Jurnal Teknologi

Full paper

Process Oriented Data Virtualization Design Model for Business Processes Evaluation(PODVDM) Research in Progress

Ayad Hameed Mousa[a*], Norshuhada Shiratuddin[b], Muhamad Shahbani Abu Bakar[c]

[a]*Computer Science, University of Karbala, Iraq*
[b]*School of Multimedia and Communication Technology, Universiti Utara Malaysia, Malaysia*
[c]*School of Computing, Universiti Utara Malaysia, Malaysia*

*Corresponding author: Ayad.H.mousa@student.uum.edu.my

Article history

Received : 15 August 2014
Received in revised form :
15 October 2014
Accepted :15 November 2014

Graphical abstract

Abstract

During process enactment in the business process management (BPM) lifecycle, information collected on execution plans are stored in the form of log files and database tables by using information systems (IS). In the past decade, a new approach based on the applications of Business Intelligence (BI) in business process management has emerged. The approach implements process-oriented data warehouse and mining techniques. However, the main issue is providing the right information at the right time to facilitate process evaluation that can be used for performance analysis and improve business process. Existing techniques have limitations, including huge data in database log files, performance of Process Warehouse (PW), which is highly dependent on specific design), complexity of PW design, lack of convergence between business processes and PW specifications, and the need for real data during process evaluation stage. Objects such as processes, storage, and data repositories can be virtualized to address these limitations. The main aim of this study is to propose a process-oriented data virtualization design model for process evaluation in BPM. The model will be validated through expert reviews and prototype development as well as through a case study. In this paper, we describe the research motivation, questions, approach, and methodology related to addressing the described limitations by designing a model for evaluation in business processes using the Data Virtualization technique.

Keywords: Business intelligence; business process management; process warehouse; process evaluation; data virtualization; data warehouse

■1.0 INTRODUCTION AND MOTIVATION

Most enterprises today perform an ordered set of activities using organizational resources to transform inputs into outputs, and achieve outlined goals and objectives. This set of activities is known as a Business Process (BP) and can generally be known as a concatenation of enactments in a business platform based on the purpose of creating worth for a business [5]. Weske (2012) states that a BP is usually composed of a combination of activities or phases that are inter-related and need to be organized in a specific manner to achieve business goals.

Business Process Management (BPM) is a systematic approach that includes concepts, approaches, and technologies to support the design, administration, implementation, execution, analysis, and evaluation of BP to optimize value creation [6]. BPM offers a systematic and consistent approach to performing a wide range of business activities, such as realizing, documenting, designing, simulating, enactment measuring, and analyzing business operations of an organization. For these reasons, BPM has emerged as a necessary instrument in assisting enterprises to achieve their business objectives and competitive advantage [7].

A BPM life cycle generally consists of six phases: goal specification organizational analysis, process design specification and modeling, process configuration (implementation), process enactment (execution), process monitoring (business activity monitoring), and process evaluation (performance analysis) [5, 6]. These phases are expounded in Figure 1.

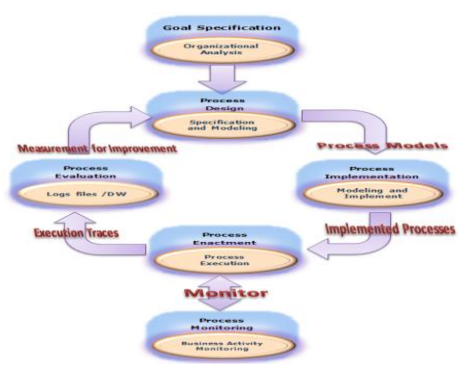

Figure 1 Business process management lifecycle

An important phase within BPM is the evaluation phase, which is also referred to in literature as Process Optimization or Process Performance Analysis and Improvement phase. This phase involves extracting data from the process performance evaluation phase to identify whether these data come from process modeling or process monitoring phase, determining bottlenecks to identify potential or existing bottlenecks and opportunities for saving costs, and making any other improvements [5, 8-10]. Subsequently, these information are used in the design processes. These enhancements would then support and create business value. Improvements to the business processes are conducted during this pahse [5, 6, 8-12].

Businesses rely heavily on data, and Data Virtualization (DV) presents a platform for business agility, action ability, information speed, and information quality. In a nutshell, using DV leads to reduced time for data integration and quicker response time to meet business information requirements. DV also provides opportunities to collect information and access new data sources that may have been added recently. DV can also deal with more complete data through a huge data size capability [13-16]. From a technical perspective, DV provides higher quality data that can be translated into business syntax and contexts, rather than delivery systems and data storage contexts. DV is also based on information delivery that is quick, efficient, and effective because data integration becomes easier in terms of information scope and timeliness [16-19].

1.1 RESEARCH QUESTIONS AND PROBLEM STATEMENT

Business processes must be planned and executed to attain specified organizational goals. During the enactment phase, information collected on execution plans are stored in the form of log files and database tables by using information systems (IS). [20, 21]. However, Grigory (2004) stated that execution logs cannot be used because of two main reasons. First, logs capture traces for a brief duration. Second, during the execution of processes, logs are continuously updated; thus, data from other sources cannot be added to the process logs because of limitations in the design. To address this issue, a data warehouse (DW) is built for BP analysis and is termed as process warehouse (PW). The performance of PW highly depends on its design [21-26]. Hence, the accuracy, speed, and live data for extraction and interpretation of information as well as providing them in a timely manner would have a positive impact on the analysis of the performance of any organization. PW could also provide support for decision-makers to make better decisions based on these analyses. Therefore, new methods, models, and guidelines must be formuated to facilitate timely and accurate analysis and definition of the basic steps in the process of analyzing performance. Without these factors, important phases in the performance analysis of BP may be overlooked [27]. However, the problem domain of this study is to provide the right information, which can be used for performance analysis and improvement of business processes, at a right time. The main concern is to facilitate the process evaluation phase of business processes by using DV technique.

A. PS1: Large size of process warehouse

Process warehouse is huge, and the magnitude of data needed for process performance analysis and decision making is too small when compared with the size of the PW and the approaches for optimal use would cause limitations to stakeholders who intend to use PW for business process improvement [25].

B. PS2: lack of convergence between business process and process warehouse

The lack of convergence between business process and PW specifications makes it difficult to determine which elements of a process should be analyzed to enhance process performance [28]. Today, business process optimization has become an integral interest in most enterprises and its key activity lies in the mining of process logs. However, log data in different information systems are often heterogeneous and thus, to ease process evaluation, must be processed together [29].

C. PS3: Data warehouse complexity

DW complexity leads to inflexibility of the data model design. Therefore, all BI tools must be able to deal with the large number of complexities of an organization's data, particularly when changing business requirements and business needs [30]. However, improving the efficiency of the extraction, transformation and loading (ETL) process can be very difficult. Designers must redesign the ETL process when changing business rules or altering data source. Consequently, rapid design and reuse of processes for majority ETL tools can be tricky [31].

D. PS4: Data warehouse refreshment

DW refreshment (integration of new data) has traditionally been done offline, which meant that OLAP applications could not access any data while updating the DW [32]. Conventional DW is unable to a) support any dynamics in its source structure and contents, b) support hidden-subjects, and c) provide data on-the fly i.e. real-time data and populate hidden-subjects on their evolution due to structure of its schema and contents.

Therefore, with the absence of sufficient and efficient virtualization design models and improved techniques that take advantage of the benefits of post-execution data for analysis and improvement of processes, the following research questions and objectives were formed. The main question that has guided this study is as follows:

How can the evaluation performance (analysis and improvement) of business processes be facilitated using process data virtualization?

To answer this question, we have divided it into the following sub-questions:

1. Existing approaches that enable process improvement in Business Process Management must be identified.
2. Data virtualization technique should be used to facilitate process-oriented data virtualization design model evaluation of processes and to decide on process improvements.
3. A business process evaluation prototype based on such design model must be developed.
4. The capabilities of the process-oriented data virtualization model for performance analysis and improvement in term of performance must be evaluated.

The overall goal of this thesis is to design a process data virtualization-based model that can facilitate and guide evaluation performance analysis and improvement of business processes. To achieve this objective we have to achieve the following sub objectives:

1. To identify existing approaches that enable process improvement in Business Process Management
2. To develop a process-oriented data virtualization design model evaluation of processes and decide on process improvements using data virtualization technique

3. To develop a business process evaluation prototype based on this model
4. To evaluate the capabilities of process oriented data virtualization model for performance analysis and improvement in term of performance

1.2 APPROACH AND METHODOLOGY

In our approach, we propose a Process Oriented Data Virtualization Design Model where data need to be extracted and consolidated from process logs and transactional sources into unified virtual form, in contrast with the traditional approach. The model aims to exploit the DV technique and facilitate the analytical capabilities of business process management systems by implementing DV to improve the business process. Figure 2 illustrates the proposed approach.

Our approach is based on integration of goals and business needs with Process Data Virtualization (PDV) to allow goal-user-based navigation of PDV.

We will propose the following:

1. Analysis of existing approaches in the evaluation of processes and improving the enabling process
2. A Process-Oriented Data Virtualization (PODV) method for BP evaluation as shown in Figure 2
3. A design model using PODV for process evaluation and performance analysis in business process
4. Evaluation Framework for Business Process Evaluation Approaches
5. Testing a prototype through a case study

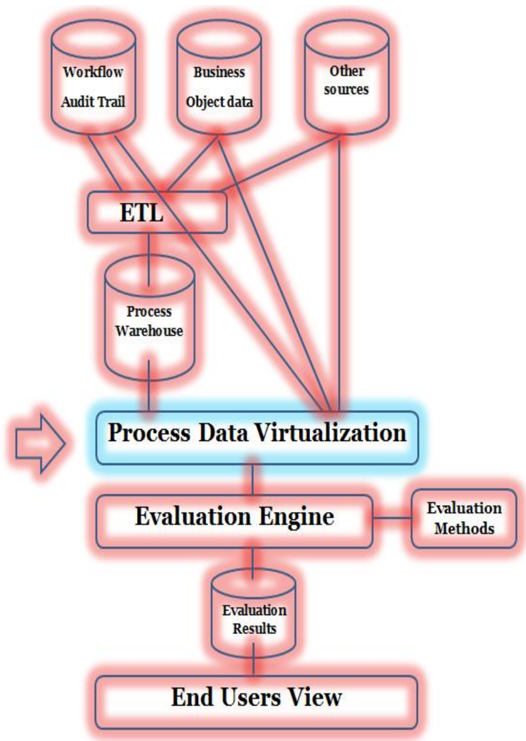

Figure 2 Proposed approach

According to Dwolatzky *et al*., 2002, systematic tasks should be planned to achieve research objectives [1]. This study is divided into three main phases as follows:

1. Theoretical study
2. Development
3. Empirical testing

In addition, five phases adapted from design science research methodology [2-4] associated with detailed activities were also utilized in achieving the objectives.

A. Theoretical Study

This phase consist of two sub phases, awareness of problem and suggestion. In the first sub phase, review of related literature and content analysis are followed by three types of comparative analysis, namely, existing BP evaluation models, existing BPM models, and existing approaches using DW for evaluation BP. The outcomes for this sub phase are the research problem and research gap. While in second sub phase, process data virtualization for evaluation BP is suggested as an early design, and discussed with experts to obtain valuable comments and achieve objective one. Figure 3 illustrates these tasks.

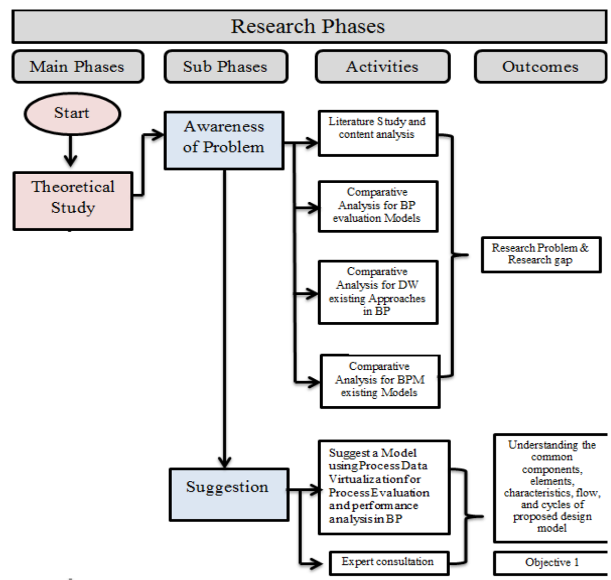

Figure 3 Theoretical study

B. Development

This phase include one sub phase (construction). During this phase, the proposed model is constructed and the second objective is accomplished. Figure 4 shows the development phase.

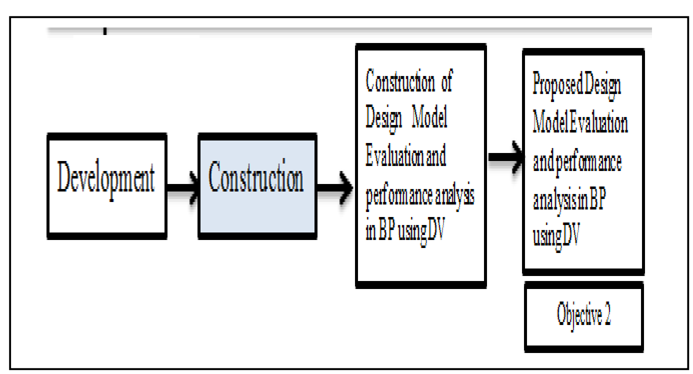

Figure 4 Development phase

C. Empirical Testing

This phase consists of two sub phases (evaluation and conclusion). During the evaluation phase, the evaluation strategies, including expert review, prototyping, and experimental methods, are developed. We then will analyze the results from the evaluation phase to write a report and publish the results. Figures 5 and 6 illustrate the empirical testing phase.

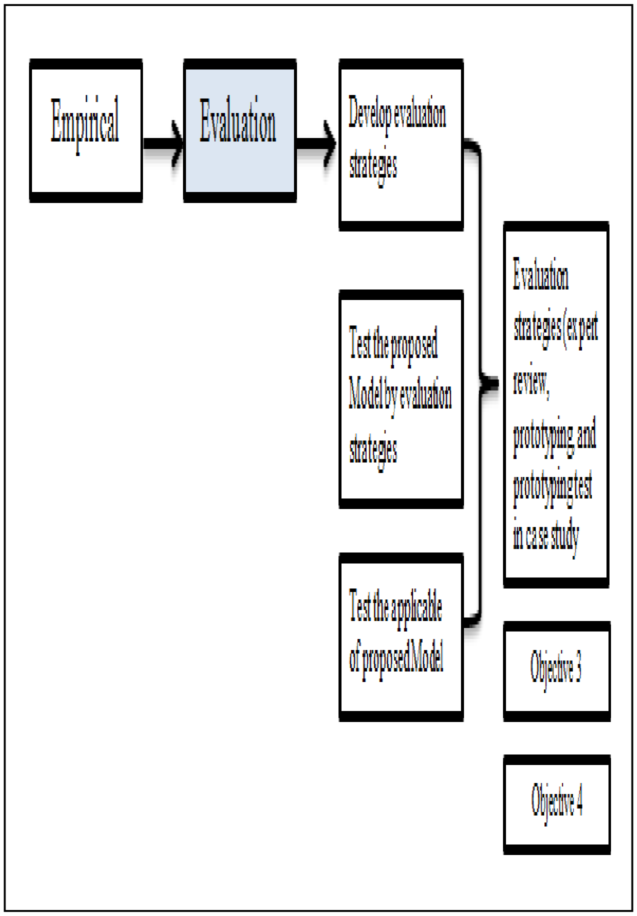

Figure 5 Evaluation phase

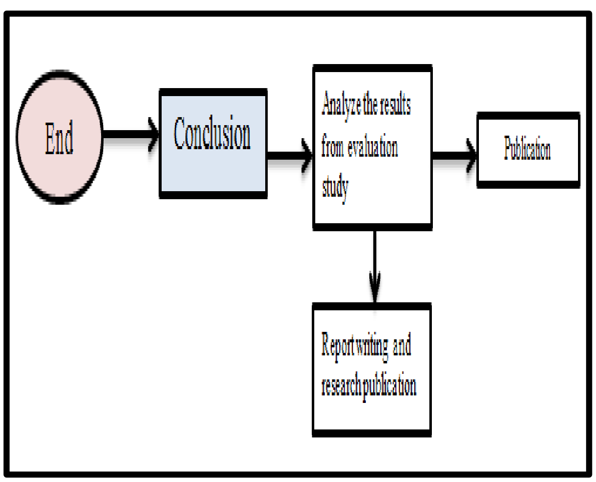

Figure 6 Conclusion phase

1.3 PROGRESS BEYOND STATE OF THE ART

Process evaluation (process improvement) is a type of "feedback" phase in the BP lifecycle. During this phase, executed processes are evaluated, and a diagnosis for bottlenecks is performed. Optimization of processes also occur during this phase place [9, 33, 34]. The Business Process Management Systems (BPMS) logs the process data and queries it using a built-in or third party reporting tool. The reports generated are used to analyze executions of operations, identifying bottlenecks and detecting deficiencies. However, the use of reports with current BPMSs presents performance, data quality, and semantic limitations [35, 36].

Business Process Cockpit has been proposed as a toolset for analyzing business process [36] and architectures such as Business Process Intelligence System architecture [33, 37] in terms of allowing stakeholders to observe and analyze processes-related data. Several studies on process warehouses have also concentrated on proposing adequate multidimensional data models for various types of warehouses, including generic [35]. In the last decade, many efforts on the business orientation in the data warehouse have been reported such as in [24, 38, 39].

In this study, we utilized three types of comparative analysis. The first type specializes in comparing models with respect to the previous evaluation of the business to determine all obstacles and past trends of the evaluation process in BPM. The comparative study included models proposed by [40-43]. The second type of comparative analysis is restricted to previous studies that have used DW as a storage for data created during the execution process in the life cycle of BPM. This comparison included 15 studies and approaches for the period from 2001 to 2012 proposed by [33, 35, 44-56] to determine all obstacles and problems that occur during the process, storage, and delivery of data in the required level. The third type of comparative analysis was conducted on five previous models for BPM [57-61] with the purpose of learning all the required components and elements that need to be saved and to understand the process of designing a model of BPM. The comparative analyses determine whether the reviewed models have certain guidelines to ensure it matches their target user. However, a lack of capacity to provide live and right data for use in evaluating BP was discovered. The scope of the previous studies is limited to data-orientated information system and a comprehensive model that can help to guide to process analysis and improvement has not been discussed. A model that uses data virtualization in BP evaluation is also lacking.

1.4 CONCLUSION

The proposed study is expected to contribute significantly by providing a comprehensive model to evaluate business processes. This model will be built depending on what has been proposed in previous studies and addresses the limitations, defects, and disadvantages of previous studies in this area. This model is applied to DV technique to overcome the problems arising during the use of DW. The problems mentioned above prevent the optimization and successful improvement of the process and should be addressed considering that most business processes are based on BI and DWs are the backbone of BI. Addressing these issues motivates us to apply such technology to limit and reduce such defects. In future work we hope to continue to apply the DV technology in most phases of the business process management lifecycle.

References

[1] B. Dwolatzky, I. Kennedy, and J. Owens. 2002. Modern Software Engineering Methods for Developing Courseware.

[2] F. Gacenga, A. Cater-Steel, M. Toleman *et al.* 2012. A Proposal and Evaluation of a Design Method in Design Science Research. *Electronic Journal of Business Research Methods*. 10(2).

[3] A. R. Hevner, S. T. March, J. Park *et al.* 2004. Design Science in Information Systems Research. *MIS Quarterly*. 28(1): 75–105.

[4] P. Offermann, O. Levina, M. Schönherr *et al.* Outline of a Design Science Research Process. 7.

[5] C. Houy, P. Fettke, and P. Loos. 2010. Empirical Research in Business Process Management–Analysis of an Emerging Field of Research. *Business Process Management Journal*. 16(4): 619–661.

[6] M. Weske, 2012. *Business Process Management: Concepts, Languages, Architectures*. Springer.

[7] G. Aagesen, and J. Krogstie. 2010. Analysis and Design of Business Processes using BPMN. *Handbook on Business Process Management 1*. Springer. 213–235:

[8] S. Adesola, and T. Baines. 2005."Developing and Evaluating a Methodology for Business Process Improvement. *Business Process Management Journal*. 11(1): 37–46.

[9] T. Bucher, A. Gericke, and S. Sigg. 2009. Process-centric Business Intelligence. *Business Process Management Journal*. 15(3): 408–429.

[10] A. Pourshahid, D. Amyot, L. Peyton *et al.* 2009. Business Process Management with the User Requirements Notation. *Electronic Commerce Research*. 9(4): 269–316.

[11] R. G. Lee, and B. G. Dale. 1998. Business Process Management: A Review and Evaluation. *Business Process Management Journal*. 4(3): 214–225.

[12] W. M. Van Der Aalst, A. H. Ter Hofstede, and M. Weske. 2003. *Business Process Management: A Survey*. Springer,

[13] M. Ferguson. 2011. *Succeeding with Data Virtualization High Value Use Cases for Analytical Data Services*.

[14] B. Hopkins. 2011. Data Virtualization Reaches the Critical Mass.

[15] R. F. v. d. Lans. 2012. *Data Virtualization for Business Intelligence Systems: Revolutionizing Data Integration for Data Warehouses*.

[16] A. H. Mousa, N. Shiratuddin, and M. S. A. Bakar. 2014. Virtual Data Mart for Measuring Organizational Achievement Using Data Virtualization Technique (KPIVDM). *Jurnal Teknologi*. 68(3).

[17] R. Eve, and J. R. Davis. 2011. Data Virtualization: Going Beyond Traditional Data Integration to Achieve Business Agility. *Composite Software*.

[18] R. Van der Lans. 2012. *Data Virtualization for Business Intelligence Systems: Revolutionizing Data Integration for Data Warehouses*. Morgan Kaufmann.

[19] L. Weng, G. Agrawal, U. Catalyurek *et al.* An Approach for Automatic Data Virtualization. 24–33.

[20] A. Lodhi, V. Küppen, and G. Saake. 2011. An Extension of Bpmn Meta-model for Evaluation of Business Processes. *Scientific Journal of Riga Technical University*. *Computer Sciences*. 43(1): 27–34.

[21] M. Zur Mühlen, and R. Shapiro. 2010. Business Process Analytics. *Handbook on Business Process Management*. 2: 137–157: Springer,

[22] B. List, J. Schiefer, A. M. Tjoa *et al.* 2000. Multidimensional Business Process Analysis with the Process Warehouse.

[23] B. List, and B. Korherr. An Evaluation of Conceptual Business Process Modelling Languages. 1532–1539.

[24] V. Stefanov, and B. List. Explaining Data Warehouse Data to Business Users-A Model-Based Approach to Business Metadata.

[25] Shahzad, and C. Giannoulis. Towards a Goal-Driven Approach for Business Process Improvement Using Process-Oriented Data Warehouse. 111–122.

[26] Shahzad. *A Data Warehouse Model for Integrating Fuzzy Concepts in Meta Table Structures.*

[27] G. Zellner. 2011. A Structure Evaluation of Business Process Improvement Approaches. *Business Process Management Journal*. 17(2): 203–237.

[28] A. Sturm. 2012. Supporting Business Process Analysis Via Data Warehousing. *Journal of Software: Evolution and Process*. 24(3): 303–319.

[29] H. Xia, Q. Yao, and F. Gao. Research and Design of Process Data Warehouse for Business Process Assessment. 377–385.

[30] C. Jossen, L. Blunschi, M. Mori *et al.* The Credit Suisse Meta-data Warehouse. 1382–1393.

[31] J. L. Liutong Xu, Ruixue Zhao, Bin Wu. 2011. A Paas Based Metadata-Driven Etl Framework. *IEEE*. 5,

[32] R. J. Santos, J. Bernardino, and M. Vieira. 24/7 Real-Time Data Warehousing: A Tool for Continuous Actionable Knowledge. 279–288.

[33] D. Grigori, F. Casati, M. Castellanos *et al.* 2004. Business Process Intelligence. *Computers in Industry*. 53(3): 321–343.

[34] M. Castellanos, A. Simitsis, K. Wilkinson *et al.* Automating the Loading of Business Process Data Warehouses. 612–623.

[35] F. Casati, M. Castellanos, U. Dayal *et al.* A Generic Solution for Warehousing Business Process Data. 1128–1137.

[36] M. Sayal, F. Casati, U. Dayal *et al.* Business Process Cockpit. 880–883.

[37] W. Tan, W. Shen, L. Xu *et al.* 2008. A Business Process Intelligence System for Enterprise Process Performance Management. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*. 38(6): 745–756.

[38] N. L. Sarda. 2001. Structuring Business Metadata in Data Warehouse Systems for Effective Business Support. *arXiv preprint cs/0110020*.

[39] V. Stefanov, and B. List. Business Metadata for the DataWarehouse. 20–20.

[40] C. McGregor, and S. Kumaran. Business Process Monitoring Using Web Services in B2B e-Commerce. 87.

[41] C. McGregor, and J. Scheifer. A Framework for Analyzing and Measuring Business Performance with Web Services. 405–412.

[42] A. Pourshahid, D. Amyot, L. Peyton *et al.* Toward an Integrated User Requirements Notation Framework and Tool for Business Process Management. 3–15.

[43] E. R. Aguilar, F. Ruiz, F. García *et al.* Evaluation Measures for Business Process Models. 1567–1568.

[44] M. K. Shahzad. 2012. Improving Business Processes using Process Oriented Data Warehouse. Dissertation, Computer and Systems Sciences, KTH-Royal Institute of Technology, School of Information and Communication Technologies.

[45] A. Bonifati, F. Casati, U. Dayal *et al.* Warehousing workflow data: Challenges and opportunities. 649–652.

[46] M. Zur Muehlen. Process-driven Management Information Systems Combining Data Warehouses and Workflow Technology. 550–566.

[47] C. W. Günther, and W. M. van der Aalst. A Generic Import Framework for Process Event Logs. 81–92.

[48] D. Grigori, F. Casati, U. Dayal *et al.* 2001. Improving Business Process Quality Through Exception Understanding, Prediction, and Prevention.

[49] J. Eder, G. E. Olivotto, and W. Gruber. 2002. A Data Warehouse for Workflow Logs. *Engineering and Deployment of Cooperative Information Systems*. 1–15. Springer,.

[50] P. Kueng, T. Wettstein, and B. List. 2001. *A Holistic Process Performance Analysis Through a Performance Data Warehouse*. Citeseer.

[51] J. Schiefer, B. List, and R. M. Bruckner. Process Data Store: A Real-time Data Store for Monitoring Business Processes. 760–770.

[52] P. Giorgini, S. Rizzi, and M. Garzetti. Goal-oriented Requirement Analysis for Data Warehouse Design. 47–56.

[53] K.-C. Pau, Y.-W. Si, and M. Dumas. Data Warehouse Model for Audit Trial Analysis in Workflows.

[54] L. Niedrite, D. Solodovnikova, M. Treimanis *et al.* Goal-Driven Design of a Data Warehouse Based Business Process Analysis System. 243–249.

[55] T. Neumuth, S. Mansmann, M. H. Scholl *et al.* Data Warehousing Technology for Surgical Workflow Analysis. 230–235.

[56] S. Mansmann, T. Neumuth, and M. H. Scholl. 2007. Multidimensional Data Modeling for Business Process Analysis.*Conceptual Modeling-ER 2007*. 23–38: Springer,

[57] K.-H. Kim, J.-H. Lee, and C.-M. Kim. 2005. A Real-time Cooperative Swim-lane Business Process Modeler. *Computational Science and Its Applications–ICCSA 2005*. 176–185: Springer.

[58] J. Ghattas, and P. Soffer. 2009. Evaluation of Inter-organizational Business Process Solutions: A Conceptual Model-based Approach. *Information Systems Frontiers*. 11(3): 273–291.

[59] M. D. Moeinaddin, Hassan Dehghan; Motahari, Saied. 2012 A Comprehensive Model for Performance Evaluation of Manufacturing Firms by Integrating Balanced Score Card and Fuzzy Analytic Network Process (A Case Study: Tile and Ceramic Manufacturers of Yazd Province). *Interdisciplinary Journal of Contemporary Research in Business*. 4(4): 18.

[60] V. C. Yen. 2009. An Integrated Model for Business Process Measurement. *Business Process Management Journal*. 15(6): 865–875.

[61] B. Kang, J.-Y. Jung, N. W. Cho *et al.* 2011. Real-time Business Process Monitoring Using Formal Concept Analysis. *Industrial Management & Data Systems*. 111(5): 652–674.