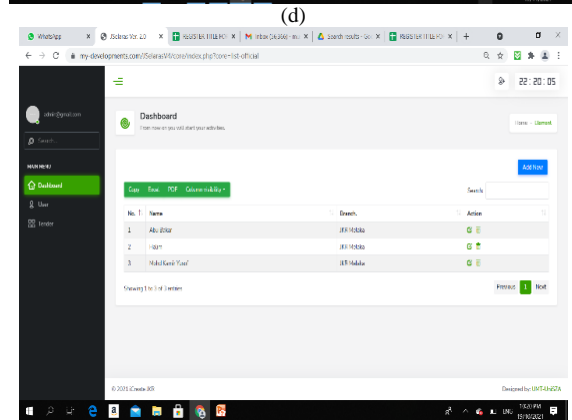
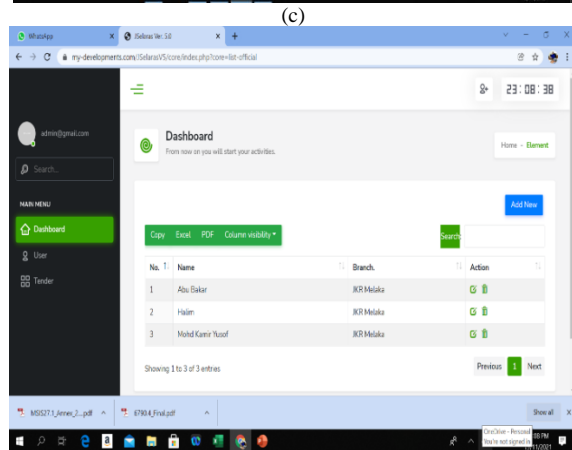
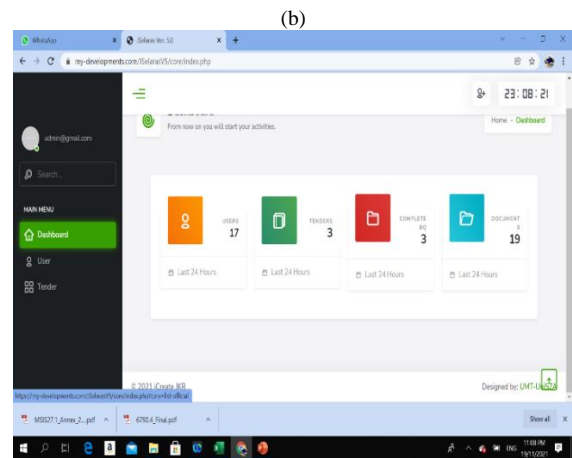
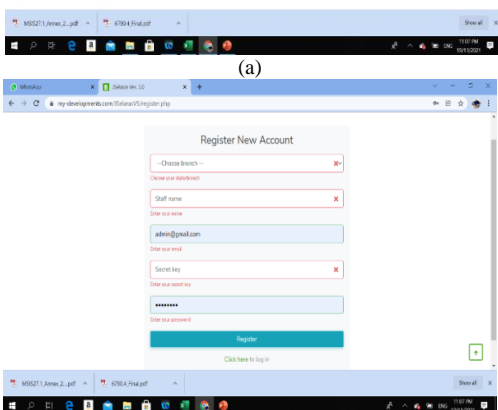
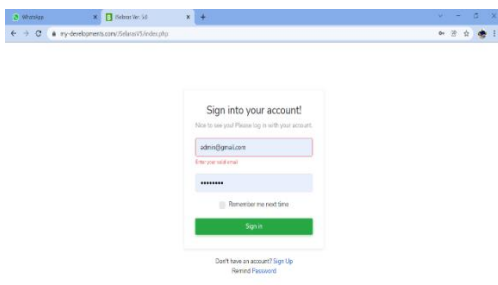
The J-Selaras system is developed under machine specification of VivoBook_ASUS Laptop X415EA_A416EA, Windows 11, i5-1135G7 and 12GB RAM as in Table I.

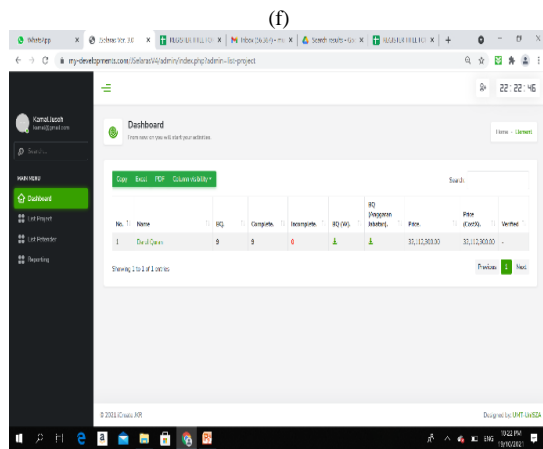
TABLE I. MACHINE SPECIFICATION

System Model:	VivoBook_ASUS Laptop X415EA_A416EA
Operating System:	Windows 11 Home Single Language 64-bit (10.0, Build 22000) (22000.co_release.210604-1628)
Processor:	11th Gen Intel(R) Core (TM) @ 2.40GHz (8 CPUs), ~2.4GHz
Memory:	12288MB RAM
Available OS Memory:	11982MB RAM

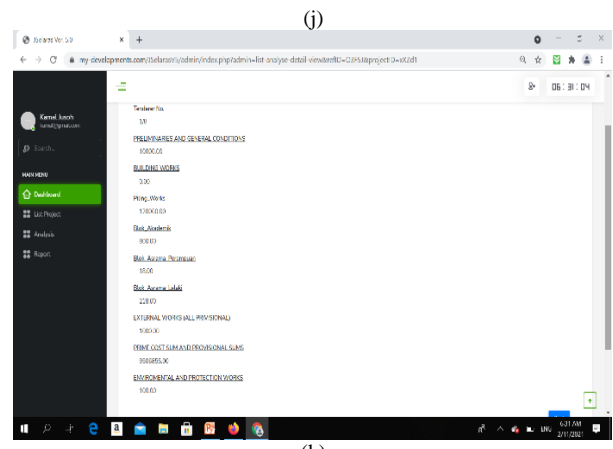
A. Graphical User Interface

J-Selaras consists of seven modules i.e., Sign-up/Login-Logout Module, Registration of User Profile module, Upload-Download of CostX file format module, Excel Standardization format module, Analysis of Price-rate module, Report generation module and Maintenance module. The J-Selaras system is accessible at <https://my-developments.com/JSelarasV5>. Fig. 6(a) depicts the login page of J-Selaras admin while Fig. 6(b) shows the new user registration of J-Selaras system. Once the registration successful, then user is directed to J-Selaras dashboard as shown in Fig. 6(c). Fig. 6(d) indicates the list of J-Selaras users that are currently active while Fig. 6(e) and Fig. 6(f) portray all J-Selaras *admin* users and *Pentadbir J-Selaras* users. Fig. 6(g) to Fig. 6(h) dictate the Darul Quran list of projects and sample of Darul Quran project respectively. Meanwhile, Fig. 6(i) states the BQ that is ready to be downloaded and edited the price and Fig. 6(j) shows the sample of generated BQ price. Finally, Fig. 6(k) illustrates the sample of summarized BQ by J-Selaras.

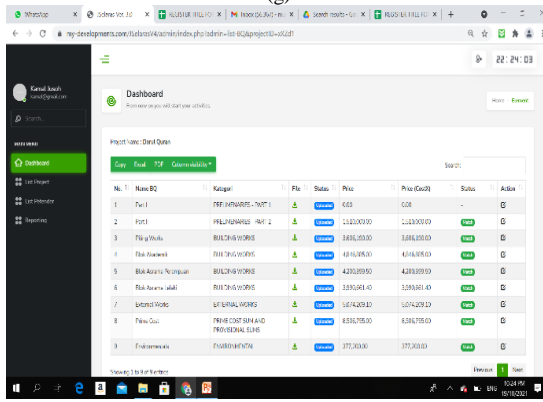




(f)

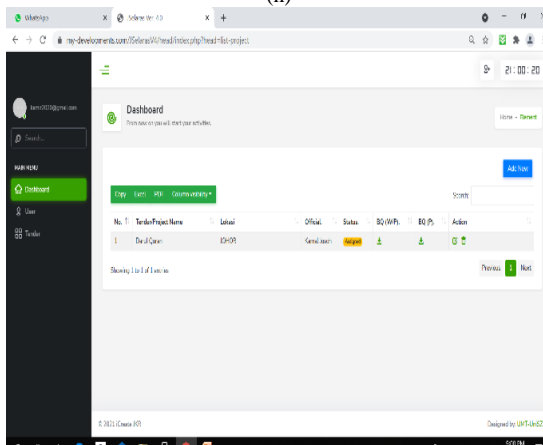


(j)



(g)

Fig. 6. a). J-Selaras Admin Login, b). New Account Registration, c). J-Selaras Dashboard, d). List of J-Selaras Users, e). List of User (Admin), f). List of User (Pentadbir J-Selaras), g). List of Projects (Darul Quran), h). Sample of Details BQ of Darul Quran, i). BQ that is ready to be Downloaded and Edit the Price, j). Sample of BQ Estimated Price Generated, k). Sample of Summarized BQ.



(h)

B. Simulation Result

Table II plots the result of simulation on BQ generation between the manual processes with the automated J-Selaras system.

TABLE II. THE EVALUATION OF THE EXECUTION TIME IN SECOND (S) BETWEEN THE MANUAL BQ GENERATION WITH J-SELARAS BQ GENERATION

Project	File Size (KB)	Manual Time Process (Ms Excel)(s)	Jselaras Tools time Process (s)
BQ for Project A	297	18000	3
BQ for Project B	301	21600	4
BQ for Project C	294	18000	3
BQ for Project D	424	25200	4
BQ for Project E	600	27000	5

The difference of time execution is recorded as the performance measurement in comparing the manual process of using CostX system with the proposed J-Selaras system for BQ generation of Project A to Project B. The processing time is recorded in second and the results from Table II is plotted in Fig. 7.

The recorded data is based on the five projects (Project A, Project B, Project C, Project D and Project E), the Excel file size in kilobyte (KB) and the execution time processing of the five projects. The manual process enquires more than 18000 seconds as compared to J-Selaras system that generates the BQ from three to five seconds only. With automated conversion tool as J-Selaras, we may have 6K times reduction in generating the BQ.

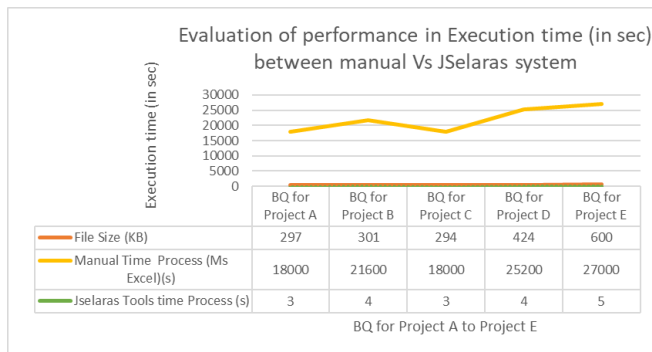


Fig. 7. The Execution Time of BQ Generation between Manual Process using CostX Versus Proposed J-Selaras System.

V. CONCLUSION

The proposed J-Selaras system has improved the time efficiency and eased the manual conversion done by the manpower. Despite the use of this converter (J-Selaras), the degree of error-rate might be reduced because of the automated conversion being done from the source file of Excel format in CostX system. Moreover, once the data is recorded and saved in centralized database such as MySQL in cloud database platform [16-17], the record keeping mechanism is structured for easy saving and retrieving at the admin own pace. The cloud database storage perhaps is resistant to difficulties in maintenance. The J-Selaras system may act as a construction tool which covers all activities in construction management to serve both top-level management as well as the operation level decision makers.

ACKNOWLEDGMENT

We thank to all entity for the contribution towards this paper writing and publications especially Sr. Norisah Abdul Ghani and Sr. Mohd Adza Arshad from Pusat Kecemerlangan Kejuruteraan dan Teknologi Jabatan Kerja Raya Malaysia (CREaTE), Sr. Raja Normawati Raja Ayob, Sr. Kamarul Azhar Mahmood, Sr. Faizul Azwan Ariffin, Sr. Mohamad Dizi Che Kadir, Sr. Lily Mariya Rosli, Pn. Nurhafiza binti Abu Yaziz from Cawangan Kontrak dan Ukur Bahan (CKUB), JKR Malaysia, Associate. Prof. Ts. Dr. Mustafa Man from Universiti Malaysia Terengganu (UMT) for the idea creation and the conceptual design and implementation, Dr. Mohd. Kamir Yusof and Dr. Wan Aezwani Wan Abu Bakar from Universiti Sultan Zainal Abidin (UniSZA) for the system development and system documentation.

REFERENCES

- [1] L.H. Aun. "Malaysia's Shared Prosperity Vision 2030 Needs a Rethink to Make a Breakthrough." 2019.
- [2] Jabatan Kerja Raya Malaysia (JKR). Strategic Plan 2021-2025 from <http://www.jkr.gov.my>. Accessed: 23/08/2022.
- [3] P. A. A. Permadi, N. I. W. Oei, and B. Hasiholan. "Comparative Study in Bill of Quantity Estimates on Reinforcement Works of Pile Cap, Single Pier and Double Pier of Flyover Between Conventional Methods and BIM (Building Information Modelling)." *IOP Conference Series: Earth and Environmental Science*. Vol. 1065. No. 1. IOP Publishing, 2022.
- [4] iTWO costX Estimating Software from <https://www.itwocostx.com/>. Accessed: 20/07/2022.
- [5] Man, Mustafa, Julaily Aida Jusuh, Mohd Shafry Mohd Rahim, and Mohammad Zaidi Zakaria. "Formal specification for spatial information databases integration framework (SIDIF)." *TELKOMNIKA (Telecommunication Computing Electronics and Control)* 9, no. 1, 2011, pp. 81-88.
- [6] What is PostgreSQL? -Definition from Techopedia from <https://www.techopedia.com/definition>. Accessed: 20/07/2022.
- [7] B. Emilian, and C. Darie. "Beginning PHP and MySQL E-Commerce: From Novice to Professional, (Beginners/Beginning Guide)." 2008.
- [8] D. Chappel, "Professional Practice for Architects and Project Managers", Wiley, 2020, pp.183-185, ISBN: 978-1-119-54007-6.
- [9] S. Sudirman, Z. M. Yusof. "Public sector ict strategic planning: framework of monitoring and evaluating process." *Asia-Pacific Journal of Information Technology and Multimedia*, 2017, 6(1). pp. 85-99.
- [10] JKR Rate On-Line System (RATOL), from ratol.jkr.gov.my. Accessed: 20/07/2022.
- [11] R. W. Thatcher, C. J. Biver, & D. M. North. "Z-Score EEG Biofeedback: Technical Foundations." Applied Neurosciences Inc., 2004.
- [12] T. C. Brown, T. C. Daniel, & R. M. Forest. "Scaling of ratings: concepts and methods.", 1990.
- [13] P. Davis, and D. Baccarini. "The use of bills of quantities in construction projects-an Australian survey." In Proceedings of the COBRA 2004 International Construction Research Conference of the Royal Institution of Chartered Surveyors. RICS Foundation, 2004.
- [14] S. Lee, W. Trench, and A. Willis. "Willis's elements of quantity surveying.", John Wiley & Sons; 2011 Mar 1.
- [15] R. A. Rashid, M. Mustapa, and S. N. A. Wahid. "Bills of Quantities-Are they still useful and relevant today." In International conference on construction industry, 2006.
- [16] W. A. W. A. Bakar, M. A. Jalil, M. Man, Z. Abdullah, and F. Mohd. "Postdiffset: an Eclat-like algorithm for frequent itemset mining." *International Journal of Engineering & Technology* 7, no. 2.28, 2018, pp. 197-199.
- [17] W. A. W. A. Bakar, M. Man, M. Man, and Z. Abdullah. "i-Eclat: performance enhancement of Eclat via incremental approach in frequent itemset mining," *Telkommika* 18, no. 1, 2020, pp. 562-570.