International Journal of Information Technology and Language Studies (IJITLS)

A Proposed ERP Data Collector to Converge the Old Legacy Data into the New ERP

Sultan T Alanazi[1,2], **Hael Al-bashiri**[3,4], and **Yousef Sanjalawe**[2,5]
Sultanalenazi2011@gmail.com; Hailealbashiri1@gmail.com; Josseph_hfs@hotmail.com

[1] Imam Muhammad bin Saud Islamic University, Riyadh, Saudi Arabia
[2] National Advanced IPV6 Center, University Sains Malaysia, Penang, Malaysia
[3] Universiti Malaysia Pahang, Lebuhraya Tun Razak, Kuantan, Pahang, Malaysia
[4] Center for Science and Technology, Sana'a University, Yemen
[5] Computer Science Department, Northern Border University, Ar'ar, Saudi Arabia

Abstract. Nowadays, one of the biggest Information Technology investments is in Enterprise Resource Planning (ERP) structure. ERP provides for the flow across the whole of an organization's information by integrating and controlling the core business processes into a single application and single data repository. Moreover, ERP is considered time-consuming and expensive in terms of money and effort. ERP faces an obstacle in transforming data from the old legacy systems into new ERP systems since there is no defined standard model that can be used. Thus, in this paper, a new practical, simple, and effective solution to convert data from the old legacy system into the new ERP is proposed. This solution assumes that the data source and the target ERP are mismatched in data type formats, data values, data incompatibility, and missing data exist. The proposed solution defined a semi-auto data conversion process based on the meta-data of the source and target databases with the consultant of predefined mapping rules. First, the mapping process suggests a mapping list based on the predefined rules and the mapping confidence level (a threshold). Next, the final uploading process, human intervention is required to guarantee the correctness of the final loading process. The paper ends with a conclusion and future work.

Keywords: *Enterprise Resource Planning, Data Collector, Data Conversion, Data mapping.*

1. Introduction

Enterprise Resource Planning (ERP) is one of the biggest IT investments adopted by organizations to gain a competitive advantage, enhance their productivity and customer satisfaction, and thereby improve their position in the market (Borovskiy & Zeier, 2010; Tarhini, Ammar, Tarhini, & Masa'deh, 2015). ERP provides the information flow across the whole organization by controlling and integrating the core business processes into one single data repository and one single application (Mahmood, 2019; Viorel-Costin, 2019). Despite the advantages posed by ERP, some researchers, in this same field, agreed that implementing traditional ERP is time-consuming and expensive in terms of money and efforts (Mahraz, Benabbou, & Berrado, 2020). Therefore, many ERP vendors search to find suitable alternatives to attract more organizations in order to invest in their ERP systems (especially small to medium-size ERPs).

With the Cloud Computing (CC) evolution, cloud ERP has a big chance to be demonstrated and considered as the affordable choice for vendors and organizations (Al-Sharafi, Arshah, & Abu-Shanab, 2019; Al-Sharafi, Arshah, Abu-Shanab, & Alajmi, 2019). Without a defined standard model that can be used in data conversion from legacy systems into ERP systems, conversion data from the current running legacy databases to the new ERP database remains a main issue (Yadati, Manchala, Pandya, & Srikumar, 2012). Therefore, in this paper, to facilitate the conversion process a new practical, simple, and effective solution called "Data Collector" is proposed. Data Collector can ease data conversion from the legacy databases to ERP database in order to satisfy the needs of organization (Monteiro, 2019). The process will passed through three mean steps which are: developing a Data Collector procedure for feeding the new ERP database (Jenab, Staub, Moslehpour & Wu, 2019; Gill, Amin & Saleem, 2020); developing a web application to help the implementer during the Data Collection; Designing and Executing a new method to integrate the current legacy databases to the new ERP, smoothly, easily and precisely.

The rest of the paper is organized as follows. Section two presents the work of the existing scholars in ERP literature. Section three explains the proposed solution. Finally, the conclusion, implementation, and analysis are presented in section four.

2. Literature Review

In this section, some of the approaches that had been used in the previous implementation of ERP and its roots are reviewed. Earlier ERP software has its roots in erstwhile inventory management and control systems of 1960s (Yadati et al., 2012). At that time, mainframe systems were the common platforms for such systems with third generation programming languages (3GL) like COBOL, PASCAL, ALGOL and FORTRAN. The main functions were to monitor inventory of raw materials and to provide inventory control reports. The Material Requirement Planning (MRP) software followed afterwards, in the 1970s. The new MRP was used to control the Sales and marketing operations with the generation of production and operation planning schedules and inventory management (Nagpal, Khatri, & Kumar, 2015).

The implementation scope of MRP got enlarged in the 1980s when the MRP functions contain the manufacturing strategy and quality control. The MRP got renamed to MRP II. The development platform changed to the fourth-generation programming languages (4GL) like Python, Ruby and Perl in addition to the updated versions of the Object-Oriented 3GL based languages like C, C++, and Java. In the 1990s the concern of the ERP was on designing production supply chain management. The name of term ERP came by the Gartner Group. The wholesome software simply means the integration between main activities of an enterprise such as operations, supply chain, production, and distribution; along with the secondary minor activities planning, marketing, inventory, human resource management, accounting, and sales and marketing (Nagpal, Kumar, & Khatri, 2017; Yadati et al., 2012).

By the beginning of the 21st century, the focus of the ERP has changed dramatically to include not only the enterprise level processes but also the peripheral processes. This ERP revolution comes along the fifth Generation Applications (5GA) like Supply Chain Management (SCM), Customer Relationship Management (CRM), Business-to-Business (B2B), Business Intelligence (BI), Sales Force Automation (SFA), etc. The ERP focus becomes on the entire value chain of an enterprise to support real-time visibility anytime anywhere throughout the chain. During the 2000s-decade, market adapted new technologies like cloud computing to the ERP processes. Cloud ERP systems become the norm for most of the organizations and known as SaaS based ERP or Cloud ERP (Nagpal et al., 2017).

Moreover, there are no definite standard guidelines that provide a step-by-step ERP implementation approach. It is totally up to the project managers, type of the project, and the steering committee. Together they could agree upon some certain guidelines toward the steps that would be accomplished; some steps might be delayed, reordered and some performed sequentially and some in-parallel for the ERP implementation (Elbahri et al., 2019; Nagpal et al., 2015). After all, visioning and planning as well as process management are the key factors in ERP implementation. There are several ERP implementations approaches according to (Kraljic, Kraljic, Poels, & Devos, 2014).

The implementation of each approach consists of different implementation steps. With no agreed international standard to adopt ERP implementations, the most common practices have shown that each ERP project has it is own characteristics which make it distinct from another similar project even when the system components are the same. This is because each organization has it is own limitations and considerations, like the budget, timeframe, steering committee, data readiness, process readiness, laws, bylaws, policies, procedures in addition to many other factors.

3. The Proposed Solution

The illustration in figure 1, shows the proposed solution architecture. The focus is to map source data into the target ERP database, the mapping process use the ERP metadata table format in the mapping process. In this process, the table matching element is created, which are converted into the mapping rules that are defined in the ERP metadata with consultancy of the source metadata as well. The final output is the mapping table which is ready to be automatically uploaded into the ERP database, and the elements to be checked table. At the end, it requires more manual and semi manual manipulation in order to determine the correctness of the elements for uploading or rejection.

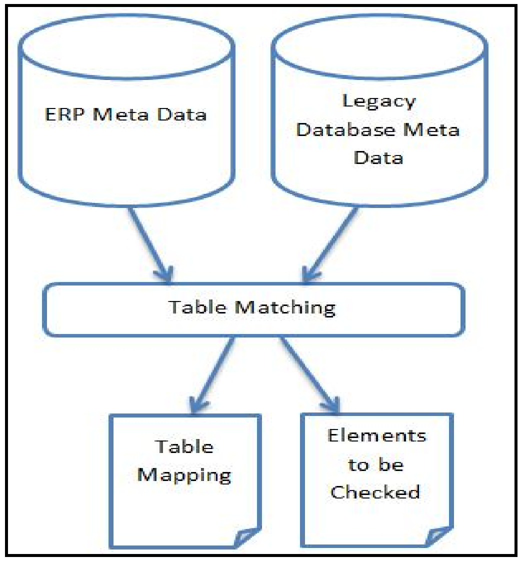

Figure 1. Proposed Solution Architecture

Taking a deeper look into the solution architecture, the detailed solution consists of three main components: The Data Source Layer; the Data Collector Layer; and the ERP layer. The API Collector receives requests from the application (i.e. HTML format), parses the request, converts it to the proper format, passes the request to the Data Collector Layer in order to extract the relevant information from the Data Source. after the Data Collector layer retrieves the desired information it refines the retrieved data according to some refinement procedures, and then it stores the refined data on the ERP database. In this work, we will use the same ETL architecture to manipulate the data mapping task.

The Data Source layer represents the legacy systems currently running in the organization, this layer might have different data formats from different sources, no matter what the data source was, this layer can adopt most of the current data connectivity adaptors being used in the market as long they comply with the international standards for data connectivity, like ODBC, JDBC or any other known adaptor, in the extreme cases a middleware application adaptor could be used.

The purpose of creating a Data Collector is to gather data from different data sources, manipulates it, processes it, and forwards it to the proper service reporting. The collected data could be saved on the ERP database.

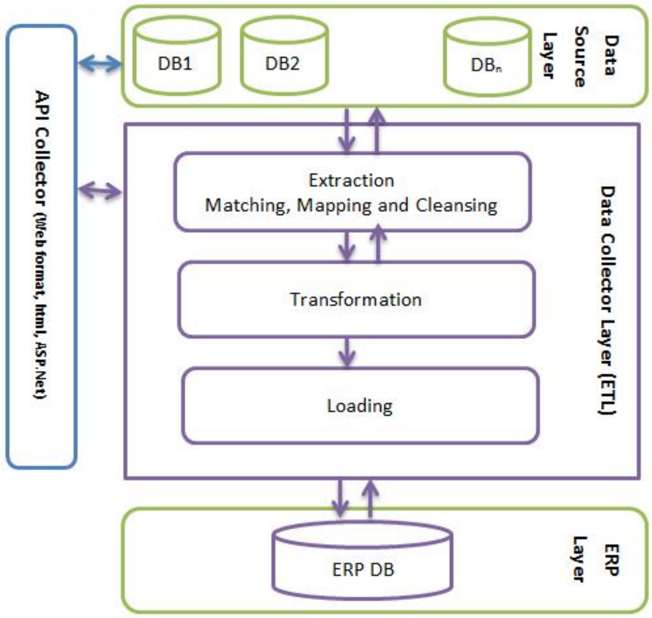

Figure 2. Abstracted Solution Architecture

The detailed solution architecture is shown in figure 3 below. The core solution is shown in the detailed solution architecture. The Extracting process is the heart of the system, where it reads the source and target meta-data and the matching rules, then it applies these rules to produce the temporary tables which are ready to be uploaded automatically or needs human intervention to upload, upon completion the updated ERP system is ready to be used in its freshly updated version.

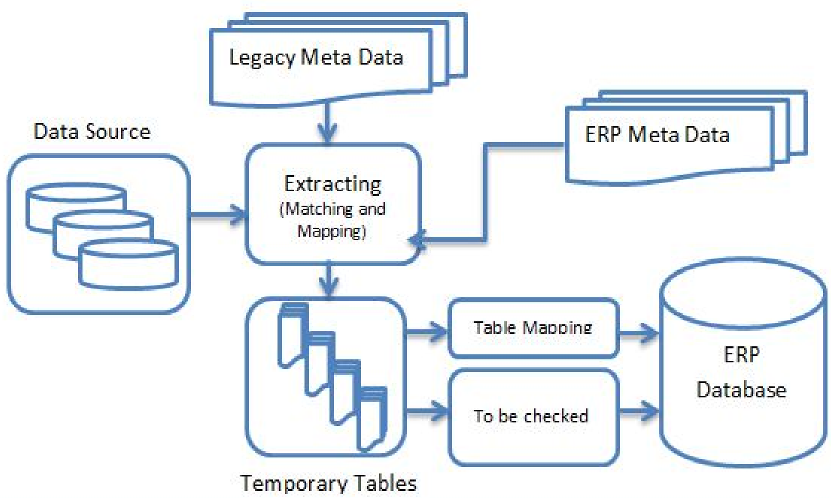

Figure 3. The Detailed Solution Architecture

Theses layers are logically connected by means of JDBC, ODBC, Source metadata and ERP metadata to complete the whole data transformation process. The data source represents the legacy systems currently running in the organization, and it is the set of existing databases and tables that represent the old running systems. Since both the SQL Server and the Access databases support JDBC, ODBC and Oracle database connectivity, the Collector Layer can communicate directly with the legacy systems after performing some system configurations, the system configurations allow the Collector Program to access the required legacy system when it is needed.

Moreover, the core component of the proposed solution is the Data Collector Layer; it consists of many subcomponents to perform data matching. These components are data mapping, data cleansing, extraction, data transformation and data loading and saving. Data collection and mapping from a legacy system(s) to ERP systems is very difficult task, since there are no formal methods to do the conversion process (Nagpal, Khatri & Kumar, 2015). To complete the data transformation into the ERP, we will adopt the ETL tool, as proposed by Kraljic *et al*. (2014). The following processes represent the major components of the ETL tool:

- Data Matching & Mapping.
- Data Extraction.
- Data Transformation.
- Data Loading and Saving into the ERP database.

The proposed solution follows the above mentioned steps in order to extract the required data and store it in the required ERP database; the proposed solution will try to find a formal representation model to map the incoming data from different sources to be in a suitable format for loading to the target ERP.

4. Implementation and Analysis

As mentioned before, the data collection and mapping from the legacy system(s) to ERP systems is very difficult task, since there are no formal methods to achieve the process (Nagpal, Khatri & Kumar, 2015). The ETL tool is adopted as proposed by Kraljic *et al*., (2014). Based on this assumption the proposed system is built. After system implementation, many issues came up. The first one is data transformation needs special attention and ambiguity in data should be resolved early before the transformation process begins. Special consideration and attention should be paid before building the metadata file, which is considered the heart of the system, the data transformation rules should be stated explicitly and precisely. Building a fully automated transformation system is considered to be very hard, since human intervention would be required to decide which data could be transferred with high confidence level.

In our system we adopted the 75% confidence level as the accepted threshold, this value was chosen randomly without scientific bases, also, the confidence criteria and attributes are selected randomly as well, and this is another problem, where it needs to be based on some well-defined attributes that represent the case, in the future the confidence attributes and components need more discussion; in our system, we depend on the data type, data value, and some other identifiers to count the confidence threshold value.

After data is being collected, then the extraction and filtering process is performed, we noted that, the filtering criteria and mapping rules are hard to be identified; we used the data type conversion (data casting) to convert data types, and some common rules to resolve the data inconsistency, incomplete data, and missing data and mismatch values.

The auto data transformation will succeed only if the matching rules and confidence level is 57% or more, the percentage of auto data transformation depends on the nature of the source data (accuracy and correctness) and the matching and mapping rules, in most cases is less than 15% of the data that will be automatically transformed.

In the semi auto transformation is when the confidence level is more than 50% and less than 75%, this amount of data represents more than 50% of the whole data, and the human intervention may take less effort or high effort. The effort paid depends on the nature and accuracy of the original source collected data. The rejected data items represent about 50% of the whole data source to be transformed. These numbers apply only in our case. In other systems, these numbers (auto transformation, semi transformation or rejected data items) depends on the nature and accuracy of the source data values and the mapping rules.

The data transformation process could be achieved in high accuracy if source data is clear and clean. It is possible to integrate the legacy data source system into the ERP system, and make it available for searching purposes for a specific period of time, in the future, the legacy system should be deleted after finding an appropriate way to keep the records for transfer to the ERP system. At the end, one system should be working, while the legacy system should be archived or deleted. Until the system is deleted, the legacy system should be available for searching and reporting purposes only.

5. Conclusion

To develop our solution method, we used our own data transformation method from the legacy system into the ERP system, this method is semi-automated process, where the mapping criteria are predefined before the transformation takes place, the predefined rules are added to the meta-data of the source and target ERP databases. These rules take into consideration the schemas characteristics of both databases, and it try to solve most of the expected mismatching problems before it happened. The mapping process could be easily implemented using simple programming language with scroll down menus to allow the user choice more convenient to manipulate source and target modules, and to add, remove tables, attributes of the source and target databases. The findings from this paper reveals several practical applications worthy of future directions. First, it would be valuable to further explore how legacy systems are migrated or integrated with new ERP systems in short time frame without any loss in data and how this impacts organization performance and save cost. On the other hand, the practical implication of this research is that findings of this paper are in conformity with theory of constraints-based methodology to success ERP implementation. As future work, we will plan to refine our proposed ERP data collector by trying to find additional factors for evaluating the results. Furthermore, we will try to compare our proposed ERP data collector with other ERP data collectors that had been proposed by other researchers and compare our approach with other approaches that had been used to show our strength points over other techniques.

References

Alhammadi, Z. H., & Shaalan, K. (2018). Lean Transformation in Information Technology: The Case of IT Services in Financial Firms. *International Journal of Information Technology and Language Studies*, 2(2).

Al-Sharafi, M. A., Arshah, R. A., & Abu-Shanab, E. A. (2019). Questionnaire Development Process to Measure the SMEs' Continuous Use behavior towards Cloud Computing Services. *Paper presented at the Proceedings of the 2019 8th International Conference on Software and Computer Applications, Penang*, Malaysia. https://doi.org/10.1145/3316615.3316723.

Al-Sharafi, M. A., Arshah, R. A., Abu-Shanab, E. A., & Alajmi, Q. (2019). The Effect of Sustained Use of Cloud-Based Business Services on Organizations' Performance: Evidence from SMEs in Malaysia. *Paper presented at the 2019 5th International Conference on Information Management (ICIM).*

Borovskiy, V., & Zeier, A. (2010). Data manipulation APIs for ERP systems. *Paper presented at the 2010 6th International Conference on Advanced Information Management and Service (IMS).*

Çınar, E., & Ozorhon, B. (2018). Enterprise resource planning implementation in construction: challenges and key enablers. Journal of Construction Engineering, 1(2), 75-84.

Elbahri, F. M., Al-Sanjary, O. I., Ali, M. A., Naif, Z. A., Ibrahim, O. A., & Mohammed, M. (2019). Difference Comparison of SAP, Oracle, and Microsoft Solutions Based on Cloud ERP Systems: A Review. *Paper presented at the 2019 IEEE 15th International Colloquium on Signal Processing & Its Applications (CSPA)*.

Gill, A. A., Amin, S., & Saleem, A. (2020). Investigation of Critical Factors for Successful ERP Implementation: An Exploratory Study. *Journal of Business and Social Review in Emerging Economies*, 6(2), 565-575.

Hamdy, M. E., & Shaalan, K. (2018). A hybrid framework for applying semantic integration technologies to improve data quality. *International Journal of Information Technology*, 2(1), 27-44.

Jenab, K., Staub, S., Moslehpour, S., & Wu, C. (2019). Company performance improvement by quality based intelligent-ERP. *Decision Science Letters*, 8(2), 151-162.

Kraljic, A., Kraljic, T., Poels, G., & Devos, J. (2014). ERP implementation methodologies and frameworks: a literature review. *Paper presented at the 8th European Conference on IS Management and Evaluation (ECIME).*

Mahmood, F., Khan, A. Z., & Bokhari, R. H. (2019). ERP issues and challenges: a research synthesis. *Kybernetes*, 10(2), 7-13.

Mahraz, M.-I., Benabbou, L., & Berrado, A. (2020). A Compilation and Analysis of Critical Success Factors for the ERP Implementation. *International Journal of Enterprise Information Systems (IJEIS)*, 16(2), 107-133.

Monteiro, J. (2019). *Factors that affect effectiveness in the use of enterprise resource planning systems* (Doctoral dissertation).

Nagpal, S., Khatri, S. K., & Kumar, A. (2015). Comparative study of ERP implementation strategies. *Paper presented at the 2015 Long Island Systems, Applications and Technology*.

Nagpal, S., Kumar, A., & Khatri, S. K. (2017). Modeling interrelationships between CSF in ERP implementations: total ISM and MICMAC approach. *International Journal of System Assurance Engineering and Management*, 8(4), 782-798.

Tarhini, A., Ammar, H., Tarhini, T., & Masa'deh, R. e. (2015). Analysis of the critical success factors for enterprise resource planning implementation from stakeholders' perspective: A systematic review. International Business Research, 8(4), 25-40.

Viorel-Costin, B. A. N. Ţ. A. (2019). The Difficulties of Transformation Sap Erp System from Actual Landscape to S/4hana Digitalization Environment. A Case Study. Annals Of'constantin Brancusi'university Of Targu-Jiu. Economy Series, (2).

Yadati, S., Manchala, V., Pandya, S. M., & Srikumar, A. (2012). An approach to enhance data management and reports in Multi-lingual ERP applications. *Paper presented at the 2012 IEEE International Conference on Computational Intelligence and Computing Research*.