Data Base Warehousing and mining –Business visual tools Interaction with UML Profile Modeling

H.V.Ramakrishnan
Phd, Dean –Information Techonolgy
DR M.G. R Educational & Research Institute
Maduravoyal-Chennai-Tamilnadu-India

Padmanathan Anantharaman
Research Scholar
DR M.G.R Educational & Research Institute
Maduravoyal-Chennai-Tamilnadu-India

ABSTRACT

This paper outlines the new approach to be taken by integrating Business Management charts with Data visualization techniques for making the understanding of the students with computer science background and business management background easy. Normally the paper Data Mining and warehousing is a computer science and engineering paper and expertise can be obtained only when business processes are understood. On the other hand this paper if taught in Business management may be very useful but the orientation of the paper has to be different. What is common in both dimensions is flow chart or business process diagrams.

General Terms

This paper outlines briefly the approach. It has used the existing commercial packages samples so that readers can buy these packages if they are further interested.

Keywords

Business Motivation, Business Process Modeling, Unified ModelingLanguage, Casetools, Order Processing,Data warehouse Usage

1. INTRODUCTION

Enormous material is available in the area of data warehousing and data mining (1-5) and all the books and papers indicate the output should be presented by way of graphs and charts as the business personal Computer Science Technologists are not given an idea of Business to enable them to understand the business. Besides Business people when they present their ideas in flow charts Computer Science understand better. Such integration is need of the hour.

Business Process Presentations – There are number of tools available to present business process. The tools available to do (Select Architect is a scalable modeling tool)

(i) Business Motivation Modeling (BMM),

(ii) To motivate Business Process Modeling (BPM),

(iii) Unified Modeling Language (UML) etc.

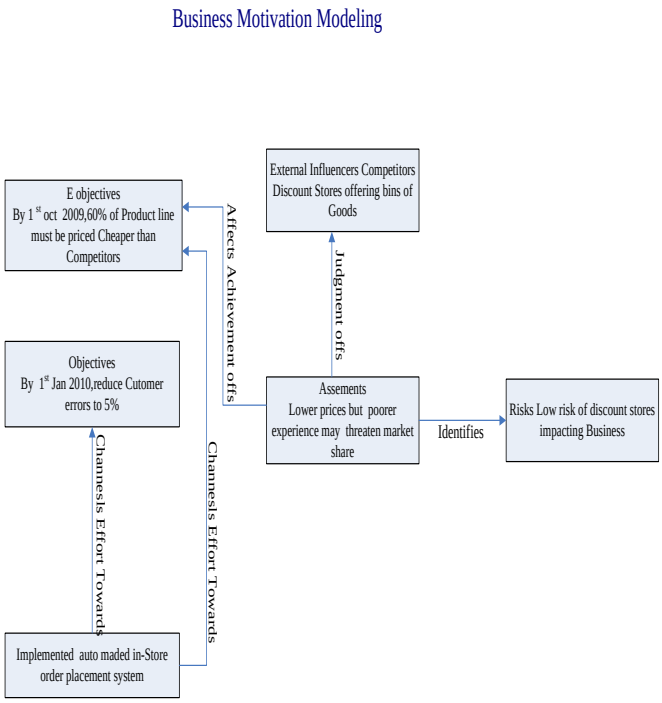


Figure 1: Business Motivation Modeling

This is an extract of the diagram for Business Motivation Modeling(http://www.selectbs.com/analysis-and-design/business-motivation-modeling-bmm)This outlines the tasks and difficulties associated. Here tasks and deadlines areoutlined which is self-explanatory and which takes pages and pages of descriptionin a conventional manner.

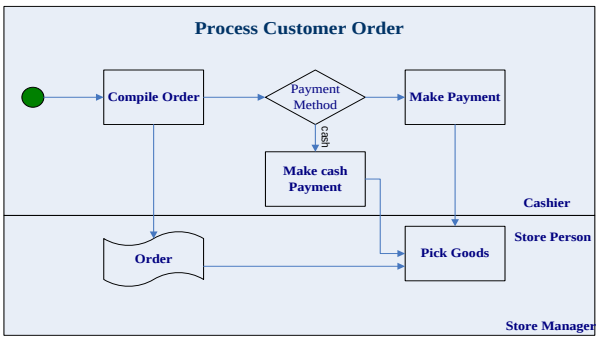


Figure 2: Customer Order

The Business process for a customer order in a business process where after paying cash or credit card payment he receives the goods is outlined. Again pages and pages of paper description can be avoided.

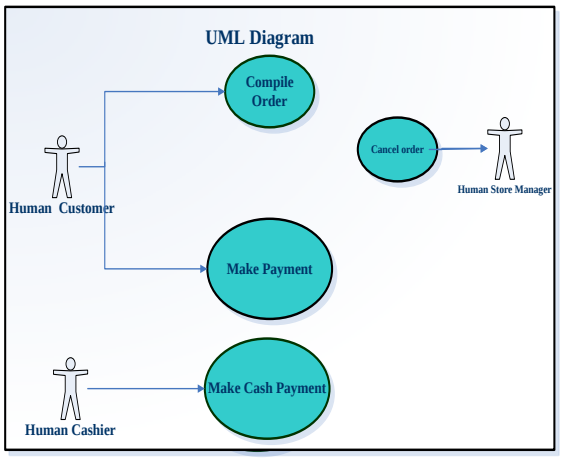


Figure 3: UML Diagram

The above process is described by an UML diagram where human beings (or) computer routines involved are shown and their activities are shown. Again it takeslarge volume of paper and still it will not be clear in standard writing.

2. CASE TOOLS

One of the famous case tools available is SMART DRAW. Case tools can easily enable visualization of complex business steps. Capital Budgeting which is essential in Government and Business sectors is easily visualized in the diagram. Indication of the steps involved such as approval by department manager, approval by CFO and finally by CEO graphically makes the understanding easy for a computer science professional that can then make necessary modules to aid different managers. (Department manager – priority resolution among existing tasks, CFO – checking with fund flow, CEO – comparative advantage over competitors). Besides review periodicity is also indicated.

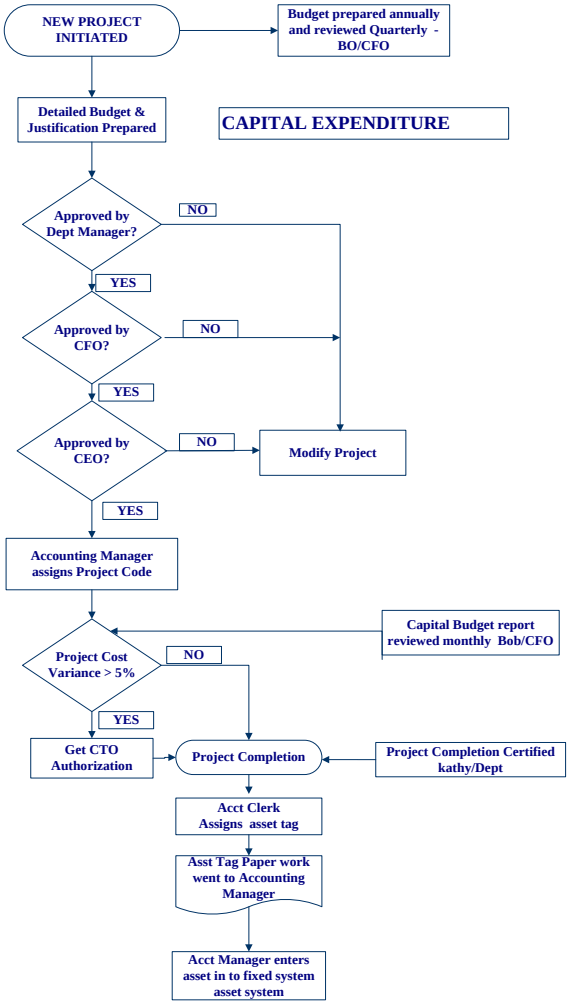


Figure 3: Capital Expenditure

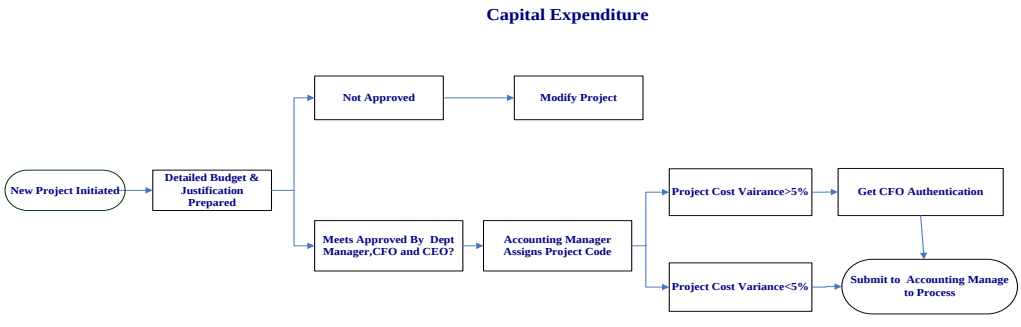


Figure 4: Capital Expenditure-Failures (Revision of Proposals)

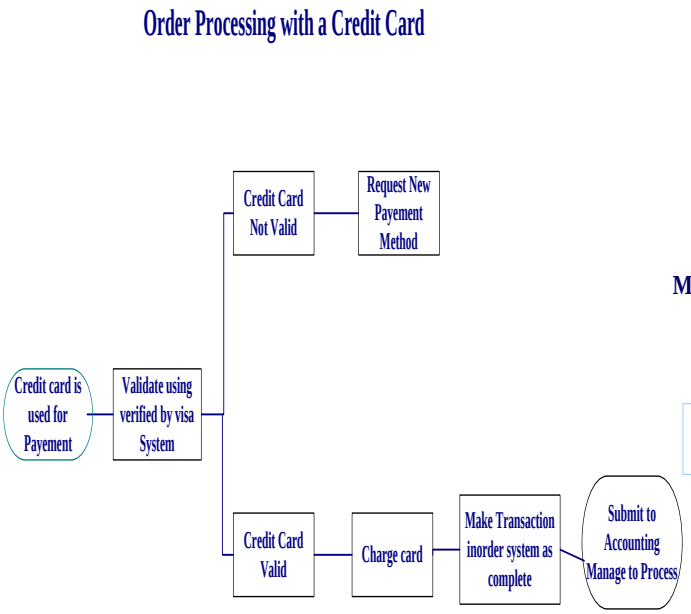


Figure 5: Order Processing with a Credit Card

Besides a workflow diagram helps people understand what to do in case of failures such as revision of proposals etc.

An order processing with credit card workflow is given

below.

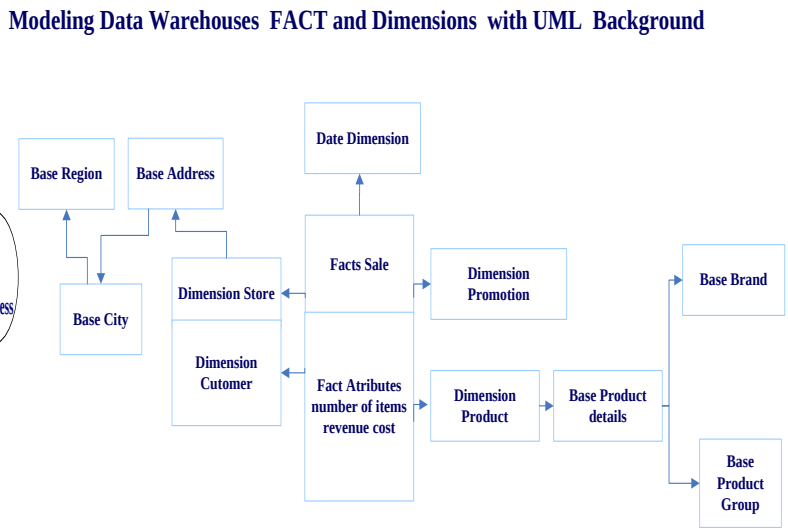
This clearly outlines in case credit card not valid new methods which may be cash payment or new credit card or handing over the person to police which has to be decided. Such operations can be automatically done through computer programs. The flow chart for such a process is given below with computer or human being clearly indicated.

For number of procedures, flow chart, UML, workflow diagrams can be drawn. After acceptance of these diagrams, using automatic tools, generation of tables and fields can be done.

3. MODELING DATA WAREHOUSES: BACKGROUND WITH UML

Our approach applies UML to the Data Warehousing domain. It is aimed at encompassing all the different ways that users may use a DWH. Our goal is to provide an overview over all aspects of DWH usage, not only focusing on the data model. Nevertheless, due to the special characteristics of DWH data, it is necessary to take the data model especially into account.

DWH applications involve complex queries on large amounts of data, which are difficult to manage for human analysts. In Data Warehousing, data is often organized according to the



multidimensional paradigm, which allows data access in a way that comes more natural to human analysts. The data is located in n-dimensional space, with the dimensions representing the different ways the data can be viewed and sorted (e.g. according to multi-dimensional model, also called star schema or fact schema, is basically a relational model in the shape of a star (see Fig 6for an example). At the center of the star there is the fact table. It contains data on the subject of analysis (e.g. sales, transactions, repairs, admissions, expenses, etc.). The attributes ofthe fact table (e.g. cost, revenue, amount, duration, etc.) are called measures or fact attributes.

The spokes/points of the star represent the dimensions according to which the data will be analyzed (sorted/aggregated by data, by store). The dimensions can be organized in hierarchies that are useful for aggregating data (e.g. store, city, region, country). Stars can share dimensions, thus creating a web of interconnected schemas that makes drill-across operations possible

Figure 6: Modeling Data Warehouses: Backgroundwith UML

3.1 UML PROFILE FOR MODELING DATAWAREHOUSE USGAGE

This section introduces our UML Profile for Modeling DWH Usage. We use the extension mechanism of UML and import elements from a well-known UML Profile of the Data Warehousing domain, in order to achieve a conceptually sound model, with (a) tool support and (b) well-known notation elements.

Its stereotypes and enumerations, and also shows which classes are used as base classes of the stereo-types. We import the Profile of Luj´an-Mora et al. and also some packages of the UML metamodel (for the convenience of not having to use fully qualified names). Table 1 describes the characteristics of the stereotypes not shown above. Usage is defined as an Information Flow, which is a type of directed relationship that specifies that information items circulate from sources to targets1.Information flows are defined in UML as a very general concept to be used in Sources and targets of an information flow may be: Actor, Node, Use Case, Artifact, Class, Component, Port, Property, and Interface, Package, and Instance Specification "preliminary models, before having taken detailed modeling decisions on types or structures. One other purpose of information items and information flows is to abstract complex models by a less precise but more general representation of the information exchanged between entities of a system" [3]. This makes information flows very for our purpose, which is to provide models that capture an overview of the general structure of DWH usage. The direction of the Usage arrow indicates whether the users actively initiate the access to the DWH or whether they wait to receive messagesfrom the system, i.e. push or pull mode. A user analyzing OLAP data would be pull, whereas an E-mail alert is push.

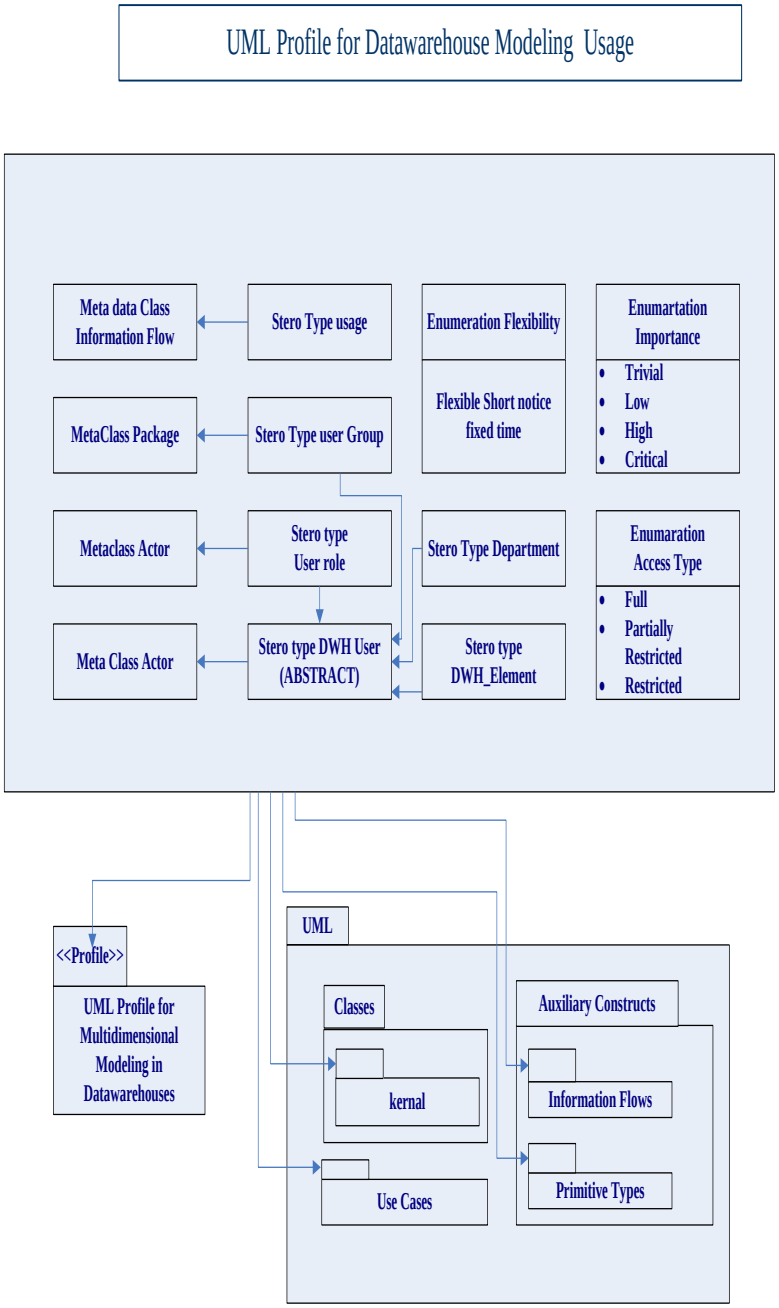


Figure 7: UML Profile for Modeling Data warehouse Usage

4. CONCLUSION AND FUTURE WORK

A suggestion for using Visual Tools has been given with examples from commercial packages such as Solution Architect and Smart Draw. Visualization makes the business man job of explaining easy and understanding by the computer engineer also very easy. The Project manager or CEO or CFO can explain visually the requirements further as complete automation is nearly impossible as country to country business practices vary. Today's DWH systems provide many different services to different kinds of users. In order to have a big picture of the current situation and to visualize future scenarios, people involved in designing and managing today's DWH systems need an overview of all these different ways the DWH is being used In this paper, we have introduced the UML Profile for Modeling DWH Usage for modeling the different kinds of DWH usage on a conceptual level. It distinguishes four perspectivesof usage (access control, temporal intensity, temporal flexibility and importance) as well as active or passive usage, and allows to model details of the users such as their skill level, number of instances, functional grouping or organizational affiliation. We base usage on UML information flows, which are intended for a general representation of information

exchanges. It will continue extending same subject Area with UML

5. REFERENCES

[1] Alex Berson-Stephen. J.Smith (2006),"Data warehousing-Data Mining &OLAP", TMH 2006.

[2]Arun K Pujari (2002), "Data mining Techniques", University Press (India) Ltd., 2002.

[3] Doug Rosenberg and Matt Stephens. (2007)" Use Case Driven Object Modeling with UML: Theory and Practice" by (Hardcover - Jan 11, 2007)

[4] JimArlow and IlaNeustadt (2005)."UML 2 and the Unified Process: Practical Object-Oriented Analysis and Design" (2nd Edition) (Paperback - Jul 7, 2005)

[5] Jiawen Han and MichelineKamber (2005), "Data mining concepts and Techniques", Mortgan Kaufmann Publishers, 2005.

[6]. Howard Podeswa (2009) "UMLfor the IT Business Analyst "by second edition (Jun-2009)

[7] H.V.Ramakrishnan (2011)" Data Base Warehousing and mining - Business visual tools"IJSTM, Vol. 2 Issue 3, July 2011http://ijstm.com/papers/ramakrishnan_0709.pdf

[8] Inmon, W.H., Hackathorn, R.D (1994)"Using the data warehouse"Wiley-QED Publishing, Somerset, NJ, USA (1994)

[9] Margaret H.DunHam (2005)."Data mining – Introduction and advanced topics", Pearson Education, 2005.

[10] Russ Miles and Kim Hamilton (2006)."Learning UML 2.0 "(Paperback - May 2, 2006)

[11] Sam Anahory, Dennis Murry (2003), "Data warehousing in the real world", Pearson Education, 2003.

AUTHOR'S PROFILE

Author 1:Dr H.V. Ramakrishnan

Dr H.V. Ramakrishnan received the PhD Degree in Electrical and Electronics Engineering from Indian Institute of Technology, Madras, 1985 and The M.Sc. (Eng.) distinction Degree in Thiagarajar College of Engineering ,Madurai 1975,Where he is currently working as Dean Information Technology and Computer science and Engineering in DR M.G.R Educational and Research Institute-Maduravoyal-Chennai Tamilnadu,India.He worked as Registrar in Bharath Deemed University in Chennai,Tamilnadu,India His Research interest areas are Data warehouse, Data mining, Project management,Network,Operating systems

Authors name(s), DR H.V.Ramakrishnan
position(s) and affiliation (s): Dean-DR.M.G.R Educational and Research Insitute
Detail address: DR.M.G.REducational and Research Insitute-Maduravoya1-chennai-Tamilnadu

Author 2: PadmanathanAnantharaman

PadmanathanAnantharaman received Master of Computer Applications Degree from Annamalai University, Chidambaram, Tamilnadu, 2007, andB.Sc (Computer Science) from Madras university, 2003, Where he is currently working toward the PhD degree in DR M.G.R Educational and Research Institute-Maduravoyal-Chennai From 2008-2012.His Research interest areas are Data warehouse, Data Mining, Project management, Project failure causes Six Sigma, Software, Database, Query performance.

Authorsname(s), PadmanathanAnantharaman
Position(s) and affiliation (s): ResearchScholar and DR.M.G.R Educational and Research Institute
Detail address: DR.M.G.REducational and Research Insitute-Maduravoyal-Chennai-Tamilnadu
Employed with: HCL America –Inc-USA