DATA MODEL QUALITY FRAMEWORK

Vassil Milev , Stefan Kalchev

Abstract: The data quality has a very important role in the business decisions process. For a good quality, a good models and good procedures are needed. The paper presents a framework for assessment of the process of data modelling, based on the maturity model of Harold Kerzner, known as Project Management Maturity Model (PMMM). Based on this framework, a concrete indicators/metrics could be described. They can be used for specifying the maturity level reached by the organization in the field of data modelling.

Key words: PMMM, Data Modeling, Maturity Model, Enterprise Architecture, TQM

РАМКА НА КАЧЕСТВОТО НА МОДЕЛА

Васил Милев, Стефан Калчев

Резюме: The data quality has a very important role in the business decisions process. For a good quality, a good models and good procedures are needed. The paper presents a framework for assessment of the process of data modelling, based on the maturity model of Harold Kerzner, known as Project Management Maturity Model (PMMM). Based on this framework, a concrete indicators/metrics could be described. They can be used for specifying the maturity level reached by the organization in the field of data modelling.

Ключови думи: PMMM, Data Modeling, Maturity Model, Enterprise Architecture, TQM

Доклада е публикуван в :

INTRODUCTION

The choice of proper data presentation is one of the most important tasks in the whole life cycle of the information system. Data modeling is only a small part of entire development of the system; nevertheless its impact on the final result to a great degree depends on it [11]. The data model is important tool for development of the enterprise architecture.

The data quality is the soul of the organization and the flow and the use of information should be clearly described. The major component of the data architecture is the data model that usually appears as a graphical presentation of the data structure. Many models are developed over and over again. For every good model, there are a dozen bad.

The aim of this paper is to present a framework for estimation of the quality of the process of development of the data model that could be used for improvement of the processes of data analysis by assessment of the maturity level of the organization in this field.

THE IMPORTANCE OF THE DATA MODEL

The data model is a major factor defining:
- The flexibility of the system.
- The integration with other systems.
- The ability of the system to fulfill user requirements [6,8].

WHAT IS THE QUALITY OF THE DATA MODEL

In publications on the quality management there are often distinction between the quality of the product and the process. The quality of the product (data model [7]) is focused on the characteristics of the product. The criteria for the quality of the data model are used for their estimation, for detecting and removing of defects and for comparison between models. This the traditional approach for ensuring the quality [3,5].

The quality of the process is focused on the activities of the creation of the particular data model. The criteria of the quality are used improvement of processes of the analysis and development of data models, by assessment of the maturity level of the organization in this field. This is the Total Quality Management (TQM [8]) approach for the quality.

REQUIREMENTS FOR HIGH-QUALITY DATA MODEL
I deally the high-quality model should [3]:
- Respond to users and business requirements.

- Be understandable to users.
- Be stable and flexible enough for extension and modification along with business processes.
- Be used repeatedly.
- Be suitable for linking with other data models used in the organization.

FRAMEWORK FOR ESTIMATION OF THE QUALITY OF THE PROCESS OF DEVELOPMENT OF THE DATA MODEL

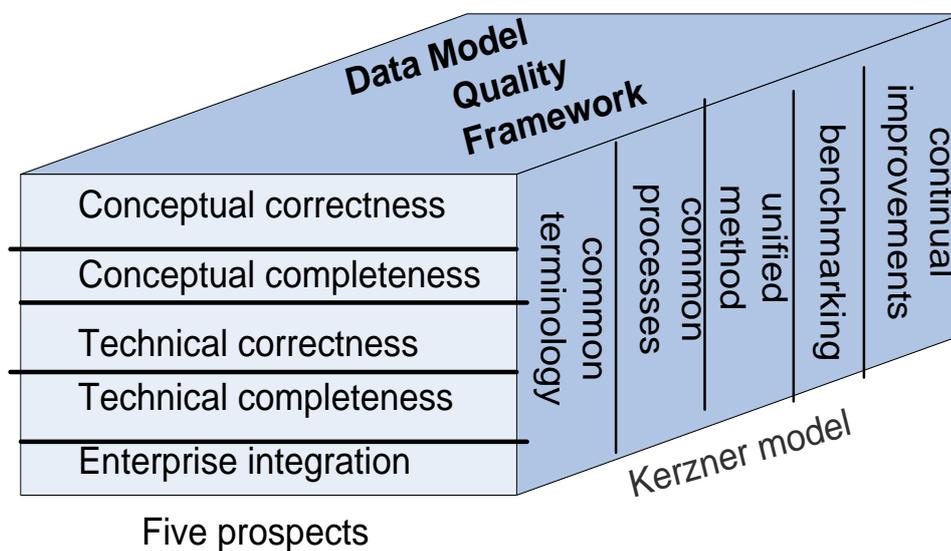Five prospects are given in next figure 1.

Figure 1. Framework for estimation of the quality of the process of development of the data model.

Conceptual correctness.

It is required that the data architecture exactly outlined business objects of the enterprise necessary for the whole business processing. Achievement of conceptual correctness depends on the ability of reasonable transformation of the information of interest from the real world towards structural representation with semantic language. Defining the conceptual correctness is one of the most difficult aspects of quality estimation and data models development.

Conceptual completeness.

Means that the data model embraces all the objects in business domain it aims to represent. The estimation of the quality of the data model depends on the external factors such as governmental and legal requirements, financial constrains etc. These circumstances exercise considerable influence and should be taken into account in data model development.

Technical correctness.

Technical correctness presumed that all of the objects in data model don't break the rules of syntax of the chosen language. Syntactical correctness means that in the data model the main conventions of graphical presentation are properly used and the model adheres to generally accepted rules of the method. After the rules are established, they should become part of the syntax and the model should correspond to this rules.

Technical completeness.

Technical completeness means that all objects and their components in the data model are present at corresponding level of data architecture.

Enterprise integration.

Enterprise integration supposes that the data model is balanced with other elements of the enterprise architecture. The data model is related and synchronized with production, business, services and the technical components of enterprise architecture.

A Kerzner model.

The Project Management Maturity Model (PMMM) of Harold Kerzner [2] consists of five levels as everyone represents a different degree of maturity of the processes in data model development (see figure 2). The following rules are valid for the model:

- Every level begins after completion of the previous one.
- The levels could overlap.
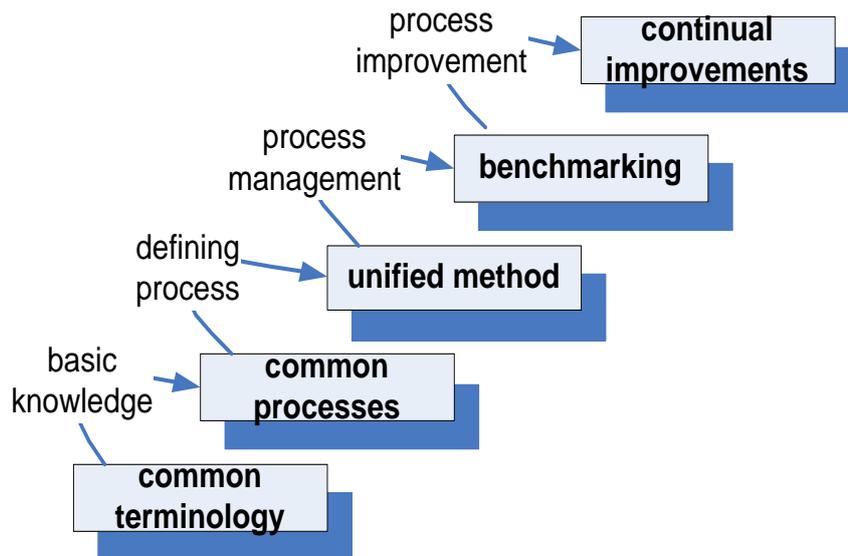- For every level are developed indicators that embraced the five prospects of framework.

Figure 2. The Project Management Maturity Model

Level 1 – common terminology.

At this level the organization has realized the importance of data management and aims at data modeling and using it of full value [1,10]. Research and collecting the basic knowledge of the field begins.

Level 2 – common processes.

The organization has collected knowledge in data modeling and has realized the necessity of common processes for management, preservation and supporting the quality of the data, guaranteeing standardization of the activities and repeated usage.

Level 3 – unified method.

The organization has realized the importance of using unified method, based on architectural approach [4,9], of simultaneously development of the architecture of the business processes, data, applications and infrastructure.

Level 4 – benchmarking.

At this level the organization study the best practices in the field of data modeling (EIM [10]) and implement the ones that correspond to its aims and purposes.

Level 5 – continual improvements.

At this level the organization has created and keeps a good level of data management and constantly checks up their quality and usage.

CONCLUSION

Presented framework in this paper is a structure bringing together indicators for quality of the data models. These indicators are represented in five prospects and phases for defining the attitude of the organization to data management and especially to data modelling.

The developed framework is based on Project Management Maturity Model (PMMM) of Harold Kerzner. It is aimed at represent in broad terms directions in which an organization should develop knowledge about data modelling.

Based on this framework, a concrete indicators/metrics could be described. They can be used for specifying the maturity level reached by the organization in the field of data modelling.

REFERENCES

1. Data model PER, http://tuj.asenevtsi.com/Public/IndexPublic.htm
2. Kerzner H. R, Using the Project Management Maturity Model: Strategic Planning for Project Management , John Wiley & Sons Inc, 2005
3. Lubyansky A., Using Data Model Patterns to Build High-Quality Data Models Analyst, Project Performance Corporation-Part of the AEA group, http://www.ppc.com/assets/pdf/white-papers/Data-Model-Patterns.pdf
4. Milev V.,Kalchev S, Stefanova M., Architectural approach, based synchronization business and information technology, International Scientific Conference "MATHTECH 2010" Shumen, Bulgaria
5. Moody D. L., School of Business Systems, Monash University Melbourne, Australia, Measuring the Quality of Data Models: An Empirical Evaluation of the Use of Quality Metrics in Practice, http://is2.lse.ac.uk/asp/aspecis/20030099.pdf
6. Moody D.L. , G.G Shanks, Improving the Quality of Data Models: Empirical Validation of a Quality Management Framework, International Journal of Information Systems, 2003
7. Reingruberand M.,T. Rinaman, Data Modeling: A Systematic Approach to Build it Right the First Time, http://www.information-management.com/infodirect/2009_129/data_modeling-10015659-1.html
8. Tujarov H., TQM, http://www.tuj.asenevtsi.com/TQM2009/Index.htm,
9. Tujarov H., V. Milev, S. Kalchev.,Enterprise Architecture, http://tuj.asenevtsi.com/EA/EA01.htm,
10. V. Milev, S. Kalchev , Extended Information Model (EIM), http://tuj.asenevtsi.com/Public/IndexPublic.htm
11. WITT G., C. SIMSION. Data Modeling Essentials: Analysis, Design, and Innovation, The Coriolis Group, 2000