Requirement Engineering for the Development of Agent Oriented Data Warehouse

S. Bharadwaj[1*], A. K. Goyal[2]

[1]College of Computing Sciences and Information Technology, TMU, Moradabad, India
[2]Guru KulKangriVishwavidhalaya, Haridwar, India

*Abstract-*A software system design was said to be succeeded only when it satisfies the purpose for which it was intended. Requirement Engineering (RE) is the progression of determining that purpose, by recognizing stakeholders and their desires and recording these in a form that is adaptable to analysis, communication and implementation. Agent–oriented perceptions are growing very popular in software engineering as demonstrating frameworks for RE. On the other hand, Data Warehouse (DW) is a field of computer science that is used to capture the historical information to provide decisions to be taken by the management. In this paper, we used a blend of these technologies with Agent-Goal-Decision-Information (AGDI) model for the RE of DW. Based on the proposed AGDI model, a new RE approach has been proposed.

Keywords—Requirement Engineering, software engineering, Data Warehouse, AGDI

I. INTRODUCTION

Data Warehouse (DW) is intended to provide historical information for decision making to management. In the development of DW, RE phase is of much importance. Most of the RE approaches for DW does not differentiate the primary and secondary phase, contrasting current methods for the traditional information system. Also in these approaches information requirements are gathered, not the decision requirements, though the DW is for decision support. Decision requirements must capture information to provide decisions to be taken by the management. This motivated us to use an Agent-Goal-Decision-Information (AGDI) model for the RE of DW. Based on the proposed AGDI model, a new RE approach has been proposed in this paper.

Rest of the paper is organized as follows, Section I contains the introduction of Requirement Engineering and data warehouse, Section II contains the work of AGDI model, Section III contains the proposed Data Warehouse SDLC, Section IV explains the methodology of proposed RE approach for Data Warehouse design , section V concludes research work with future directions.

II. AGDI MODEL

The GDI model starts off evolved with the dedication of the dreams of the employer with the support of choice makers primarily based on the postulation that handiest the choice makers are the agents liable for selection making sports within the organization [8, 11]. Nevertheless, the other agents and their cravings for accomplishing the

desires of the enterprise also are crucial to be modelled. Since of this postulation, the GDI version [8, 11] absences the capability to model agents and dependencies among them for achieving desires of the enterprise. To address this problem, this paper suggested announcing the perception of agent to the traditional GDI version [8, 11] to signify the agents of the business enterprise. Within the paper, the prolonged GDI model is known as an AGDI model, as shown inside the scattered line frame is determined 2.1. In AGDI model, agents are categorized to version various agents within the employer. Further, agents' dependencies had been modelled over agent enslavements in the suggested AGDI version.
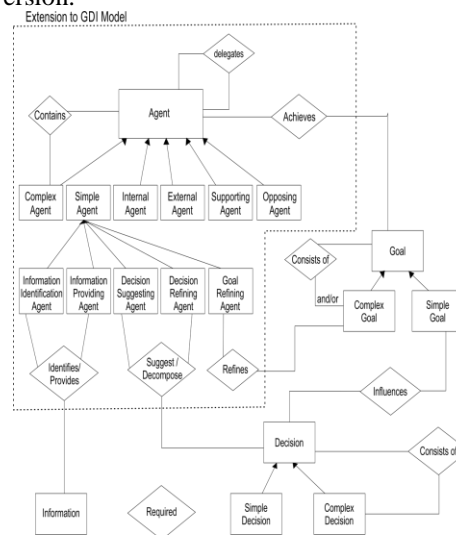
III. PROPOSED DATA WAREHOUSE SYSTEM DEVELOPMENT LIFE CYCLE

A new SDLC model for data warehouse has been provided. It includes 5 ranges primary RE section, past RE phase, conceptual format segment, logical format segment and physical layout phase as shown in table 3.1. The primary phase identifies numerous agents, their desires and the choices to gain the respective desires. The past RE section is used to elicit the preferred facts to help the diverse decisions as diagnosed in primary RE phase. The yield of primary RE phase is characterized as agency version and choice version, whereas the output of secondary RE section is signified as records model. The conceptual format section ambitions to obtain the multidimensional (MD) model. The primary and past RE fashions are utilized in conceptual layout phase to determine the information, magnitudes and measures. These facts, dimensions and measures are organized inside the form of a multidimensional theoretical schema. The primary and secondary RE fashions are based totally mostly on AGDI version, while the enterprise demonstrating, decision demonstrating and facts modelling are finished to apprehension DW requirements.

Phase	Output
Primary RE phase	Organization model, refined organization model and decision model showing agents, their goals and dependencies among agents to achieve the goals
Secondary RE phase	Information model showing dependencies among agents to get the required information
Conceptual design	Data warehouse multidimensional conceptual schema to be derived from primary and secondary RE models
Logical design	Data warehouse logical schema from multidimensional conceptual schema
Physical design	Data warehouse physical model from logical model

Table 3.1 Proposed SDLC for Data warehouse

IV. PROPOSED REQUIREMENT ENGINNERING APPROACH FOR DATA WAREHOUSE DESIGN

In latest beyond, RE phase for classic information systems had been divided into sections specifically, primary RE segment and secondary RE segment [5, 6]. In step with Yu, the primary RE section targets to model and observe agents pursuits and the way they might be addressed, or compromised with the aid of diverse tool and environment options. The accentuation in essential RE section is on data the 'whys' that underlies machine requirements [5] i.e. to seize essential requirements and consciousness of past due RE area is on the perfect and determination of 'what' the gadget need to do, i.e. to seize past due requirements. Be that as it may, most extreme of the RE methodologies [2, 3, 4, 9, 10, 12] for DW don't recognize the essential and auxiliary requirements building fragment. In the ones methodologies, the essential insight is on determining 'what' the DW framework should do, i.e. caught certainties to be kept up inside the DW. Just a couple of objective and model driven strategies [1, 7] for DW have been concentrating on 'whys' that underlies DW measurements requirements (i.e. to catch reason for measurements to be kept up inside the DW). Notwithstanding, none of those methods catch decision requirements [8], i.e. picks aren't found to relate with the objectives and information to be kept up inside the DW. In their system [8], decision requirements are caught without demonstrating merchants and their conditions, i.e. essential requirements aren't demonstrated expressly. This stimulated to signify a unique RE method within the paper to capture every primary and overdue (preference) requirements after which proceed in the direction of conceptual design phase of DW.

Prakash D. [13] told that initially, data warehouse requirements engineering was highly influenced by software/information systems. However, in recent years, new concepts and models specific to data warehouse requirements engineering have been proposed. They discussed the new models and concepts for data warehouse requirements engineering.

The proposed RE approach is exact out inside the subsequent section. Previous to describe the proposed RE method, its legends are proven graphically inside the figure 4.1. All the legends with the exception of fact, size and quantity are based on proposed AGDI model. The suggested RE method for DW layout includes three levels: (i) primary RE, (ii) secondary RE and (iii) conceptual design, schematically shown in figure 4.2. All through primary RE section, kinds of modelling activities specifically, corporation demonstrating and decision demonstrating are finished to seize 'whys' that motivates selection requirements. At some point of past RE phase, statistics modelling activities is done. In facts modelling, marketers pick out necessary data to guide various decisions for attaining their desires. Eventually, all through conceptual

layout phase, information dimensions and measures are identified from primary (organisation model, decisionmodel) and past requirement version (information version).
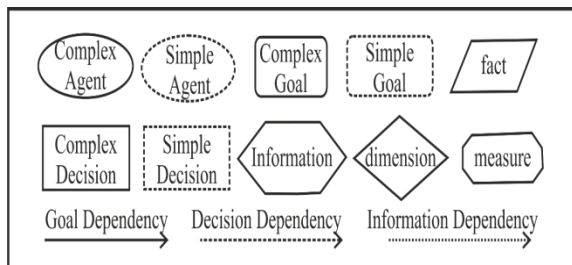

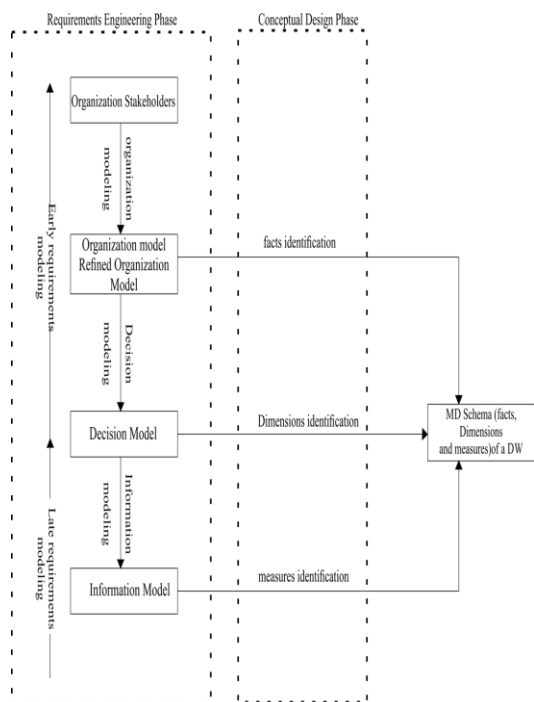Figure 4.1 : Legends of a new requirement engineering approach for data warehouse

Figure 4.2 : New requirements engineering (RE) approach for data warehouse

A.   Primary Requirement Engineering (RE) Section

The organization modelling and decision modelling sports of primary RE phase are designated out as follows:

1)   Organizational Modelling: The business enterprise modelling looks on the administrative background and recognizes the agents and their related goals. It includes sections: (a) agent analysis and (b) aim analysis, which can be provided beneath.

*a) Agent Analysis:*Here operators and their wants are analysed for the business undertaking. Likewise, the reason oppressions between specialists are perceived. The specialists and their motivation subjugations are approved as association rendition. The perceived operator might be a basic/confused specialist or inward/outside operator. The muddled specialist is what's more delicate into entangled/straightforward advertisers in this way producing a chain of importance of merchants. This specialist assessment is rehashed until the point when the greater part of the muddled merchants is inconspicuous into smooth venders. The recognized objectives of simple venders can likewise be straightforward/complex objective. The perplexing dreams are refined inside the objective investigation, offered inside the accompanying.

*b) Goal Analysis:*Here the rationale refining operator may moreover refine the intricate reason into complex/simple objectives and pass on a progression of objectives. That is rehashed till the greater part of the entangled wants is unobtrusive into smooth wants. Additionally, the advertisers may besides designate those simple wants to various retailers for moreover development. Appropriate here, the business association adaptation delivered all through operator examination is comparatively inconspicuous. The unobtrusive organization display demonstrates the chain of command of dreams and the objective conditions among merchants. From that point inside the determination displaying interest, the retailers may also advocate imperative decisions to pick up their straightforward objectives as noted inside the consequent sub stage.

2)   Decision Modelling: The yield of enterprise demonstrating drives decision displaying segment. It involves (i) choices recognizable proof wherein specialist indicates assorted significant picks to profit their wants and (ii) inclination assessment wherein muddled decisions are diffused into simple determinations. This determination displaying leisure activity is preceded till every smooth objective of touchy enterprise variant are depleted. The yield is created as determination demonstrates. The organization form, diffused office model and choice model, as a final product created, develop to be purpose behind displaying data to be kept up in the DW, i.e. catch `whys' that underlies decision requirements as sure in the accompanying.

B)Secondary Requirement Engineering (RE) Section: The output of choice modelling drives this phase. The goal of overdue RE section is to pay attention on 'what' the DW should do for the agency (i.e. to capture secondary requirements). Here, the DW wants to provide applicable statistics to assist numerous decisions for achieving desires of the company. As a consequence, in statistics modelling interest, sellers explore critical information to manual the choices below interest. The records modelling pastime is certain out inside the subsequent sub phase.

*1) Information Modelling:*It includes two levels, (i) statistics identity, wherein agent, itself, may perceive the appropriate data to help specified decisions or can also hang on

supplementary agent for achieving goals within the enterprise. ii) Facts resource identity -agents are find out to present the facts.

C)   *Conceptual Design Section:*This section is used to find out the statistics, dimensions and measures for the implementation of the information warehouse. Primary and overdue requirement fashions are used to identify those and to create MD conceptual schema.

1)   *Fact Identification*: The subtle organization version is used to discover the statistics of a data warehouse. Fact is a crucial detail for facts warehouse schema. Statistics are diagnosed from dreams. Complex desires provide the facts that encompass sub facts. All the goals are used to discover all the records.

2)   *Dimensional Identification:*The choice model is utilized for distinguishing proof of measurements for examining a reality. Estimation character additionally takes after zenith down strategy that begins from the mind boggling want and developments toward straightforward decisions. The progressive system of choice model may also in addition cause measure pecking order. The measurement ID is proceeded until all the supported decisions are mulled over. The refined hierarchical rendition and the determination show each are eluded, while investigating a dating among the perceived records and measurements.

3)   *Measure Identification:*A confirmation is a trait of a reality that is required to be respected in various settings for helping the picks. The ones settings will likewise be utilized as measurements and could expand the measurements as of now analysed in past sub-stage. The measures and their setting of assessment, i.e. measurements are analysed from the records show.

V.   CONCLUSION

The aim of this paper is to capture decision requirements of DW to support decision making activities in the organization that involves various stakeholders.

The paper proposed that the requirements engineering phase of DW must be divided into early RE phase and late RE phase. Early RE phase captured 'why' that underlies decision requirements and late RE phase captured 'what' the DW should do, information related to the decisions. A new RE approach based on AGDI model has been proposed. In the proposed approach, three kinds of modelling activities namely: early RE, late RE and conceptual modelling are described. The early phase identifies various agents, their goals and the decisions to achieve the respective goals. The late RE phase is used to elicit the required information to support the various decisions as identified in early RE phase. The output of early RE phase is represented as organization model and decision model, whereas the output of late RE phase is represented as information model. The conceptual design phase aims to obtain the multidimensional (MD) model. Future efforts can be made for transition from the conceptual model to logical model of data warehouse.

REFERENCES

[1] P. Bresciani, P. Giorgini, F. Giunchiglia, J. Mylopoulos, and A. Perini. TROPOS: An agent oriented software development methodology. *Autonomous Agents and Multi-Agent Systems*, 2003.in Press

[2]. M. Frendi and C. Salinesi Requirement Engineering for data warehouse. In proceeding of 9th international workshop on requirement engineering : foundation for software quality (REFSQ' 2003) pp. 75-82

[3]. M. Boehnlein, and U.A. Ende, Business process oriented development of data warehouse structures. In proceedings of data warehousing 2000, Verlag.

[4]. A Bonifati,. F. Cattaneo, S.Ceri, , A. Fuggetta, and S. Paraboschi, Designing datamarts for data warehouses. ACM Transactions on software engineerin and Methodology, Vol. 10, Issue 4, pp. 452-483, 2001.

[5]. Yu. Modeling Strategic Relationships For Process Reengineering. Ph.D. Thesis. Department of Computer Science, University of Toronto, 1995.

[6]. E. Yu. Towards Modelling and Reasoning Support for Early-Phase Requirements Engineering.In Proc. Third IEEE International Symposium on Requirements Engineering.

[7]. J.N Mazon,., J. Trujillo, M. Serrano, and M. Piattini, Designing Data warehouses: From business requirements analysis to multidimensional modeling. In proceeding of 1st international workshop on requirement engineering for business needs and IT alignments (REBNITA' 2005) pp. 44-53.

[8]. N. Prakash, and A. Gosain An approach to engineering the requirements of data warehouse.Requirement Enginnering Journal, Vol.13, Issue 1, pp. 49-72, Springer 2008.

[9]. J. Shiefer, and B. List and R.M. Bruckner A holistic approach for manageing requirements of data warehouse systems. In proceeding of of 8[th]Americas conference on Information system. pp. 77-87, 2002.

[10]. R. Winter, and B. Strauch A method for demand driven information requirement analysis in data warehousing projects. In proceeding of 36th Hawaii international conference on system science, 2003.

[11]. N. Prakash, and A. Gosain, Requirement driven data warehouse development. In short-paper proceeding of 15th conference on advanced information system engineering (CAiSE' 2003), pp. 13-16.

[12]. F.R.S Paim,. and J.F.B. De Castro, An approach for requirement definition and management of data warehouse system.In Proceeding of the 11th international requirement engineering conference (RE 2003) pp. 75-84, IEEE computer society.

[13]. D. Prakash and D. Gupta "Data warehouse requirements engineering: an emerging discipline" in International Journal of Business Information System Volume 23 Issue 2, January 2016 Pages 194-211.

Authors Profile

Mr. S. Bhardwaj pursed Masters in Computer Management from University of Pune, India in 2001 and Masters of Technology (Compter Science and Engineering from GGS Indraprastha University in year 2008. He is currently pursuing Ph.D. and currently working as Associate Professor in College of Computing Sciences and Information Technology, Teerthanker Mahaveer University, Moradabad since 2001. He has published more than 15 research papers in reputed international journals and conferences including IEEE. His main research work focuses on Software Engineering, Data warehouse and Agent concepts based education. He has 15 years of teaching experience and 6 years of Research Experience.

Dr A. K. Goyal pursed his M.phil (Computer Science) and Ph.D. (Computer Science) from Gurukul Kangari Vishwavidhyalaya, Haridwar, Uttrakhand and currently working as System Manager and Head, University Computer Centre, Gurukul Kangari Vishwavidhyalaya, Haridwar, Uttrakhand. He is a life member Computer Society of India & Indian mathematical society. He has published more than 20 research papers in reputed international journals and conferences including IEEE. His main research work focuses on Computer Network & Security, Software Enginnering, Data Mining, DBMS based education. He has more than 30 years of experience.