

ETL Maturity Model for Data Warehouse Systems: A CMMI Compliant Framework

Musawwer Khan¹, Islam Ali¹, Shahzada Khurram², Salman Naseer³, Shafiq Ahmad⁴, Ahmed T. Soliman⁴, Akber Abid Gardezi⁵, Muhammad Shafiq^{6,*} and Jin-Ghoo Choi⁶

¹Department of Computer Science, COMSATS University Islamabad, Wah Cantt, 47000, Pakistan

²Faculty of Computing, The Islamia University of Bahawalpur, Bahawalpur, 63100, Pakistan

³Department of Information Technology, University of the Punjab Gujranwala Campus, Gujranwala, 52250, Pakistan

⁴Industrial Engineering Department, College of Engineering, King Saud University, P.O. Box 800, Riyadh, 11421, Saudi Arabia

⁵Department of Computer Science, COMSATS University Islamabad, Islamabad, 45550, Pakistan

⁶Department of Information and Communication Engineering, Yeungnam University, Gyeongsan, 38541, Korea

*Corresponding Author: Muhammad Shafiq. Email: shafiq@ynu.ac.kr

Received: 17 January 2022; Accepted: 09 June 2022

Abstract: The effectiveness of the Business Intelligence (BI) system mainly depends on the quality of knowledge it produces. The decision-making process is hindered, and the user's trust is lost, if the knowledge offered is undesired or of poor quality. A Data Warehouse (DW) is a huge collection of data gathered from many sources and an important part of any BI solution to assist management in making better decisions. The Extract, Transform, and Load (ETL) process is the backbone of a DW system, and it is responsible for moving data from source systems into the DW system. The more mature the ETL process the more reliable the DW system. In this paper, we propose the ETL Maturity Model (EMM) that assists organizations in achieving a high-quality ETL system and thereby enhancing the quality of knowledge produced. The EMM is made up of five levels of maturity i.e., Chaotic, Acceptable, Stable, Efficient and Reliable. Each level of maturity contains Key Process Areas (KPAs) that have been endorsed by industry experts and include all critical features of a good ETL system. Quality Objectives (QOs) are defined procedures that, when implemented, resulted in a high-quality ETL process. Each KPA has its own set of QOs, the execution of which meets the requirements of that KPA. Multiple brainstorming sessions with relevant industry experts helped to enhance the model. EMM was deployed in two key projects utilizing multiple case studies to supplement the validation process and support our claim. This model can assist organizations in improving their current ETL process and transforming it into a more mature ETL system. This model can also provide high-quality information to assist users in making better decisions and gaining their trust.

Keywords: ETL; maturity model; CMMI; data warehouse; maturity model



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1 Introduction

BI is defined as getting the right information to the right people at the right time. The term encompasses all the capabilities required to turn data into intelligence that everyone in an organization can trust and use for more effective decision-making. A DW system is a large pool of data collected from different data sources to guide management decisions. DW is a key component of the BI system. ETL a core process of the DW brings data from different types of data sources into the target DW. The importance of accurate and timely information to BI and improving data quality has become a top management priority. The data integration through an ETL process brings up a reliable and consistent view of the organization to ensure better decisions.

Fig. 1 depicts the ETL process for bringing data from various sources into the DW, which serves as the foundation for the BI system. The quality of the ETL process enhances the data quality and, as a result, the information quality. Resultantly, the quality of overall BI systems improves allowing for better decision-making. An enterprise BI model based on process, technology, and organization has been proposed in [1]. The model was verified by four different companies to assess its maturity. Maturity models like CMMI play an important role to achieve higher levels of maturity for the development process and product quality [2]. The Spruit & Sacu Data Warehouse Capability Maturity Model (DWCMM) was used to examine the National Narcotics Board’s DW maturity and to provide an assessment to the DW working group team. The results demonstrated that numerous improvements are required to raise the DWCMM maturity level [3].

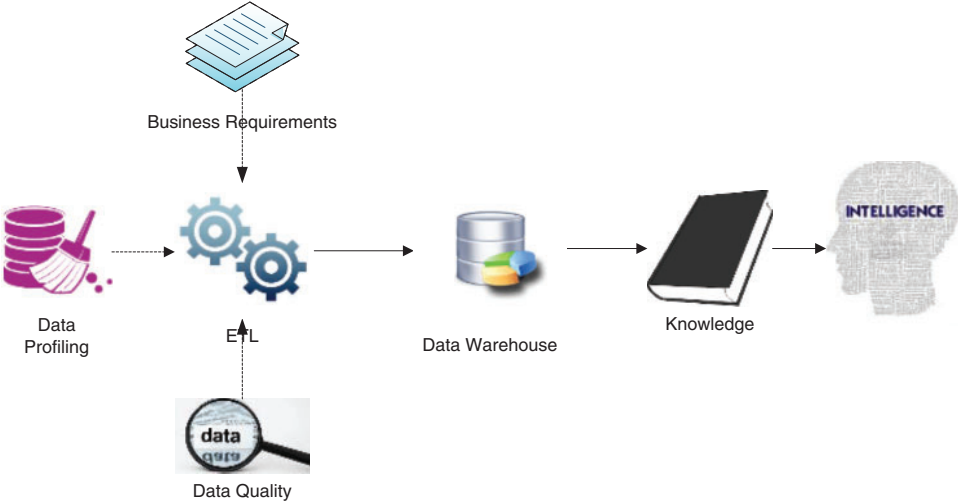


Figure 1: Traditional ETL process in the context of DW and BI

1.1 Extraction, Transformation and Load (ETL)

Extraction is the first step in the ETL process. In this step, data from multiple source systems is extracted into the staging area, which can be in various formats such as relational databases, XML, No SQL, and flat files. The cleansing and consolidation of data that may be required to prepare it for analysis are referred to as transformation. Data is extracted and moved to a staging area, where it is transformed before being loaded into the warehouse. Finally, the data is loaded into the target DW to be further used by OLAP, Data Mining tools, or other BI applications. The load stage depends mainly on what you want to do with the data after it is shifted into the DW.

1.2 Capability Maturity Model Integration (CMMI)

CMMI framework has been used to measure process capacity and organizational maturity all around the world since 1993 [4]. CMMI is a well-known process improvement framework with five maturity stages. Five tiers are made up of 22 process areas. Each process area has well-defined goals that are met through a collection of associated best practices. According to Kimball et al. [5], in the development of DW systems, a major effort is consumed during the execution of the ETL process. Hence the maturity of DW systems mainly depends on the maturity of the ETL process. Nevertheless, we hardly find research that addresses the maturity of this core area of DW systems. We propose a maturity model for ETL to address this issue. The key contributions of this paper are summarized as follows.

- The proposed EMM would help businesses evaluate and enhance the quality of their present ETL process by adhering to the CMMI, a globally recognized process improvement framework.
- The EMM improves the worth of information provided, allowing users at all levels of an organization to make better decisions and increase their trust.
- The EMM transforms a simple ETL process into a resilient and improved ETL system by employing KPAs and QOs, which are a collection of best practices.

The following is a breakdown of the paper's structure. The literature review is discussed in Section 2. The proposed EMM, as well as its implementation specifics, are introduced and elaborated on in Section 3. Section 4 discusses the proposed framework's validation procedure, including a case study and expert review and compliance with the CMMI framework. Section 5 concludes with a conclusion and recommendations for future work.

2 Literature Review

Researchers are not new to the field of building maturity models. Various researchers have suggested several maturity models assess the quality and reliability of software products and processes. In this section, we discuss those relevant contributions.

The author of [6], contributed with a maturity model to gauge the maturity of the DW. This model consisted of six maturity levels i.e., Prenatal, Infant, Child, Teenager, Adult, and Sage. Each level denotes a business value of DW and consists of defined characteristics to achieve the business value. Achieving the highest level enables an organization to get a high business value. Nevertheless, the model focuses on the enterprise level DW only not specifically on the ETL. All the phases have been addressed but the same characteristics have been applied to each phase of the DW. The metrics for DW conceptual model understandability have been discussed by the author in reference [7]. According to the author, these metrics shall help understand the conceptual model and have empirically proved its validation. The author has not discussed metrics about the ETL process in his contribution. The author of reference [8] has compared various BI maturity models and has concluded that different maturity models are there but no maturity model focuses on the maturity of the organization itself. In our research, we have addressed the problem highlighted by the author. Likewise, the author of reference [9] has proposed a maturity model for BI called EBIM with five levels of maturity and four key dimensions. But this model has not given enough attention to the process that integrates data in the DW. The concept has been discussed a bit in the information quality and master data management factors but only with a little attention. Similarly, the author of [10] suggested the five stages of maturity, in the data management factor of the technical aspect, though the ETL was not discussed. But again not enough details could be found related to data integration. No special attention is being given to

the ETL process. Kimball et al. [5] a leading visionary in the field of DW points out that a properly designed ETL system that extracts the data from source systems enforces data quality and consistency standards. The author further says that a principal goal of the ETL system is to deliver data in the most effective way to end-user tools as users look at the DW as a source of trusted information to build upon the management metrics, strategies, and policies. The author in [11], has proposed DWCMMS consisting of 60 questions related to each process of DW. ETL has been assigned 7 questions to assess the maturity of this process. The model is purely based on the questions being verified by various DW experts and does not specifically focus on the ETL process but the entire DW system.

The author of the paper [4], has contributed to assessing the maturity of the BI model in an organization. Four dimensions have been considered i.e., organizational, process, technology, and outcome. The ETL aspect has been addressed in the technology dimension. Merely, the tool-based approach has been focused on the ETL process and showed the effectiveness of the ETL process only. This is again not enough to emphasize the very important process of the DW system. In another study [12] the author described how a domain-specific BIA maturity model for HEIs was established, the methodological design process was followed, the model itself, and lastly, demonstrations and validation of its coverage, accuracy, and usability by practitioners in the field. In reference [2], the author has introduced a very relevant DWP-M (Data Warehouse Process Maturity) model. It is an evolutionary contribution and its initial versions have been explained in [13]. This model consists of process areas and practices. The model is very comprehensive that covers the entire DW development process. The validation by 20 different types of experts has given strength to the author's contribution. But we could simply say that our proposed Maturity Model is for the ETL process particularly and this claim makes our research work more special. In reference [14], the authors have done a similar type of work but focus on the Implementation aspects of the BI model. The authors have proposed an enterprise-based BI maturity model to assess the implementation of a BI system in an organization. The model was prepared with the help of industry experts using the Delphi method. Another model named Information Quality Management Maturity (IQMM) model [15] was presented based on the CMMI framework. This model addresses information quality improvement. The model has five levels of maturity consisting of various activities, sub-activities, and input/output examples based on the IQ management process. In reference [16], a model-driven development framework for the ETL process has been proposed by the author. The model addressed the issues of technology-based ETL code generation. The framework drives the models into an independent development code where it could be implemented in the relevant technology as and when required. In reference [17], a maturity model for software maintenance was presented. The five levels of maturity are based on a CMM that includes KPAs that are specifically focused on software life cycle maintenance. Kimball et al. [18] provides more detailed guidance on the foundation for building a DW/BI system as well as ETL implementation. The author has extensive experience in the DW/BI field and has authored several well-known publications on the subject.

The EMM complies with the CMMI paradigm that is the key constituents of Maturity Levels and Process Areas (PAs). This feature adds to the concerns about quality and dependability. A CMMI compatible Requirements Change Management (RCM) Model for the CMMI Level II has been proposed by the author in reference [19] to assist software development firms in implementing the best practices of REQM. Literature research was conducted in two domains, Software Process Improvement (SPI) and Requirements Engineering (RE), to discover elements that aid in the execution of the requirements change management process. In addition, two companies were interviewed regarding the RCM process. This model was reviewed by an expert panel and is divided into five stages: Request, Validate, Implement, Verify, and Update. Scrum paradigm can be mapped to the

CMMI Level 2 process domains including Project Planning, Requirements Management, and Project Monitoring and Control [20].

3 Proposed Work

Capability Maturity Model Integration (CMMI) is an internationally known and widely used framework for process maturity. We built our proposed model on the basic paradigm of the CMMI framework, which brings more reliability and acceptance. In this regard, there are four building blocks of our proposed model called model constructs. The first two are called “Maturity” and “Maturity Level” constructs. The maturity of the ETL system shall be represented using the “Maturity Level” construct of the model. The third construct is the “Key Process Area (KPA)” which is a collection of linked activities that must be carried out as part of an ETL system to improve the relevant process area. The “Activity” construct is the fourth building block of the proposed model. Finally, the model shall integrate all of these elements so that it can deliver a full solution as a benchmark for ETL system maturity.

3.1 Model Development Roadmap

We chart out a model development roadmap to cope with this challenge. The roadmap is given as:

As shown in Fig. 2, preparing the initial version of the model was the first step. After preparing an initial draft of the model it was presented to the panel of experts. The first certified version of the model was developed after the detailed deliberation sessions. To strongly affirm the proposed model, different types of case studies were conducted and the model was revised in the light of these case studies. The updated and enhanced version of the model was finally issued after the revised model was presented to the experts once again for review.

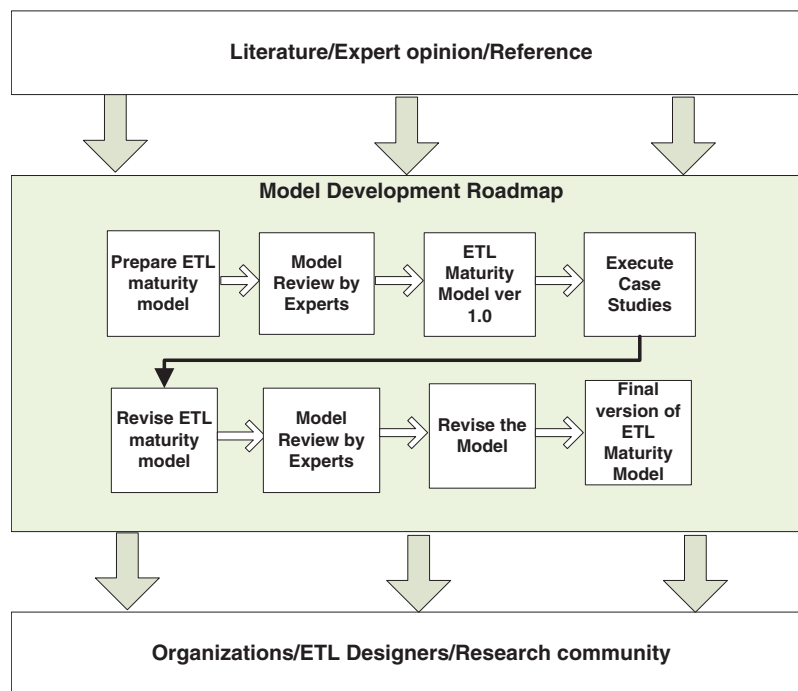


Figure 2: Model development roadmap

3.2 The ETL Maturity Model (EMM)

The EMM is the ultimate solution for producing high-quality knowledge by implementing quality processes during the ETL of data from multiple sources into a single target system. In terms of overall structure and paradigm, as previously stated, our suggested model is compliant with the CMMI framework. The proposed model’s overall closeness to the CMMI framework makes it more dependable and quality-oriented.

As shown in Fig. 3, there are five levels of maturity and each level contains a different no of KPAs except the “Chaotic” level. The higher the level, the more mature the ETL system is, with the “Chaotic” level being the least mature and the “Reliable” level being highly mature. Each level contains a specific number of KPAs. We placed these KPAs in their most relevant level of maturity. Each KPA consists of the QOs.

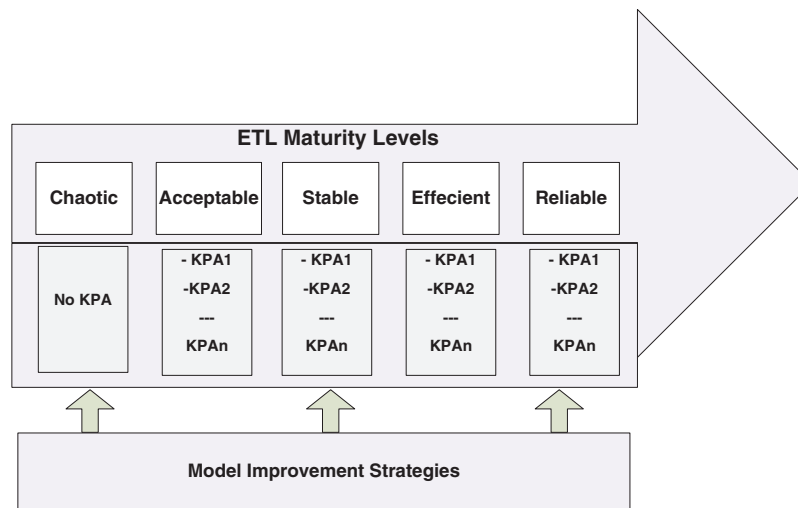


Figure 3: Overview of EMM

We devised the following initial version of the model, thoroughly evaluated by the panel of experts using multiple brainstorming sessions according to the design science research guidelines [21]. The model’s initial version is shown in Tab. 1.

Table 1: The initial version of EMM

Level no	Level	KPAs	QOs
5	Reliable	5.1	Causal analysis and resolution QO1: Analyze performance results QO2: Fine-tune the process
		5.2	ETL performance management QO1: Manage and enhance the performance results

(Continued)

Table 1: Continued

Level no	Level	KPAs	QOs
4	Efficient	4.1	ETL performance review QO1: Apply performance analytics QO2: Test and measure the performance results
		4.2	Data quality management QO1: Analyze data quality QO2: Maintain data completeness
3	Stable	3.1	Data profiling QO1: Analyze data QO2: Document the business rules QO3: Establish and maintain data mappings
		3.2	Data staging QO1: Decide on stage or not-to-stage QO2: Design the staging area
		3.3	Source system analysis QO1: Manage requirements QO2: Establish and maintain the requirements traceability
		3.4	ETL planning QO1: Understand business objectives QO2: Analyze the source system QO3: Plan the ETL process
		3.5	Architectural decisions QO1: Decide on technology requirements QO2: Discuss solutions QO3: Analyze pre-requisites
2	Acceptable	2.1	Requirements management QO1: Elicit requirements and the commitment to requirements QO2: Maintain requirements
		2.2	Dimensional modeling QO1: Design the data model QO2: Establish and maintain the dimensional model
		2.3	Data extraction QO1: Analyze data quality QO2: Establish and maintain the extraction method
		2.4	Data cleansing and loading QO1: Format data QO2: Apply cleansing methods
1	Chaotic	ETL is unpredictable and poorly executed. ETL management is as good as the ETL manager	

In the five levels of the model, there are 13 KPAs. The panel of specialists discussed and reviewed each KPA in great detail. This was the first iteration of our research's output work product. This first version of the model was selected to be enhanced by the execution of various types of case studies to affirm our work. After executing the case studies (discussed later) we came up with a revised and more

refined version of the model. This revised version had more process areas than the previous one as shown in [Tab. 2](#).

Table 2: Revised version of EMM

Level no	Level	KPAs	QOs
5	Reliable	5.1	Causal analysis and resolution QO1: Analyze performance results QO2: Fine-tune the process
		5.2	ETL performance management QO1: Manage and enhance the performance results
4	Efficient	4.1	ETL performance review QO1: Apply performance analytics QO2: Test and measure the performance results
		4.2	Data quality management QO1: Analyze data quality QO2: Maintain data completeness
3	Stable	3.1	Data profiling QO1: Analyze data QO2: Document the business rules QO3: Establish and maintain data mappings
		3.2	Dimensional model review QO1: Test drive the model QO2: Maintain and improve the model
		3.3	Data staging QO1: Decide on stage or not-to-stage QO2: Design the staging area
		3.4	Data flow control QO1: Plan and apply the data flow controls
		3.5	Source to target mapping QO1: Map the source and the target
		3.6	Source system analysis QO1: Manage requirements QO2: Establish and maintain the requirements traceability
2	Acceptable	2.1	ETL planning QO1: Understand business objectives QO2: Analyze the source system QO3: Plan the ETL process
		2.2	Architectural decisions QO1: Decide on technology requirements QO2: Design the CDC/complete refresh strategy QO3: Design staging selections

(Continued)

Table 2: Continued

Level no	Level	KPAs	QOs
		2.3	Requirements management QO1: Elicit requirements and commitment to requirements QO2: Maintain requirements
		2.4	Dimensional modeling QO1: Design the data model QO2: Establish and maintain the dimensional model
		2.5	Data extraction QO1: Analyze data quality QO2: Establish and maintain the extraction method
		2.6	Data cleansing and loading QO1: Format data QO2: Apply cleansing methods QO3: Apply loading script QO4: Validate data
1	Chaotic	ETL is unpredictable and poorly executed. ETL management is as good as the ETL manager	

In comparison to the earlier version, the refined version has 3 more KPAs along with other changes in the ordering and minor adjustments in the QOs. This version is more detailed and comprehensive to address all facets of a good ETL system.

4 Validation of EMM

EMM has been validated both quantitatively and qualitatively. The EMM comprised of established maturity levels and accompanying KPAs was thoroughly assessed by a team of specialists. Using a rating-based methodology, the data was processed and produced using the Delphi method. Multiple systems incorporating ETL operations were used to test the proposed framework. The conclusions acquired after finishing the case study have been thoroughly detailed in the latter part of this section.

4.1 Expert Judgment

The proposed CCMF was validated by a panel of experts. A total of 22 experts were engaged from 6 different relevant firms. The Delphi method was adopted and the “Likert scale” was used to rate the opinions.

Tab. 3 below lists various characteristics of the experts (To maintain anonymity, firm names are not revealed):

The first version of the model was prepared and presented to the industry experts (profile shown in **Tab. 3**) for review. Delphi method was adopted using five points Likert scale to evaluate the initial version of the model. The choices according to the Likert scale were *Strongly Disagree* (1.00–1.80), *Disagree* (1.81–2.60), *Neutral* (2.61–3.40), *Agree* (3.41–4.20), and *Strongly Agree* (4.21–5.00) with associated weights of 1, 2, 3, 4, and 5 respectively. The get more quality results the weighted averages below 3.40 were ignored. In the model review, 22 industry experts were involved.

Table 3: Profile of experts

Company	Specialty	Experience	No of experts
Unrevealed	DW	20 years	3
-	ETL designer	15 years	5
-	ETL developer	5 years	4
-	Information system designer	15 years	2
-	DW/BI architect	5 years	3
-	CMMI appraisal team member	15 years	5

All of the experts were requested to look at the model's KPAs and the associated QOs. The experts examined the KPAs and QOs from three perspectives: the rationale of their existence, appropriate placement in the model, and their intended purpose.

The average ratings for all three questions were found well above the desired outcome. The experts had to assess all of the KPAs/QOs and confirm their inclusion in the model in the first step. That means a KPA or QO had to pass a rigorous review to be included in the model. They had to validate the placement of KPAs and QOs in their respective maturity levels in the second aspect, which included selecting the most appropriate KPA and QO for each level. They were supposed to assess the KPA/QO in the third aspect. The third part required them to examine the KPA/QO in the context of their purpose and to give it a more appropriate title based on syntactical, semantic, and logical dimensions as shown in [Tab. 4](#).

Table 4: Results of expert reviews

	Perspective	Measure	'1'	'2'	'3'	'4'	'5'	Avg	Result						
Answers of respondents Answer to RQ-C	EPR-R1	Maturity levels represent the quality of a process and include the necessary components.	Occurrence	0	0	3	10	9	4.3	Strongly agree					
	EPR-R2	The proposed KPAs and QOs are reasonable and completely support their existence.	Percentage Occurrence	0	0	14	45	41	0	0	1	11	10	4.4	Strongly agree
	EPR-R3	All of the KPAs are appropriately placed in their respective maturity levels.	Percentage Occurrence	0	0	5	50	45	0	2	0	9	11	4.3	Strongly agree
			Percentage	0	9	0	41	50							

Tab. 4 displays the astonishing findings obtained from responders for each question. Each statement received a “Strongly Agreed” rating on the Likert scale. That implies a highly desirable outcome for validating the model’s performance.

4.2 Case Studies

In the company (name withheld), we conducted two case studies in two different domains. The first case study concerns the organization’s departments that supply various administrative services to industrial departments. The Human Resource (HR) department, which handles numerous Human Resource management duties, is the subject of the second case study.

4.2.1 Services Information System (SIS)

The data mart for the Services Information System (SIS) was processing data linked to user requests for various maintenance services such as heating/cooling systems, computer equipment, and electrical appliances, as well as maintenance records and the services items. The proposed maturity model was used to integrate maintenance data in an existing ETL process. All KPAs and QOs at level 2 were implemented (levels 3, 4, and 5 shall be discussed in a separate publication).

The performance comparison of traditional and model-driven ETL is displayed in Fig. 4. The model-driven ETL method improved data quality, decreased maintenance time, and increased customer satisfaction. Customer feedback from roughly 60 end-users was used to assess the overall data quality and customer satisfaction. Customers were given a survey once the new ETL system was installed. This questionnaire included questions about data quality, expected outcome change, and performance-related questions. The availability of traditional and model-driven ETL was calculated using the standard availability formula i.e., $MTBF / (MTBF + MTTR)$, Where MTBF is the Mean-Time-Between-Failure and MTTR is the Mean-Time-To-Repair. The outcome of the availability aspect is shown in Fig. 5.

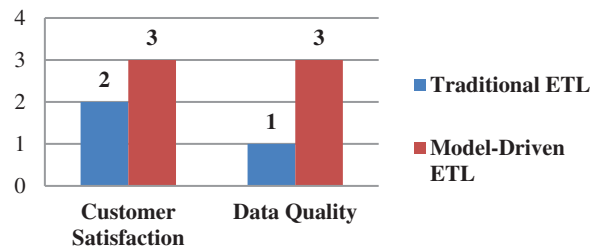


Figure 4: Maturity model performance analysis for SIS

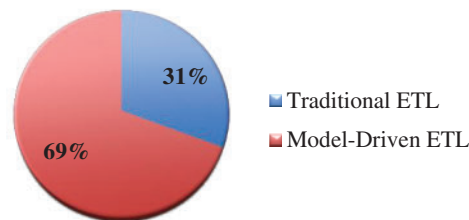


Figure 5: The availability of model-driven vs. traditional ETL for HCIS

The model-driven has outclassed the traditional ETL by a big margin in the context of the system’s availability as depicted in Fig. 5.

4.2.2 Human Capital Information System (HCIS)

The HR domain was chosen for validation of our proposed model in the second case study. Personnel profile, Staff Postings, Career Progression, Performance Evaluation, and Punishments/Rewards were some of the sub-functions of the HR Department. HR data mart was already up and operating to suit the users’ basic requirements. The ETL procedure was in place, and it met the DW’s data integration requirements. All the KPAs/QOs at level 2 were again implemented here. After transforming the existing ETL process into a quality ETL system using our EMM, considerable improvement was found in the ETL vis-à-vis in the DW already in place. We utilized the same yardstick to validate the HCIS as we did with the SIS.

As seen in Fig. 6, the ETL based on our suggested model outperforms the standard ETL procedure in every way. Users were urged to look at their earlier work and compare it to the current state. Users had high trust in the system since it had a complete data set, uniform data formats, and correct data values. In comparison to the previous version, the data values were found to be more consistent, complete, and correct. Due to its complexity, inappropriate execution, and disorganized set of operations, the previous ETL process was difficult to maintain, but the newly changed ETL was found to be more manageable and easy to maintain.

As shown in Fig. 7, model-driven ETL has almost double the system availability of traditional ETL. The implementation of KPAs and QOs enhanced system uptime substantially, confirming the model’s robustness. This is because the HCIS was not a single example, but rather five sub software modules, these two case studies helped us refine our suggested model. In other words, we validated our model on six different software systems, which is sufficient for any model. Following the above-mentioned practical experiences, we redesigned the model and made a few necessary improvements. Using the same procedure, the model was given to the panel of experts for assessment again, and we came up with the more refined version of the EMM displayed in Tab. 2.

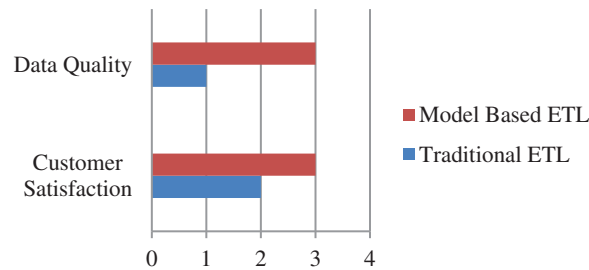


Figure 6: Maturity model performance analysis for HCIS

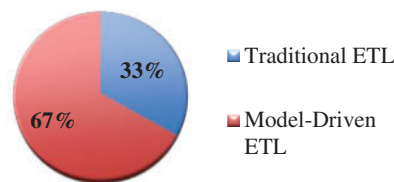


Figure 7: The “Availability” of model-driven vs. traditional ETL for HCIS

4.3 Compliance with the CMMI Framework

In terms of its overarching paradigm, our proposed EMM adheres to the CMMI architecture. The Maturity Levels, Process Areas, and Specific Goals make up the CMMI structural hierarchy. The following comparison illustrates how closely EMM adheres to the CMMI structure. A quick comparison of the two structures, EMM and CMMI, is shown in [Tab. 5](#).

Table 5: EMM's compliance with the CMMI paradigm

Structure components	Availability in the model		Result
	<i>CMMI</i>	<i>EMM</i>	
Maturity Level (ML)	ML	Level	✓
Process areas	KPAs	KPAs	✓
Specific goals	SGs	QOs	✓
	Features		
No of levels	5	5	✓
Ordering of maturity levels	5~Highest 1~Lowest	5~Highest 1~Lowest	✓
Components arrangement	ML, KPAs, SGs	Level, KPAs, QOs	✓

The comparison plainly shows that the two architectures are very similar. The essential components and features of both structures have been assessed. In both aspects, the results reveal that EMM is fully compliant with the CMMI paradigm and conforms to the structural as well as semantic features of a world-renowned framework.

5 Conclusion

We introduced EMM, a ground-breaking effort to turn normal ETL operations into a high-quality ETL solution. The model has sixteen (16) KPAs and has five stages of maturity. The applicable QOs are found in each KPA. The model was assessed by a panel of 22 highly skilled professionals throughout several brainstorming sessions. Besides that, two case studies (i.e., six software modules) were used to validate the EMM for maturity level 2. The validation results were promising in terms of three important quality indicators, customer satisfaction, data quality, and system availability. When compared to the previous ETL procedure, the results for all parameters for the new ETL system were remarkable. This approach will assist organizations in improving the performance of their DW systems by adopting a quality process framework for their ETL infrastructure, which will lead to better and more informed decision-making. Furthermore, ETL designers and developers can make use of important KPAs and QOs that are available at various levels of maturity to design and construct a more robust and successful ETL system. We can extend this EMM to all three phases of the ETL system in the future, resulting in a more fine-grained approach that could aid in quality improvement by getting insight into the ETL system.

Acknowledgement: The authors extend their appreciation to King Saud University for funding this work through Researchers Supporting Project Number (RSP-2021/387), King Saud University, Riyadh, Saudi Arabia.

Funding Statement: This work was supported by King Saud University for funding this work through Researchers Supporting Project Number (RSP-2021/387), King Saud University, Riyadh, Saudi Arabia.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] C. W. Khuen and M. Rehman, "A maturity model for implementation of enterprise business intelligence systems," in *Proc. Int. Conf. on Mobile and Wireless Technology*, Kuala Lumpur, Malaysia, pp. 445–454, 2018.
- [2] A. Sen, K. Ramamurthy and A. P. Sinha, "A model of data warehousing process maturity," *IEEE Transactions on Software Engineering*, vol. 38, no. 2, pp. 336–353, 2012.
- [3] A. N. Hidayanto, H. S. Indriany, A. Prastya and S. F. Mardiansyah, "Data warehouse capability maturity model assessment for efficient monitoring process: A case study in national narcotics board," in *Proc. IOP Conf. Series: Earth and Environmental Science*, NIT Raipur, India, vol. 969, no. 1, 2022.
- [4] I. L. Ong, P. H. Siew, and S. F. Wong, "Assessing organizational business intelligence maturity," in *Proc. Conf. on Information Technology and Multimedia*, Kuala Lumpur, Malaysia, pp. 1–6, 2011.
- [5] R. Kimball and J. Caserta, *The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning Conforming, and Delivering Data*, Hoboken, NJ, UAS: John Wiley & Sons, 2004.
- [6] W. Eckerson, "Gauge your data warehouse maturity," *Information Management*, 14, vol. 11, pp. 34, 2004.
- [7] M. Serrano, J. Trujillo, C. Calero and M. Piattini, "Metrics for data warehouse conceptual models understandability," *Information & Software Technology*, vol. 49, no. 8, pp. 851–870, 2007.
- [8] G. Lahrman, F. Marx, R. Winter and F. Wortmann, "Business intelligence maturity models: An overview," in *Proc. Int. Conf. on Information Technology & Multimedia*, Kajang, Malaysia, pp. 1–6, 2011.
- [9] C. Tan, Y. W. Sim and W. Yeoh, "A maturity model of enterprise business intelligence," in *Proc. Knowledge Management and Innovation: A Business Competitive Edge Perspective*, Cairo, Egypt, pp. 20–29, 2010.
- [10] B. Dinter, "The maturing of a business intelligence maturity model," in *Proc. Americas Conf. on Information Systems*, Seattle, Washington, USA, pp. 3948–3957, 2012.
- [11] M. Spruit and C. Sacu, "DWCMM: The data warehouse capability maturity model," *Journal of Universal Computer Science*, vol. 21, no. 11, pp. 1508–1534, 2015.
- [12] E. Cardoso and X. Su, "Designing a business intelligence and analytics maturity model for higher education: A design science approach," *Applied Sciences*, vol. 9, no. 12, pp. 4625, 2022.
- [13] A. Sen, A. P. Sinha and K. Ramamurthy, "Data warehousing process maturity: An exploratory study of factors influencing user perceptions," *IEEE Transactions on Engineering Management*, vol. 53, no. 3, pp. 440–455, 2006.
- [14] M. -H. Chuah and K. -L. Wong, "Construct an enterprise business intelligence maturity model (EBI2M) using an integration approach: A conceptual framework," in *Proc. Business Intelligence-Solution for Business Development*, London, UK, pp. 1–12, 2012.
- [15] I. Caballero, A. Caro, C. Calero and M. Piattini, "IQM3: Information quality management maturity model," *Journal of Universal Computer Science*, vol. 14, no. 22, pp. 3658–3685, 2008.
- [16] H. Carvalho, D. Battaglia, D. Montini, G. Moreira, L. Dias *et al.*, "ETL process model for a manufacture cells production line integration," in *Proc. Int. Conf. on Information Technology: New Generations*, Washington, DC, USA, pp. 1165–1171, 2010.
- [17] A. April, J. H. Hayes, A. Abran and R. Dumke, "Software maintenance maturity model (SMmm): The software maintenance process model," *Journal of Software Maintenance & Evolution: Research & Practice*, vol. 17, no. 3, pp. 197–223, 2005.
- [18] R. Kimball, M. Ross and A. A. Anisimov, "The data warehouse toolkit: The complete guide to dimensional modeling," in *John Wiley & Sons*, vol. 32, no. 3, pp. 101–102, 2003.

- [19] M. Niazi, C. Hickman, R. Ahmad and M. Ali Babar, "A model for requirements change management: Implementation of CMMI level 2 specific practice," in *Proc. Int. Conf. on Product Focused Software Process Improvement*, Berlin, Heidelberg, Springer, pp. 143–157, 2008.
- [20] J. Diaz, J. Garbajosa and J. A. Calvo-Manzano, "Mapping CMMI level 2 to scrum practices: An experience report," in *Proc. Communications in Computer & Information Science*, New York, USA, vol. 42, pp. 93–104, 2009.
- [21] A. Hevner and S. Chatterjee, "Design science research in information systems," in *Proc. Design Research in Information Systems*, Boston, MA, Springer, pp. 9–22, 2010.