# THE SUCCESS FACTORS OF EXTRACT-LOAD-TRANSFORM PROCESS IN DATA INTEGRATION IMPLEMENTATION

Kamsuriah Ahmad[1*], Siti Zarith Sofia Itmin[2]

[1]    Research Center for Software Technology and Management, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Malaysia
       Email: kamsuriah@ukm.edu.my
[2]    Information Technology Department, Government Administrative Center, Putrajaya, Malaysia
       Email: szsrz88@gmail.com
*      Corresponding Author

**Article Info:**

**Abstract:**

The primary purpose of data integration in data warehouse is to consolidate data from various heterogeneous data source. ETL (Extract-Transform-Load) is a traditional approach used to extract, transform and load data from external sources into data warehouse environment. Since there exists limitation on ETL approach, many organization move to ELT (Extract-Load-Transform) approach to improve the loading time and performance. To ensure the success of ELT, then identifying the factors that contribute to this success is important. Yet not many studies have been done in identifying the success factors of ELT approach in organizations so that the process can be performed efficiently and effectively. Therefore identifying the influencing factors that contribute to the success of this process is important. There are various factors proposed in previous studies; however studies focusing on identifying the success factors of ELT approach are missing and this becomes the objective of this paper. Based on the reviews, this study has collected fourteen factors where eight elements are classified under system quality factor and six elements are classified under data quality factors. All these factors will be verified by the domain experts. The success factors proposed in this study could be used as a reference to the system developers in improving the performance of data integration process.

**Keywords:**

## Introduction

Organizations need to access a large amount of business data from various locations to make a comprehensive and accurate decision. Thus, consolidation and sharing of data has become an

important service among the organizations (Mohd Hassan & Ahmad, 2019). This service could be achieved through a data integration process. Since data are accessed from multiple sources, it needs to be integrated and stored in a data warehouse. Data integration is a process of combining data from various locations and stored into a data warehouse environment. The ETL (Extract-Transform-Load) is a key role and a common processing method used in data integration strategies. During this process, business data is extracted from multiple operational databases, transform the data according to the data warehouse rules and regulation, and load it into a single and centralized location. However, due to some unexpected situations, integration processes failed to produce a good output (Jamil & Mohammad Yusof, 2018). Among the cause of failures are due to data integration requirements not being met, critical data is difficult to access, long processing times and the resulting data does not meet the standard quality (Zainon & Singh, 2019). Moreover, ETL is reported to have some shortcoming in terms of cost, process, performance, and continuous improvement (Haryono et. al, 2020), therefore the ELT (Extract-Load-Transform) approach has been used as an alternative (Tayade, 2019). There are many organizations have considered using ELT process because of its performance (Haryono et. al, 2020); (Tayade, 2019). A study has been conducted by comparing the ELT and ETL processes using seventeen parameters, among those are performance, process, cost, easy to use, data availability and others (Haryono, 2020). The survey was conducted with various industries that implemented data warehouse. The results suggest that ELT process is better than ETL based on the parameters used. Even though many organizations have started their initiatives in ELT implementation, unfortunately these initiatives are often fail. Therefore identifying factors that contribute to the success of ETL implementation is urgently needed. The information systems (IS) literature suggests that there are some factors affect the success of a system; however, there are limited discussions on the success factors that affect the ELT implementation in the organizations. These factors can also be used to evaluate the effectiveness of the ELT approach in terms of its compliance with the data processing standards (Mohd Hassan & Ahmad, 2019). Thus, this paper will focus on investigating these key factors. To discuss the achievement of the study; this paper is organized as follows: section II compares the difference between ETL and ELT implementation. Section III discusses the issues of success model in data processing. Section IV investigates the existing underpinning theories for data processing. Section V reviews the factors in contributing to the success of ELT implementation. Based on the limitations faced by the existing methods, an evaluation model is proposed in Section VI. Section VII discusses the findings from this study. Section VIII concludes the study and future work suggestion.

**ETL vs ELT Implementation**

The ETL is a traditional method to integrate and reform data into a unified format. However, in a modern architecture with data warehouse based on cloud technology, the ELT approach has emerged as the newer approach for data gathering and data preparation (Tayade, 2019). Big data phenomena have changed the way how organizations handle huge amounts of data. These data come in various types whether structured, semi-structured, or unstructured. Therefore the data extracted from the sources need to be cleaned, transformed, and enriched before integrating it into one data source. Both ETL and ELT approaches serve the same purpose but the difference is in the implementation process. Basically, these two approaches have three main steps:

- Extract: The data from various sources will be extracted. The data can be in any format such as structured data, unstructured, semi-structured images, emails, or voices.
- Transform: The data will be cleaned, processed and converted into the warehouse format
- Load: The data will be loaded into the warehouse and analysed using any business intelligence and visualization tool.

The difference in these two approaches is on the handling of the data. ETL transform the data from the data source, convert the data into the existing format and load the data into the warehouse. The transformation processes are performed on the data before it is loaded into the warehouse. Figure 1 shows the ETL implementation process.
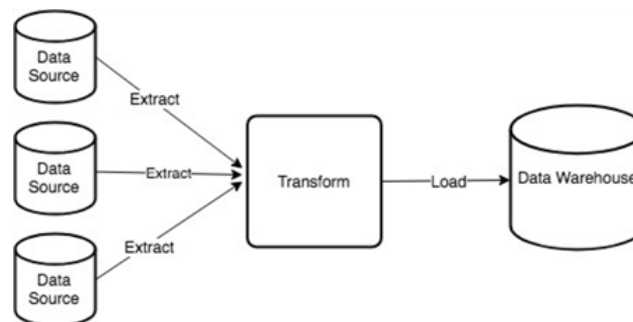


**Figure 1: ETL Implementation Process**

(Haryono et. al, 2020)

ELT offers an improvement process to ETL, where the data is loaded first into the warehouse before the transformation process took place. The data does not have to be loaded into a temporary staging area. The advantage of this process is its flexibility that allows the user to save the data to the target database before the transformation process (Tayade, 2019). This will give access to the information whenever the user wants it. Figure 2 shows the ELT implementation process.
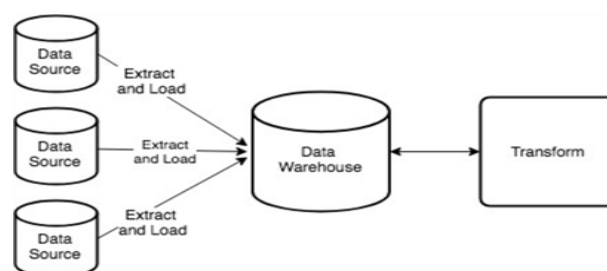


**Figure 2: ELT Implementation Process**

(Haryono et. al, 2020)

Earlier researchers indicate that with the technology advancements, the popularity of ELT has increased (Haryono et al., 2020); (Wass et. al, 2013). These studies claim that the ELT process is able to improve the data integration process in terms of reducing operating costs, low maintenance and improve the performance when compared with ETL, as explained below:
- Low operating cost: ELT stores the data into the warehouse after the data is extracted from the sources. This allows the user to access the data needed for analysis and transformation.

- Performance: The data transformation is done inside the warehouse; this will reduce the loading time.
- Maintenance: ELT is a cloud-based process, hence maintenance is automated.

Therefore the implementation of ELT process in the organizations to overcome these limitations. Even though, some organizations have started to implement it but the success rate is low. Often it is due to technical and non-technical factors. Finding the success factors in implementing ELT is now deemed important and becomes the aim of this paper. To achieve this aim, the steps in the methodology are as follows:

i.   Review the success model and underpinning theories in data processing
ii.  Identifying the contributing factors to the success of ELT implementation
iii. Identifying the evaluation criteria in order to test the level of achievement of each factor
iv.  Propose a model based on the identifying factors
v.   Validating on the suitability of the model in evaluating ELT implementation with domain experts

## Success Model in Data Processing

ELT is a data processing approach and it involves accessing data from various locations and stored in the warehouse. To ensure this process achieves the objectives, it needs to be done effectively and accurately. Therefore it is important to identify the factors and elements that contribute to the success of the ELT process. Earlier researchers have helped in developing robust models for data processing and it could be used to identify the success of ELT (Ibrahim et al., 2021). To investigate further, this study refers to success model in data processing as an underpinning theory. During the ELT process, it is important to ensure the processing system achieved its standard quality and the datasets generated after the processing are of good quality. ELT processes reach its quality level if the same methods, requirements and controls are applied consistently (Azeroual et al., 2019). This level of system quality can be outlined as best practices and can serve as a guideline in the ELT implementation. Producing a quality data after the ETL process is also important since this data will be used by the management to make decisions. Data is said to be quality if the data is accurate, complete, consistent, reliable and updated (Jaya et al., 2017). Quality data is easier to process and analyze by other users, such as database systems, data warehouse visualizations, or systems analysis. Therefore it is important to consider data quality and system quality factors in determining the success of ELT approach.

## Underpinning Theories for Data Processing

Since producing quality system and quality data is important to ELT process, therefore investigating other sub-factors that relates with these factors are needed. This study has reviewed four underpinning theories in Information Systems (IS) as the basis in developing a success model for ELT implementation.

i.   DeLone and McLean have proposed the Information Systems success model (Delone & McLean, 2003). This model is among the most popular theory and been referred by many researchers. It states that the user satisfaction and the intention to use the system are influenced by three quality features which are: information quality, system quality, and service quality. Even though, this theory could be referred when investigating the success factors for ETL implementation, however, this model is developed to evaluate the success of IS implementation; the focus on evaluating ELT process might be different. Studies

done by many researchers (Al-Mamary et al., 2014); (Sabeh et al., 2021); (Urbach & Müller, 2012) discovered various constraints in DeLone and McLean evaluation model. Thus, this model has been refined and improved by other researchers to be used in other disciplines and perspectives (Yakubu & Dasuki, 2018); (Ojo, 2017).

ii. The success model proposed by (Yang & Wu, 2015) consists of three influencing factors for the success of data processing during information sharing process namely Information Quality, System Quality and Service Quality. Each factor has certain elements in studying the impact of information sharing process. Information Quality factor has eleven elements of success namely consistency, accuracy, timeliness, reliability, completeness, comprehensibility, real-time, variation, relevance, thoroughness and accessibility. System quality factor has eight elements namely availability, usability, capability, reliability, maintenance, security, compatibility and responsibility. The Service Quality factor has three elements, namely responsibility, assurance and communication. The researchers argued that these three factors and its elements are important in evaluating the effectiveness of data processing in the organization. Sharing of information relies on trusted social network, data exchange and interconnection of information systems in the organization. This study argued that to enhance the public system services, these three quality factors are needed in determining the effectiveness of cross-boundary information sharing.

iii. The model proposed by (Wixom & Watson, 2001) consists of seven elements that influence the success of a data warehouse project. These factors are management support, priorities, resources, user participation, resource systems, technology development and team skills. These seven elements are then group under three factors namely: organizational support, project implementation, and technical success. These three factors strongly influence the success of data warehouse implementation which implies the system quality and data quality of the warehouse. This study verified that there exist strong relationships between system quality and data quality factors in determining the success of data warehouse projects.

iv. The quality model proposed by (Azeroual et al., 2019) is aims to measure the research information during ETL process. The process of extracting the data from the sources and transformed it in the research information system must be guarded securely. He claims that the success in data integration is highly depends on information quality and data quality which are determined by the extraction and transformation process. This study emphasizes the quality criteria that need to be achieved during the ETL process are completeness, correctness, consistency and timeliness. These quality criteria ensure the data able to be used in generating new information for the benefits of the organization.

Based on the reviews, these four models have a particular focus and aim; ironically identifying the success factors in the implementation of ELT data processing is missing. This indicates that lack of studies conducted in identifying the factors of ELT success. The factors used to measure the success of ELT process can also be used as evaluation factors based on the criteria that need to be achieved as well as comply with data processing standards. Therefore, the purpose of the study is to identify the factors and elements of evaluation in determining the successful implementation of ELT process in the organization. Identifying these factors and elements is important so that the data integration implementation is more efficient and effective while services among organizations could be improved. Based on the existing studies, there are four main factors that influence the success of data processing implementation. These factors are

information quality, system quality, service quality and data quality. Table 1 outlines these factors and the models that proposed them.

**Table 1: The Four Quality Factors**

| Authors | Information Quality | System Quality | Service Quality | Data Quality |
|---|---|---|---|---|
| Azeroual et al., 2019 | | | | √ |
| Jaya et al., 2017 | | | | √ |
| Yakubu & Dasuki, 2018 | √ | √ | √ | |
| Ibrahim et al., 2021 | | √ | | √ |
| Delone & McLean, 2003 | √ | √ | √ | |
| Yang & Wu, 2015 | √ | √ | √ | |
| Edmondson et al., 2019 | | √ | | √ |
| Theodorou et al., 2015 | | √ | | √ |

Table 1 indicates that the data quality and system quality factors have a high frequency where many researchers agree that these two factors are important in evaluating the success of ELT data processing implementation (Yang & Wu, 2015); (Haryono, 2020); (Al-Mamary et al., 2014). However, information quality and service quality factors are also important but when considering ELT data processing, system quality and data quality factors are more prominent. The quality system factor is able to ensure the hardware and software usage is effective and efficient during data processing. While the quality data factor ensures the output from ETL processing is accurate, complete and meets the user requirements. Therefore this study will take into account these two factors and identify the sub-factors that contribute to the success of ELT data processing.

**Contributing Factors to the Success of ELT Implementation**
ELT has the potential to improve the ETL approach (which is the traditional method of data processing during data integration), owing to its ability to improve the data processing in terms of cost, better performance and low maintenance (Haryono, 2020). To ensure successful implementation of the integration process, data access between application systems must be done effectively and accurately. Therefore, it is important to identify data integration requirements to ensure this process is successfully implemented. The requirement can be determined by identifying the factors that lead to its success. These factors can be used as a yardstick in evaluating the success of data integration process (Kanmani & Ezhilarasi, 2017); (Zellal & Zaouia, 2017).

Identifying success factors and used these factors in evaluation is an effective approach where it has been used widely in various domains and studies (Mohd Hassan & Ahmad, 2019); (Ibrahim et al., 2021). However, reviews indicate that not many studies have been conducted in identifying factors in evaluating the success of ELT implementation in organizations (Haryono, 2020). During ELT data processing, it is important to ensure the system achieves its quality level during processing and the resulting data is of good quality (Cai & Zhu, 2015); (Mukherjee & Kar, 2017). This study refers to fourteen existing data processing models where

all of these models considered system quality and data quality factors in identifying the elements of success for data processing. Table 2 displays the evaluation elements used by past researchers in evaluating the success of data processing.

From the reviews, fourteen elements are identified where eight elements are classified under system quality factor and six elements are classified under data quality factors. These elements are identified as contributing factors to the success of ELT data processing and could be used in evaluating whether ELT meets system quality and data quality standards. The descriptions of these elements are as follows:

*Availability*
System must be readily available to users or application when and where they need it. The warehouse system is said to achieve its quality if the ELT process is able to be used whenever needed by the user. Studies conducted by (Al-Mamary et al., 2014), (Yang & Wu, 2015); (Edmondson et al., 2019); (Theodorou et al., 2015) state that the availability factor is important in determining system quality.

*Responsive*
Responsive refers to the ability of the system to complete the tasks within a given timeline. The warehouse system is said to achieve its quality if the ELT process is able to response to the user requests within a given period of time. Studies conducted by (Al-Mamary et al., 2014); (Yang & Wu, 2015); (Wixom & Watson, 2001) state that the responsive factor is important in determining system quality.

**Table 2. The Proposed Evaluation Elements**

| Author | System Quality | | | | | | | | Data Quality | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Availability | Responsive | Reliability | Usability | Capability | Compatibility | Safety | Maintenance | Perfection | Validity | Consistent | Accuracy | Integrity | Timeliness |
| Azeroual et al., 2019 | | | | | | √ | | | √ | | √ | √ | | √ |
| Schmidt et al., 2021 | √ | | √ | | √ | | | | √ | | √ | √ | √ | |
| Ibrahim et al., 2021 | | | √ | √ | √ | | √ | √ | | √ | | | √ | |
| Yakubu & Dasuki, 2018 | √ | | √ | √ | | | | | √ | | | √ | | √ |
| Jaya et al., 2017 | | | | | | | | | | √ | √ | √ | | √ |
| Al-Mamary et al., 2014 | √ | √ | √ | √ | | | | | √ | | √ | √ | | |
| Yang & Wu, 2015 | √ | √ | √ | √ | √ | √ | √ | √ | | | √ | √ | | √ |
| Edmondson et al., 2019 | √ | | | √ | | | √ | | | | | √ | | √ |
| Theodorou et al., 2015 | √ | | | √ | | | √ | √ | | | √ | √ | √ | √ |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Kanmani & Ezhilarasi, 2017 | | | | | √ | √ | | | √ | √ | √ | √ | √ | |
| Zellal & Zaouia, 2017 | | | | | √ | √ | | | √ | | √ | √ | | √ |
| Cai & Zhu, 2015 | √ | | √ | √ | | | | √ | | | √ | √ | √ | √ |

### Reliability

Refers to the ability of the system to execute its function without failure at the given time interval with stated conditions. The warehouse system is said to achieve its quality if the ELT process is able to perform its tasks without fail according to the given time frame and specific requirements. Studies conducted by (Al-Mamary et al., 2014); (Yang & Wu, 2015); (Wixom & Watson, 2001); (Cai & Zhu, 2015); (Ibrahim et al., 2021) state that the reliability factor is important in determining system quality.

### Usability

Usability refers to the ability of the system to enable its users to perform tasks safely, effectively, and efficiently. Users are able to achieve their objectives through the use of the system. The warehouse is said to achieve its quality if the ELT process enable the users to perform their tasks efficiently and they are satisfied with the services offered. Studies conducted by (Al-Mamary et al., 2014), (Yang & Wu, 2015), (Wixom & Watson, 2001) Ibrahim et al. (2021); (Edmondson et al., 2019), (Cai & Zhu, 2015) state that the usability factor is important in determining system quality.

### Capability

System capability refers to the ability of a system to perform a particular task or to achieve a specific outcome. The warehouse is said to achieve its quality if the ELT process capable to perform all the required tasks in warehouse as outlined in the methodologies and produce the outcomes that enable the managers to make decisions. Studies conducted by (Yang & Wu, 2015); (Kanmani & Ezhilarasi, 2017); (Zellal & Zaouia, 2017) state that the capability factor is important in determining system quality.

### Compatibility

System compatibility refers to the ability of a system to work with other systems. The main challenge in the warehouse system is to overcome the compatibility or heterogeneous issues among the systems involved. The warehouse is said to achieve its quality if the ELT process is compatible and able to work with other systems. Studies conducted by (Azeroual et al., 2019); (Yang & Wu, 2015); (Kanmani & Ezhilarasi, 2017); (Zellal & Zaouia, 2017) state that the compatibility factor is important in determining system quality.

### Safety

System safety refers to the ability of a system to ensure the tasks are performed in the safest condition and free from any risks. The warehouse is said to achieve its quality if the system is able to secure the transaction process that may harm the data (Gabriel et al., 2020). Studies done by (Yang & Wu, 2015); (Edmondson et al., 2019); (Theodorou et al., 2015) claim that the safety factor is important in determining system quality.

## Maintenance

System maintenance purposes are to sustain the capability of a system to provide a good service. Maintenance activities are usually done after the evaluation process. This activity helps the organization to identify the effectiveness of the implemented system. Maintenance process involves improving the system through monitoring, evaluating, and modifying the system used. In a warehouse system, maintenance involves processes and methodologies to represent changes made to data and schema of data warehouse application. The warehouse is said to achieve its quality if the system is able to sustain the capability of providing good service through consistent maintenance activities. Studies done by (Yang & Wu, 2015), (Theodorou et al., 2015); (Cai & Zhu, 2015) claim that the maintenance factor is important in determining system quality.

## Perfection

Perfection in data quality aims to reduce the number of defects in the output during data processing and to maximize the organization profit. Data extracted from various sources during the ELT process should be comprehensive, perfect and useful for decision making. Data perfection is important in determining data quality and ensures data meets the standards in data processing (Azeroual et al., 2019); (Al-Mamary et al., 2014).

## Validity

Validity is the process to ensure the data is correct, accurate and useful before the data is stored in the database. During ELT process, the data will be cleansed to ensure the quality of the data before it is stored in the warehouse. Studies done by (Jaya et al., 2017); Ibrahim et al. (2021) (Wixom & Watson, 2001); (Kanmani & Ezhilarasi, 2017) claim that the validity factor is important in determining data quality.

## Consistent

Consistency refers to the measurement of synchronicity of the data stored across different systems. This is a particular concern when data is aggregated from various sources. Discrepancies in the data will create inaccurate and unreliable data. Studies done by (Azeroual et al., 2019); (Jaya et al., 2017); (Al-Mamary et al., 2014) claim that the consistency factor is important in determining data quality.

## Data Accuracy

Data accuracy refers to the data stored having correct values without errors and can be referred as reliable information. Data must be uniform and comply with terms and conditions. The data values used in the warehouse and the other systems must be consistent and unambiguous. Studies done by (Azeroual et al., 2019), (Jaya et al., 2017); (Al-Mamary et al., 2014); (Yang & Wu, 2015) claim that the accuracy factor is important in determining data quality.

## Integrity

Data integrity is the assurance of data accuracy and consistency over the data life-cycle. To maintain the integrity, data needs to be recorded exactly as intended and prevent unintentional changes of information to the data. During ETL process, maintaining the integrity and ensuring the completeness of data is essential. Studies done by (Theodorou et al., 2015); (Kanmani & Ezhilarasi, 2017); (Cai & Zhu, 2015) claim that the integrity factor is important in determining data quality.

*Timeliness*

Timeliness refers to availability and accessibility of data when it is expected and needed. It can be measured as the time between data is expected to access and when it is readily available for use. Clean and well-structured data are important for the organization to make decision. The success of applications that relies on data depends on its consistency and timely information. Data or information that is not timely will lead organization in making wrong decision (Machado et al., 2019). Studies done by (Azeroual et al., 2019), (Jaya et al., 2017); (Al-Mamary et al., 2014); (Yang & Wu, 2015) claim that the timely factor is important in determining data quality.

The proposed fourteen elements as described in Table 2 can be used to evaluate the impact of ELT process in achieving data processing standards. Evaluation criteria are used to test whether the ELT process successfully achieves the standards. The ELT process is said to be in comply with the data processing standards if the test is successful. Table 3 shows these elements and the criteria in evaluating the ELT process.

**Table 3: The Evaluation Elements and Criteria**

| Factor | Evaluation Element | Evaluation Criteria |
|---|---|---|
| System Quality | Availability | The system should be available 24 hours and should always accessible for use. |
| | Responsive | The system reaction time used is acceptable within set expectations and system feedback is achieved; efficient operating performance. |
| | Reliability | No missing data during data transmission |
| | Usability | The system should be real time and able to accommodate large amounts of data. |
| | Capability | The system needs to be compatible with other systems to enable data transmission |
| | Safety Perfection | The system should be able to transmit data securely over the network and should protect the confidentiality of data. |
| | Compatibility | The system should be able to resolve inequality issues by using latest technology |
| | Maintenance | The level of effort and cost required to keep the system running smoothly. |
| Data Quality | Perfection | The data is sufficient and usable |
| | Validity | The data is valid and verified. |
| | Accuracy | Data is correct, updated and according to the format. |
| | Consistent | Data is consistent and the relationship between entities and attributes are also consistent |
| | Integrity | Data records and attributes are verified and referred. |

| | Timeliness | Data transmission and data reception are according to a set schedule |
|---|---|---|

## The Proposed Conceptual Model

The proposed factors and elements are used as a guide in building a conceptual model of ELT implementation evaluation based on the underpinning theories. The system quality and data quality factors are an exogenous variable which is an external variable that affects or impacts the evaluation of data processing in data warehouse. While ELT success is an endogenous variable which is an internal variable that affects the dependent variable. This model states that the success of data integration is the result of the successful implementation of ELT process. Based on the understanding of the theory, the evaluation model is constructed as shown Figure 3.
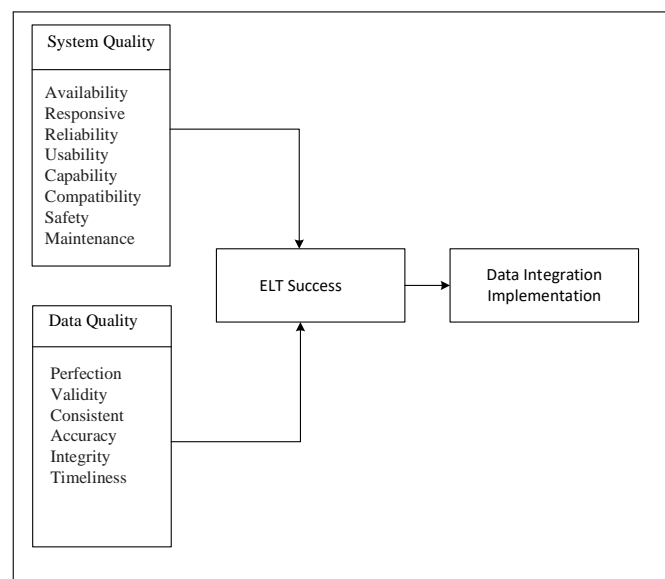


**Figure 3: The Proposed Conceptual Model**

## Model Verification and Validation

The developed model needs to be verified and validated in terms of its functionality. Two approaches are used which are quantitative and qualitative. Qualitative methods are used to test the effectiveness of the relationship between factors. While qualitative methods are used to confirm the appropriateness of the proposed model. Through quantitative methods, a set of questionnaires was developed and tested for effectiveness in supporting the proposed model. Based on the proposed factors, 42 questions are constructed and each item is tested using Cronbach's coefficient Alpha. The results of this test show that the instrument used is in the excellent category with a coefficient rate exceeding 0.90. This questionnaire was then distributed to 28 respondents who were directly involved in the development of data integration. Descriptive analysis, factor analysis and residual correlation values are used to analyze the collected data. The result of these tests shows that the proposed factors have significantly influenced the success of ELT implementation in the organizations.

To validate the developed model three evaluators were appointed. This effort was to obtain views and comments on the appropriateness of the factors and elements used in determining the success of ELT implementation. These evaluators are individuals who are experienced and

involved in data warehouse development. The profile of each evaluator is listed in Table 4. The interview method was used to obtain the views of the evaluators. Several questions are constructed based on the design of the model and the suitability of the factors and elements in the formation of the model. The evaluators agreed that the proposed factors and elements were appropriate and confirmed that the model can be used in evaluating ELT process. Through Cohen Kappa (K) analysis, the K value obtained is 1.00 which indicates that the evaluator agreement is very strong. According to Denzin & Lincoln (2003) if the Kappa agreement value reaches 1.00, the agreement obtained is very strong and if it is related to a model construction then the model is robust.

**Table 4: Evaluators Profile**

| Evaluator | Position | Working Experiences |
|---|---|---|
| Evaluator 1 | Data warehouse developer, public sector | 18 years |
| Evaluator 2 | Data Warehouse developer, public sector | 20 years |
| Evaluator 3 | Data Warehouse developer, public sector | 15 years |

**Discussion**

The data extracted from various sources needs to be evaluated and cleaned before being transformed into a warehouse. Data needs to be cleaned in terms of errors (such as data duplication, inconsistent data, missing data, and others) which are an important step during the ELT process. Only data that meets the quality such as consistent, accurate, integrity, timely, perfect and valid are stored in the warehouse, which will be used by the organization to make decision. The quality of the data integration processing system, which includes software and system components must be technically sound. The system quality standard such as availability, responsive, reliability, usability, capability, compatibility, safety and maintenance must be met. This study has proposed two important factors in evaluating the effectiveness of ELT processes which are data quality and system quality, and another fourteen evaluation elements as sub-factors of these two factors. This study emphasizes that these are the influence factors and sub-factors that contributes to the success of ELT process and the data integration implementation in the organizations. The evaluation criteria are designed based on the proposed factors whereby the ELT process is said to comply with the data processing standards if it is able to meets these criteria. An evaluation model is developed based on the extension of underpinning theories in IS. The experts have confirmed that the proposed model is suitable to be used in evaluating the effectiveness of ELT processes.

**Conclusion**

Nowadays, most organizations use data to exchange, transform or share the data to make better decision. To maximize the data potential, organizations need to establish ways to control the quality of the data and the systems and remove inconsistencies or errors that may reduce the organizational performance. This study has proposed an evaluation model which is the extension of the IS existing theories through the investigation of ELT implementation evaluation factors. The current success model cannot be used to evaluate ELT process without some modification. For example, this study found two factors namely data quality and system quality are important in evaluating the effectiveness of ELT process during data warehouse

implementation. Fourteen elements have been proposed to further explain the data quality and system quality for the data warehouse. If the ETL process produces quality data then the resulted output is perfect, valid, accurate, consistent, integrity, and timely. For future works, there is various measurement elements shall be explored such as information quality, service quality or other technical aspects. Further investigation is needed in identifying the evaluation elements from these aspects. However, this issue may vary as the infrastructure of the project is highly dependent on the organization's practices.

**References**

Al-Mamary, Y. H., Shamsuddin, A. & Nor Aziati, A.H. (2014). The relationship between system quality, information quality, and organizational performance. *International Journal of Knowledge and Research in Management & E-Commerce*, 4(3), 7-10.

Azeroual, O., Saake, G. & Abuosba, M. (2019). ETL best practices for data quality checks in RIS databases, *Informatics*, 6, 1-10.

Cai, L. & Zhu, Y. (2015). The challenges of data quality and data quality assessment in the big data era. *Data Science Journal*, 14(2), 2015, 1-10.

Delone, W. & and McLean, E. (2003). The DeLone and McLean model of information systems success: a ten-year update, *Journal of Management Information Systems,* 19(4), 9-30.

Denzin, N. K. & Lincoln, Y. S. (2003). Introduction: The discipline and practice of qualitative research". In N. K. Denzin & Y. S. Lincoln (Eds.), *Strategies of qualitative inquiry* (2nd Ed.), Sage, 1 - 45.

Edmondson, M., McCollum, W.R., Chantre. M & Campbell, G. (2019). Exploring critical success factors for data integration and decision-making in law enforcement, *International Journal of Applied Management and Technology*, 18(1), 1-16

Gabriel, J., Mayzira, A., Aditya, J., Itsari, M., Satrio, S., & Ruldeviyani, Y. (2020). Critical success factors of data integration on digital human capital information system to support digital transformation - A case study at PTXYZ, *8th International Conference on Cyber and IT Service Management (CITSM)*, 1-7.

Haryono, E. M., Fahmi, A. S. Tri W, I. Gunawan, A. Nizar Hidayanto & Rahardja, U. (2020). Comparison of the E-LT vs ETL method in data warehouse implementation: a qualitative study. *International Conference on Informatics, Multimedia, Cyber and Information System* (ICIMCIS), 115-120.

Ibrahim, A., Mohamed, I. & Mohd Satar, N.S. (2021). Factors influencing master data quality: a systematic review, *International Journal of Advanced Computer Science and Applications*, 12(2), 181-192

Jamil, A. & Mohammad Yusof, Z. (2018). Information security governance framework of Malaysia public sector. *Asia-Pacific Journal of Information Technology and Multimedia*, 7(2), 85 – 98.

Jaya, I., Sidi, F., Ishak, I., & Jabar, M. A. (2017). A review of data quality research in achieving high data quality within organization, *Journal of Theoretical and Applied Information Technology*, 95(12), 2647-2657

Kanmani, D. & Ezhilarasi, G. (2017). Data quality problems in data warehousing with descriptive classification, *Imperial Journal of Interdisciplinary Research (IJIR)*, 3(7), 694-701.

Machado, G.V., Cunha, Í. & Pereira, A.C.M. (2019). DOD-ETL: distributed on-demand ETL for near real-time business intelligence, *Journal of Internet Services and Applications*, 10(21), 1-15.

Mohd Hassan, N. H. & Ahmad, K. (2019). A review on key factors of data integration implementation in public sector. *International Conference on Electrical Engineering and Informatics (ICEEI),* 13-18.

Mukherjee, R & Kar, P. (2017). A comparative review of data warehousing ETL tools with new trends and industry insight. *IEEE 7th International Advance Computing Conference (IACC)*, 943-948.

Ojo, I. A. (2017). Validation of the DeLone and McLean information systems success model", *Healthcare Informatics Research*, 23(1), 60-66.

Sabeh, H. N., Husin, M. H., Kee, D. M. H, Baharudin, A. S. & Abdullah, R. (2021). A systematic review of the DeLone and McLean model of information systems success in an E-Learning context (2010–2020), *IEEE Access*, vol. 9, 81210-81235.

Schmidt, C.O., Struckmann, S., Enzenbach, C., Reineke, A., Stausberg, J., Damerow, S., Huebner, M., Schmidt, B., Sauerbrei, W., & Richter, A. (2021). Facilitating harmonized data quality assessments. A data quality framework for observational health research data collections with software implementations in R. *BMC Medical Research Methodology*, 21:63

Tayade, D. M. (2019). Comparative study of ETL and E-LT in data warehousing. *International Research Journal of Engineering and Technology* (IRJET), 6(6), 2803-2807

Theodorou, V., Abelló, A., Lehner, W. & Thiele, M. (2015). Quality measures for ETL processes: from goals to implementation, *Journal of Concurrency and Computation Practice and Experience*, 28(15), 3969-3993.

Urbach, N. & Müller B. (2012). The updated DeLone and McLean model of information systems Success. In: Dwivedi Y., Wade M., Schneberger S. (eds) *Information Systems Theory*. Integrated Series in Information Systems, vol 28. Springer, New York, NY.

Waas, F., Wrembel, R., Freudenreich, T. & Furtado, P. (2013). On-demand ELT architecture for right-time BI: extending the vision. *International Journal of Data Warehousing and Mining*, 9(2), 21-38

Wixom, B.H. & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly*, 25(1), 17-41.

Yakubu, M. N. & Dasuki, S. I. (2018). Assessing eLearning systems success in Nigeria: An application of the DeLone and McLean information systems success model. *Journal of Information Technology Education: Research*, volume 17, 182-202.

Yang, T.M. & Wu, Y.J. (2015). Exploring the effectiveness of cross-boundary information sharing in the public sector: the perspective of government agencies, *Information Research*, 20(3), no.685.

Zainon, Z. & Singh, D. (2019). "Review of interoperability practices for Enterprise Information System (EIS) in public sector", *Asia-Pacific Journal of Information Technology and Multimedia*, 8(2), 41-70.

Zellal, N. & Zaouia, A. (2017). An examination of factors influencing the quality of data in a data warehouse, *International Journal of Computer Science and Network Security*, 17(8), 161-169.