Enterprise Knowledge Modeling and Data Mining Integration

Auksė Stravinskienė
Kaunas Faculty of Humanities of
Vilnius University
Kaunas, Lithuania
Aukse.Stravinskiene@khf.vu.lt

Saulius Gudas
Kaunas Faculty of Humanities of
Vilnius University
Kaunas, Lithuania
Saulius.Gudas@khf.vu.lt

***Abstract*—The control view-based approach to integration of enterprise modeling and data mining is the focus of the paper. The analysis and evaluation of enterprise modeling methods and languages (ARIS, UEML, BPMN and Value Chain Model) for elicitation of enterprise knowledge components are presented. The knowledge components of control view-based Enterprise Meta-model are discussed. The principles of identification of enterprise knowledge components and their integration to data mining process are defined.**

Keywords-Enterprise Modeling; Data mining; Control view; Enterprise Meta-model; Knowledge

I. INTRODUCTION

The rapid development of Business Intelligence (BI) systems requires elicitation of enterprise management experience as knowledge and using it for development of Intelligent Information Systems [1]. It is important to develop an adequate business model that matches the organization's management information interactions.

The enterprise modeling for information systems (IS) engineering is aimed at to describing the scope of the organization's management activities – knowledge and data transformations in enterprise management systems. Enterprise model is required to improve the company's performance, describing the operational sequences, responsibilities and relationships as well as to identify must systems that duplicate each other or are otherwise inappropriate [2].

Even though enterprise models are different, they all serve the purpose of developing enterprise models and methods, which could be used for the IS development as knowledge-based integration of data mining, online analytical processing (OLAP). Using such models it is easier to understand enterprise process, enterprise knowledge and structure, and to improve the enterprise efficiency.

Enterprise models for information systems engineering are developed using DFD, ARIS, UML, IDEF3, IDEF0 notations [2,8,9,10]; especially, BPMN and UEML notations are effective for enterprise modeling.

It is important to identify the enterprise management knowledge which can be useful to improve enterprise efficiency, to help user and to manage data mining process.

There are two types of models for IS engineering: a) enterprise models that describe empirical information about domain and b) enterprise management models that verify empirical information against formal theories or predefined structures [3]. Empirical enterprise models as well as enterprise management models can be developed using all previously mentioned business process modeling notations. A specific characteristic of the Enterprise management models is that they are based on obligatory information relationships of enterprise components. For instance, they form a feedback loop between the components of managed process and enterprise management function.

Information need of enterprise activities depend on the certain patterns of enterprise management process, identified in the Enterprise management model [4]. The knowledge-based approach to IS development is based on the Enterprise Meta-model [4] and co-relates with knowledge modeling and discovered of enterprise knowledge components. **T**ypes of enterprise model components and their interactions are defined in Enterprise Meta-model [4]. Therefore, an Enterprise Meta-model component could be used for acquisition of enterprise knowledge components.

These principles of enterprise management modeling could be used for improvement data mining process. Reasonable selection of a data mining method and an algorithm is important to successful implementation of the process of analysis and the final result of data mining.

Enterprise and enterprise management models are composed of components, which will be regarded as sources of enterprise knowledge.

On the basis of knowledge-based engineering principles it is essential to identify the enterprise's knowledge components that are involved in enterprise management.

This paper analyses EMM, ARIS, UEML, BPMN enterprise modeling methods and identifies knowledge-based components [2,8,9,10]. A control view-based approach to identification of knowledge components and their integration to data mining process are presented in the paper.

 ISBN: 978-1-61208-160-1

The paper structured into 3 parts. First, Enterprise knowledge and data mining integration are defined. Second, Knowledge-based analysis of Enterprise modeling methodology is described. Third, Enterprise knowledge model's integration into data mining is analyzed.

II. ENTERPRISE KNOWLEDGE AND DATA MINING

Data mining is an iterative process of searching in large amounts of data or in large data repositories for hidden, previously unknown and potentially useful information, knowledge, a set of patterns and relationships between data. Data mining is a process (based on the CRISP-DM model) composed of consistent set of six stages [5]: Business understanding, Data understanding, Data preparation, Modeling, Evaluation, Deployment.

These six data mining stages construct a single cycle of process. In order to gain better data mining results, data mining process could be aligned with Enterprise management modeling [6]. We propose to integrate the knowledge base in a data mining process, on the basis of which the data mining process could be controlled.

The component of Enterprise management model could be used in data mining such as Enterprise knowledge held by an expert. Its use can develop a more rational data mining queries. We propose to use the Knowledge base in which Enterprise management knowledge would be collected. More accurate and structurally interesting information can be extracted from Knowledge base by adding business rules, knowledge of experience about management functions.

Knowledge-based management model (Figure 1) is being established according to this principle: Enterprise management is structured as interactions of Management functions and Enterprise processes [6]. This Knowledge-based management model illustrates the interaction between Enterprise processes (Pj) and Management functions (Fi), which serves for the Knowledge–based Enterprise's information system. The model (Fi x Pj) consists of Enterprise's information system, Management functions (F1xFi), Business goals (G), Enterprise processes (P1xPj).

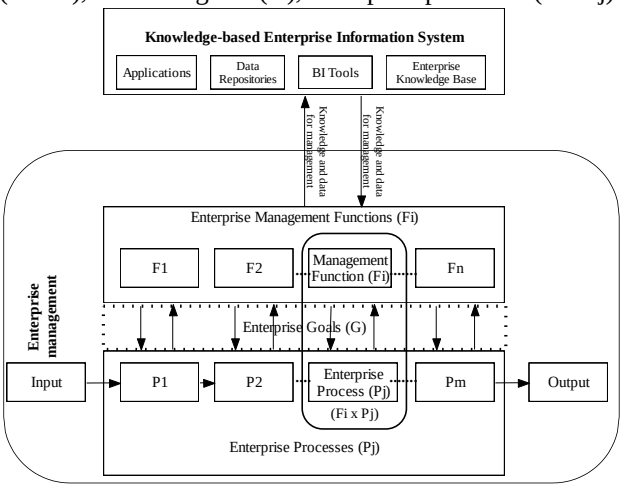


Figure 1. Knowledge – based management as interactions of Management functions and Enterprise processes

Knowledge-based Information System consists of Data Repositories, Applications (Management IS), as well as of Enterprise Knowledge base, which supports Applications and BI Tools with knowledge components.

The influence of Enterprise Goals to Enterprise management function is critically important as the final result of management depends on the Enterprise Goals.

Control view-based Enterprise management structure is presented in Figure 2.

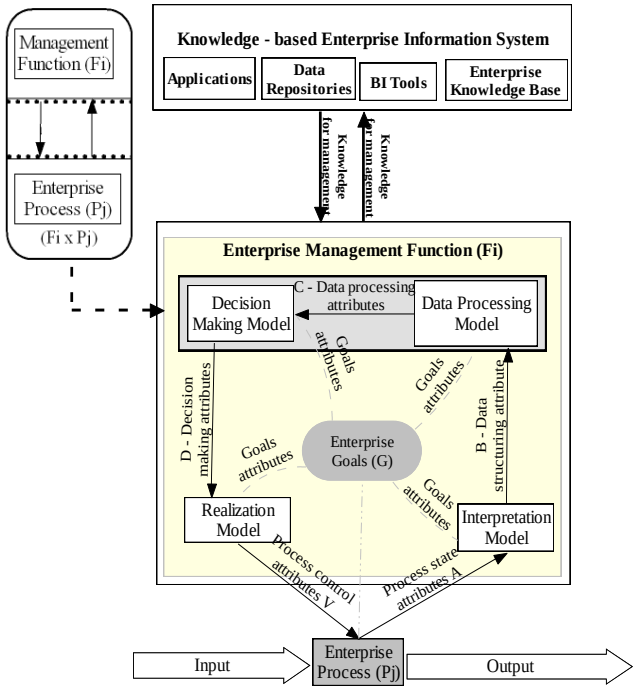


Figure 2. Control view-based structure of Enterprise management function Fi (Gudas S., Skersys T., Lopata A., 2004)

Enterprise management is a model identifying Enterprise components, components of Enterprise management (Data, Knowledge, Goals) and their interactions [6]. "EMC (Elementary Management Cycles) is formalized description of the interaction of Process and Function – as two core components of Enterprise from the control point of view"[7].

Application of this model is important because this model describes the essential elements of Enterprise management and control function and their interactions. Besides, information flows are identified (Figure 2). The model reflects the interaction of management function and Enterprise process in which the elementary management activities (Interpretation, Data processing, Decision making, Realization of decision) compose a cycle– an Elementary Management Cycle (EMC).

Decomposition of Knowledge management system interactions with Business Intelligence Tools and Applications are presented in Figure 3.

Figure 3 displays a detailed scheme of Enterprise management function integration into data mining (aligned with Figure 1 and 2), in which it is clearly shown what interfaces there are between Business Intelligence Tools,

Applications, Data repositories and Enterprise knowledge base and which data is integrated from the Enterprise Knowledge base.

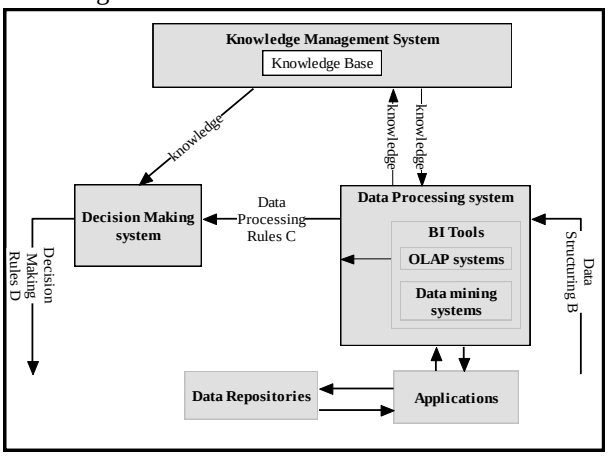


Figure 3. Integration of Knowledge management system and data mining

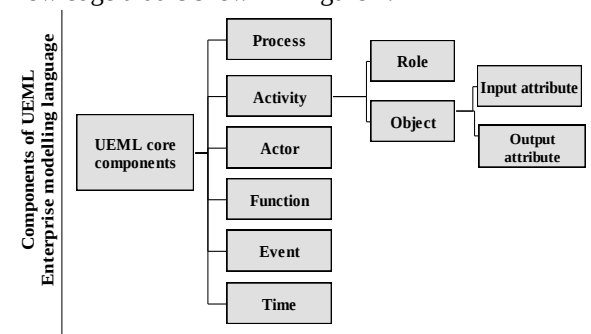
Knowledge, which is stored in Knowledge Base, is going to be used in the process of data mining for algorithm modification.

III. KNOWLEDGE-BASED ANALYSIS OF ENTERPRISE MODELING METHODOLOGY

This section provides the analysis of UEML, ARIS, BPMN and EMM enterprise modeling methodologies and languages from the aspect of knowledge. Components of structure of an enterprise model are described and knowledge components of a certain method are identified.

A. *UEML Enterprise modeling language*

Unified Enterprise Modeling Language (UEML) is a language of Enterprise processes' modeling, which is aimed at facilitating the integration of different modeling languages in a company [8]. UEML model stores Enterprise knowledge that is shown in Figure 4.



Enterprise Knowledge elements	
Process	Process ID, Process Attributes
Activity	Activity ID, Activity attributes
Actor	Actor ID, Actor Attributes
Function	Function ID, Function Attributes
Event	Event ID, Event Attributes
TIme	Time ID, Time Attributes
Role	Role ID, Role Attributes
Object	Object ID, Object Attributes
Business Rules	

Figure 4. The major components of the UEML

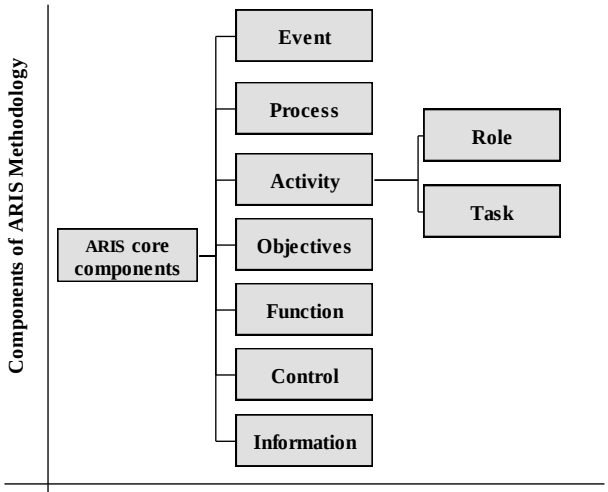
Figure 4 shows the composition of components of the UEML model: process, activity, actor, function, event, time, role, object, input and output. Function is defined as a component, which consists of processes and interactions of activities (including feedback loop). At the Knowledge base, each component stores a certain attribute or value that identifies it. The structure of UEML model can be described by these constructs:

UEML= {Process, Activity, Actor, Function, Event, Role, Enterprise Object (Product, Order, Resourse)}

B. *ARIS Enterprise modeling methodology*

Architecture of Integrated Information Systems (ARIS) - method of Enterprise modeling [9]. The main concepts of ARIS are a) the architecture for describing business processes; b) modeling tools for Event-Driven Process; c) comprehensive computer-aided Business process management. This methodology includes design, implementation, optimization and controlling of Business process. The uniqueness of this methodology is modeling from the control view point. Figure 6 shows a hierarchical structure of components of the ARIS method, in which constituent parts of the method are reflected.

ARIS= {Process, Activity, Objectives, Function, Control, Information}



Enterprise Knowledge elements	
Process	Process ID, Process Attributes
Activity	Activity ID, Activity attributes
Task	Task ID, Task Attributes
Objectives	Objective ID, Objective attributes
Control	Control ID, Control attributes
Function	Function ID, Function Attributes
Information	Information ID, Information attributes
Role	Role ID, Role attributes
Business Rules	

Figure 5. The major components of the ARIS method

C. BPMN Enterprise modeling notation

Business Process Management Notation (BPMN) is a notation for the formation of business processes' flow models (Figure 6). It is coordinated by Object Management Group (OMG) standards' organization that focuses on improving business processes. The basis of this methodology is the possibility for representatives of different fields of business to understand business processes that take place in an organization and to reveal them in a single modeling notation [11]. Composition of BPMN model can be described by constructs that are presented in Figure 6:

BPMN= {Pool, Process, Task, Activity, Gateway, Lane, Association, Event, Data Object}

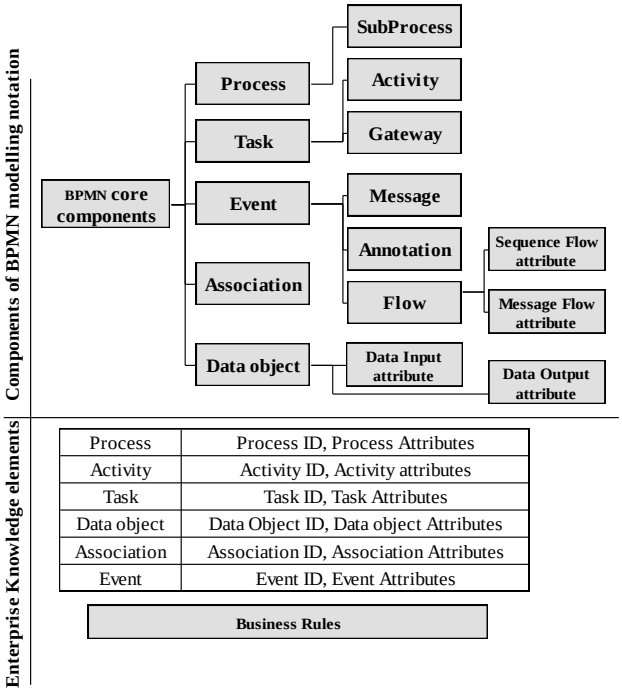


Figure 6. The major components of the BPMN model

D. Enterprise Meta-model

EMM is a model of managed processes based on the structure of Elementary management cycle (EMC) that can be used for Knowledge-based systems modeling [4]. The main principle of this model's structure is the evaluation of an aspect of Enterprise's management by modeling informational interactions of Enterprise processes and Enterprise management functions.

The control view-based approach to Enterprise modeling results in the Enterprise Meta-model (EMM)[4]. The EMM model is a formal description of the Enterprise management knowledge, in which major types of Enterprise components, their types of interactions are reflected (Figure 7).

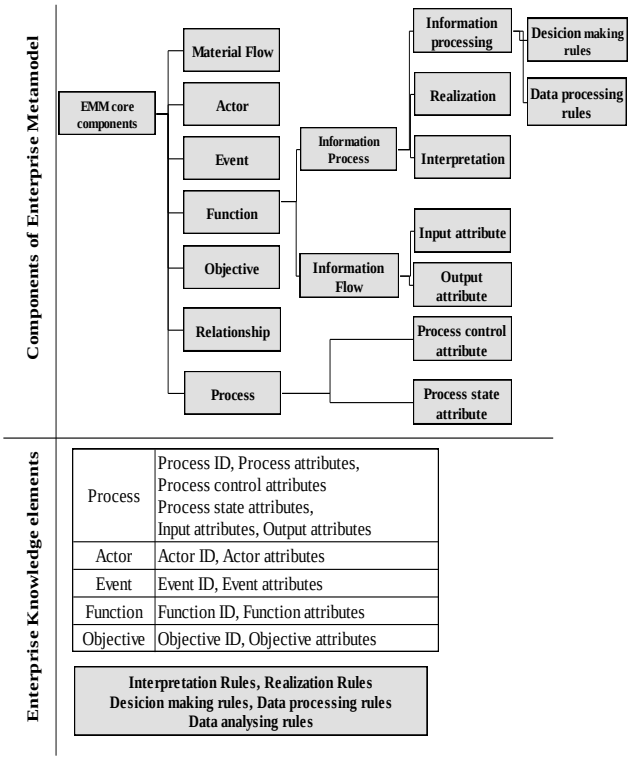


Figure 7. Components of Enterprise Meta-model (EMM) (Gudas S. et al., 2005)

E. Comparison of Enterprise Modeling Methods

Table 1 show a summary of the Enterprise modeling methods (Enterprise Meta-models) and it's, which demonstrates what knowledge elements are identified by each Enterprise modeling method. Each component expresses an aspect of Enterprise knowledge, which can be accumulated in the Knowledge base and developed on the basis of such an EMM. It indicates which Enterprise model accumulates more knowledge of the organization. A comparative analysis shows that Enterprise Meta-model (EMM) accumulates more knowledge components about the activities of the organization (Figure 7) as it models Enterprise management function in detail because of the control-view approach [4].

 ISBN: 978-1-61208-160-1

TABLE I. COMPARISON OF ENTERPRISE MODELING METHODS

Components of Enterprise Knowledge		ARIS	UEML	BPMN	EMM
Process	Process ID	+	+	+	+
	Process State Attributes				+
	Process Control Attributes				+
Function	Function ID	+			+
	Process State Attributes A				+
	Process Control Attributes V				+
	DP Input Attributes B				+
	DP Output AttributesC				+
	DM Output AttributesD				+
	IN Rules				+
	DP- Data Processing Rules				+
	DM -Decision Making Rules				+
	Decision Realization Rules				+
	Data Analysis Rules				+
Event		+	+	+	+
Role		+	+		+
Actor			+		+
Activity		+	+	+	+
Goal, Objective		+		+	+
Control		+			+
Time			+	+	
Material Flow				+	+
Input		+	+	+	+
Output		+	+	+	+
Object			+		
Interpretation Rules					+
Realization					+
Decision control					+
Data structuring					+
Business Rules		+	+	+	

Table 1 lists the components of different Enterprise modeling approaches, indicates Enterprise knowledge components and attributes. After the analysis of the models and their constituent components we may presume that the EMM model retains more Enterprise knowledge components that are needed in order to enhance the IT based organization's management functions.

IV. ENTERPRISE KNOWLEDGE MODEL'S INTEGRATION INTO DATA MINING

Enterprise knowledge model is a model of information management processes, identifying Enterprise components and their interactions, components of Enterprise management (data, knowledge, goals) and their interactions [6]. Application of this model for the development of a Knowledge-based information system is important because it reflects the essential elements of the model and their interactions and displays information flows (Figure 5). The element of knowledge management is important in this model. Adding certain rules, knowledge of experience to functions performed by process, more accurate and structurally interesting information can be extracted from data sets.

Figure 8 represents a diagram of entities' connections of Enterprise management model, which consists of the following components: Process, Function, Goals, Interpretation rules, Realization rules, Decision making rules, Data processing rules and Interactions between them. This is a transformed diagram of Enterprise Meta-models defined in Figures 2 and 3, the aim of which is to show the interfaces of Enterprise knowledge with process and function.

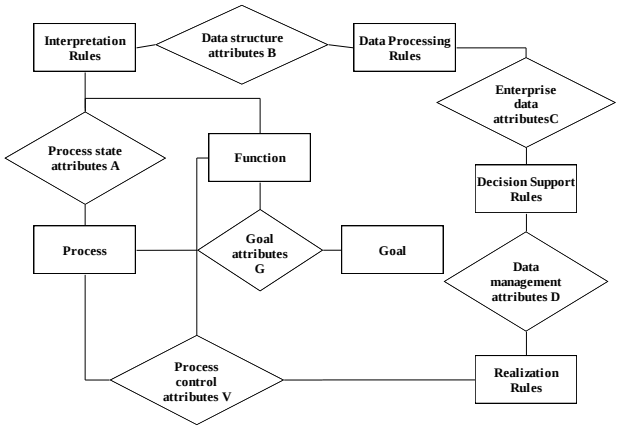


Figure 8. Enterprise Knowledge-based model - Entity relationships diagram (Gudas S., 2004)

In particular, Enterprise model helps to define and optimize processes in a company. Sequence of events, their interactivity, responsibility and data sources are clearly defined in the model. The composed Enterprise model allows the user to clearly imagine the data mining process and to adapt it to their own company easily and efficiently.

Components of knowledge that can be expressed in a form of Enterprise rules are accumulated in Knowledge base. In this base, rules are stored independently from software code, thus, ensuring the flexibility of a system and the opportunity to quickly adapt the functionality of the system under the changed conditions, but without changing the software code. Template of Enterprise rules can be expressed as follows:

IF <Variable><Condition><Value> THEN
<Variable><Condition><Value>

Figure 9 depicts mapping of the structure of Enterprise Meta-model into data mining process. Data mining process consists of several stages: creation of an OLAP cube, which will be used for data analysis; making a cube scheme; creation of a new data mining model; creation of additional fields, columns for implementation of a data mining algorithm; implementation of data mining process [11]. The table “Enterprise Meta-model” (Figure 9) provides knowledge components of Enterprise Meta-model that can

be transmitted to a certain stage of data mining process and monitors its implementation. Knowledge bases and interaction between database and data mining process are also highlighted.

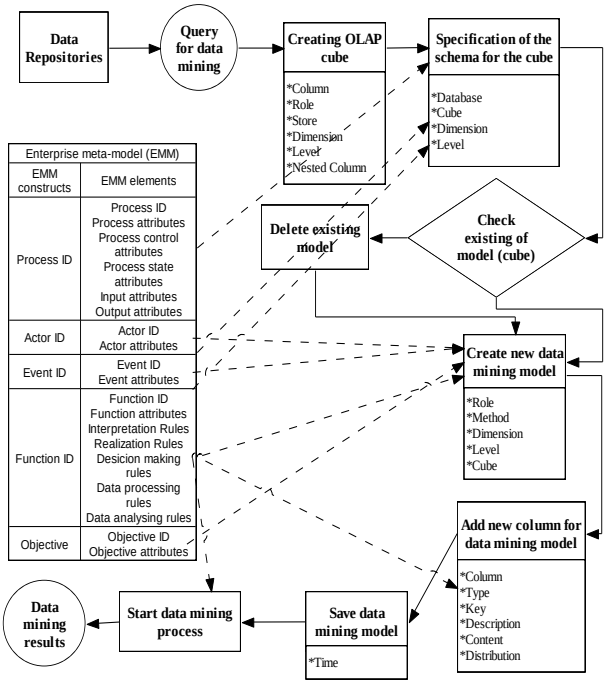


Figure 9. Mapping of EMM components into data mining process steps

A few types of Business rules are identified by EMM (IN rules, DP rules, etc.) usually have a declaratory form, note certain events and express them in conditional sentences. Figure 9 presents the model, in which the components of Enterprise management knowledge can be used for a specific data mining process phase. They may be useful in choosing the data mining algorithm, through which the data will be analyzed. According to the formed rules that would define the rules for algorithm model selection, the stage of model selection would be controlled.

Such Enterprise business rules, accumulated in Knowledge base, can be used in the process of data mining: Data processing, Realization, Decision making, and Interpretation rules.

Interpretation rules may be useful in shaping dimensions and determining the variables needed for data mining model, by determining roles and selecting methods for data analysis. Data processing rules will allow controlling data selection for data mining. Realization rules can be used in choosing the format of displaying results gained during the data mining process.

V. CONCLUSIONS

In this paper the control view-based approach to integration of data mining and Enterprise knowledge modeling was discussed.

The analysis of Enterprise modeling methods and languages (ARIS, UEML, BPMN and detailed Value Chain Model) was performed for elicitation of Enterprise knowledge components.

The control view-based Enterprise Meta-model EMM (Figure 7) is selected as a rational alternative for Enterprise management knowledge modeling. A comparative analysis of these Enterprise modeling methods in terms of knowledge components has shown that EMM is distinguished from other Enterprise modeling methods by a detailed structure of identified knowledge.

Our analysis has shown that it is possible to integrate a Knowledge base, which is designed for EMM, into data mining process to streamline the implementation of a data mining algorithm (Figure 9). Such a data mining system is enhanced by Enterprise knowledge subsystem, and can give additional semantics and functionality for experts to mine more accurate and relevant Enterprise data and knowledge components.

REFERENCES

[1] Josef W. Seifert. Data Mining: An Overview, CRS Report for Congress, Received through the CRS Web. Order Code RL31798, 2004

[2] Saulius Gudas. Žiniomis grindžiamos IS inžinerijos metodų principai. International conference, IT-2005, KTU, Kaunas, 2005

[3] Saulius Gudas, Tomas Skersys and Audrius Lopata. Framework for Knowledge-Based IS Engineering. ADVIS 2004: 512-522

[4] Saulius Gudas, Tomas Skersys and Audrius Lopata. Approach to Enterprise Modeling for Information Systems Engineering. Informatica, Vol. 16, No. 2, Institute of Mathematics and Informatics, Vilnius, 2005, pp. 175-192

[5] The CRISP-DM consortium (2011-06-25) Process Model. The CRISP-DM project. <http://www.crisp-dm.org/Process/index.htm>

[6] Saulius Gudas, Tomas Skersys and Audrius Lopata. (2004). Framework for knowledge-based IS engineering. Advances in Information Systems: 3rd international conference, ADVIS 2004, Izmir, Turkey, October 20-22, 2004: Proceedings. Berlin, Springer, pp. 512-522. ISBN 3-540-23478-0

[7] Saulius Gudas and Rasa Brundzaitė. Decomposition of the Enterprise Knowledge Management Layer, 2006 Seventh International Baltic conference on databases and information systems-proceedings, pp. 41-47, 2006 ISBN 1-4244-0345-6

[8] François Vernadat. UEML: Towards a Unified Enterprise modeling language. International Conference on Industrial Systems Design, Analysis and Management (MOSIM'01), Troyes, France, 2001

[9] IDS Scheer. ARIS Platform Products (2011-06-25) http://www.bptrends.com/publicationfiles/04-08-PR-BPM-Tools Report-IDS Scheer.pdf

[10] OMG Modeling and Metadata Specifications (2011-06-28) http://www.omg.org/technology/documents/modeling_spec_catalog.htm

[11] Daniel T. Larose. Data mining methods and models, IEEE Computer Society Press, pp. 344, ISBN-13 978-0-471-66656-1., 2006

 ISBN: 978-1-61208-160-1