



Ontology-Based Knowledge Model for Multi-View KDD Process

El Moukhttar Zemmouri, Hicham Behja, Abdelaziz Marzak, Brigitte Trousse

► **To cite this version:**

El Moukhttar Zemmouri, Hicham Behja, Abdelaziz Marzak, Brigitte Trousse. Ontology-Based Knowledge Model for Multi-View KDD Process. International Journal of Mobile Computing and Multimedia Communications (IJMCMC), IGI Global, 2012, 4 (3), pp.21-33. .

HAL Id: hal-00769786

<https://hal.inria.fr/hal-00769786>

Submitted on 3 Jan 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Ontology-Based Knowledge Model for Multi-view KDD Process

EL Moukhtar Zemmouri

Ecole Nationale Supérieure d'Arts et Métiers, Meknes, Morocco

Hicham Behja

Ecole Nationale Supérieure d'Arts et Métiers, Meknes, Morocco

INRIA Sophia Antipolis - Méditerranée, Equipe-projet AxIS, Sophia Antipolis, France

Abdelaziz Marzak

Faculté des Sciences Ben M'sik, Casablanca, Morocco

Brigitte Trousse

INRIA Sophia Antipolis - Méditerranée, Equipe-projet AxIS, Sophia Antipolis, France

ABSTRACT

Knowledge Discovery in Databases (KDD) is a highly complex, iterative and interactive process that involves several types of knowledge and expertise. In this paper we propose to support users of a multi-view analysis (a KDD process held by several experts who analyze the same data with different viewpoints). Our objective is to enhance both the reusability of the process and coordination between users. To do so, we propose a formalization of viewpoint in KDD and a Knowledge Model that structures domain knowledge involved in a multi-view analysis. Our formalization, using OWL ontologies, of viewpoint notion is based on CRISP-DM standard through the identification of a set of generic criteria that characterize a viewpoint in KDD.

Keywords: KDD process, viewpoint, domain knowledge, ontology, CRISP-DM, OWL.

INTRODUCTION

Knowledge Discovery in Databases (KDD) is a highly complex, iterative and interactive process, with a goal-driven and domain dependent nature (Fayyad et al., 1996). It involves three main steps (data preprocessing, data mining and post-processing) with many decisions made by the analyst (see Figure 1). The complexity of KDD is mainly due to the nature of the analyzed data (distributed, incomplete, heterogeneous, etc.) and the nature of the process itself (since the KDD is by definition interactive and iterative).

Given this complexity of KDD, the analyst faces two major challenges. On the one hand, he must manipulate prior domain knowledge to better understand the data and the business objective. On the other hand, he must be able to choose, configure, compose and execute tools and methods from various fields (e.g., machine learning, statistics, artificial intelligence, databases) to achieve goals. The first challenge involves analyzed domain knowledge, while the second involves the analyst domain knowledge (see Figure 1).

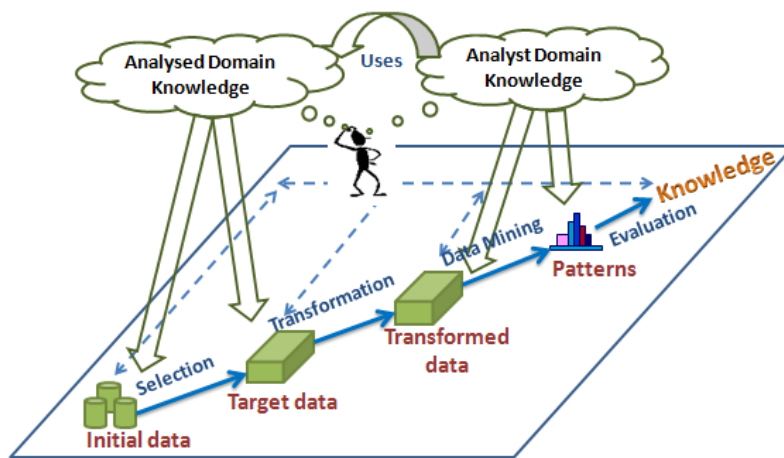


Figure 1. Interaction between KDD and the two types of domain knowledge. Analyzed domain knowledge is used during the early stages of the process; mainly to understand and prepare data. Analyst domain knowledge is used during the latter stages to choose, configure and execute data mining methods, and to evaluate extracted patterns (Behja et al., 2005).

A multi-view KDD process is usually held by one or more experts who consequently manipulate several types of knowledge and know-how. They will have different objectives and preferences, different competences, and different visions of analyzed data, KDD methods and functions. In brief, they have different viewpoints. In this context, the KDD process will be guided by the analyst's viewpoint (Behja et al., 2005) and several types of knowledge and expertise are incorporated.

Figure 2 below shows an example of a multi-view analysis of data from an e-learning system (mainly: log files, database, and courses material). These data can be analyzed by different actors of the system (learners, teachers, administrator, marketing ...). The objective of a teacher (e.g., evaluation of a course) is not the same as the administrator's one (e.g., ensuring system reliability). Attributes used for evaluating a course are different from those used for studying the reliability. Similarly, chosen data mining methods, techniques and tools will be different, and the interpretation of data mining results depends on the analyst's viewpoint. Therefore, it is fundamental to take into account the viewpoint of each analyst and incorporate the two types of domain knowledge in the KDD process.

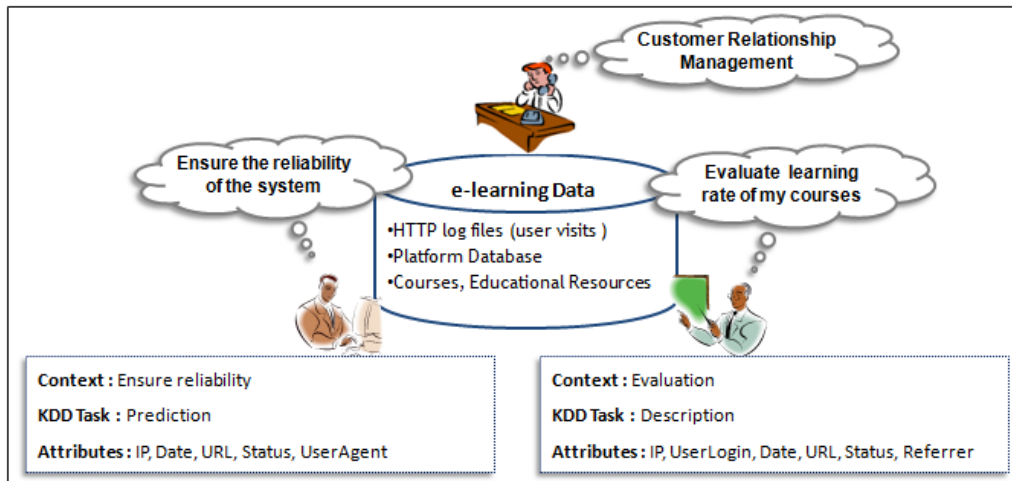


Figure 2. Multitude of viewpoints to analyze data from an e-learning system. The teacher may have as an objective the “evaluation of learning rate of a course”, description as KDD task, and may use (IP, UserLogin, Date, URL, Status, and Referrer) as attributes. While the administrator may have as an objective “ensuring the reliability of the system”, prediction as KDD task, and may use (IP, Date, URL, Status, and UserAgent) as attributes.

In this paper we propose to assist the users of a multi-view KDD process. Our objective is to enhance both the reusability of the process and coordination between its different users. We propose a formalization of viewpoint notion in KDD following a knowledge engineering approach: eliciting, structuring, and formalizing information and knowledge involved in a multi-view analysis (Schreiber et al., 2000). Elicitation will be based on CRISP-DM standard (Chapman et al., 1999) to identify a set of generic criteria that characterize a viewpoint in KDD. Knowledge involved in a multi-view analysis will be structured as a knowledge model containing four hierarchical sub-models: domain model, task and method model, viewpoint model and viewpoint organizational model. The viewpoint sub-model will be formalized using Ontologies in OWL¹ (Web Ontology Language) language.

The remainder of the paper is organized as follows. Section 2 presents the state-of-the-art in supporting users of KDD process using ontologies. Section 3 gives our definition of viewpoint in KDD, its situation in relation to other approaches in knowledge engineering, and our proposed set of generic criteria that characterize a viewpoint in KDD. Then, in section 4 we present our knowledge model integrating viewpoints. In this section we also present our OWL formalization of the viewpoint sub-model. Finally section 5 draws some conclusions and opens up new avenues for future work.

RELATED WORK

Several related works have addressed the complexity of KDD with different approaches with the aim of supporting both expert and novice users using ontologies. Most of these approaches offer the user the advantage to explore the large space of valid data mining processes (Bernstein et al., 2005, Diamantini et al., 2009), to discover and access distributed data mining services (Euler 2005, Diamantini et al., 2009), to reuse successful data mining experiences (Morik et al., 2004), etc. but without taking into account the multi-view aspect of a KDD analysis.

One of the first ontologies proposed to support users of KDD is DAMON (Data Mining ONtology) (Cannatro and Comito, 2003), that is designed to simplify the development of distributed KDD applications on Grids. DAMON ontology concerns only the data mining

phase of a KDD process, and offers a taxonomy for discovering tasks, methods and tools deemed more suitable for a given data mining goal.

In MiningMart project (Morik et al., 2004) a case-based reasoning (CBR) system to support end users during data preprocessing is proposed. This system is based on a meta-model (called M4) of KDD preprocessing chains that contains ontology for describing conceptual domain knowledge. In the same project, Euler (2005) proposes a web-based platform (which is a case base containing MiningMart successful experiences) to publicly display preprocessing models in a structured way, together with descriptions about their business domains, goals, methods and results.

Bernstein et al., (2005) propose an Intelligent Discovery Assistant (IDA) for valid data mining processes enumeration and ranking. IDA focuses mainly on preprocessing and data mining phases of the KDD process. It is based on a formal ontology that contains input/output, preconditions constraint, and performance (accuracy, complexity, and comprehensibility) of each data mining operator. This ontology allows selection and composition of data mining operators suitable for user's data and goal. A similar approach is proposed by Diamantini et al., (2009) in a project called KDDVM (KDD Virtual Mart), which is a web services based system that aims to support users in the design of valid KDD process. It represents KDD operations as services which can be "annotated, introduced, accessed, described, composed and activated". KDDVM is based on KDDONTO ontology and concerns only data preprocessing and data mining steps.

A recent European project (e-LICO² : e-Laboratory for Interdisciplinary Collaborative Research in Data Mining and Data-Intensive Science) deals with the problem of supporting users of KDD in a collaborative way (Hilario et al., 2011). One of the products of this project is eProPlan (Kietz et al., 2010), an ontology based environment for planning KDD workflows. It is based on two ontologies DMWF-DMOP and uses IA planning techniques to automatically generate KDD execution plan for solving data mining problems. DMWF (Data Mining Work Flow Ontology) formalizes IO-objects, operators, goals, tasks and methods as well as the decomposition of tasks into methods and operators (this ontology is equivalent to our OntoECD ontology described in (Zemmouri et al., 2009)). DMOP (Data Mining Optimization Ontology) provides a unified conceptual framework for analyzing data mining tasks, algorithms, models, datasets, workflows and performance metrics, as well as their relationships.

Our approach focuses on the reusability and coordination between multi-users of a KDD process, rather than the automatic generation of KDD execution plan. In addition, we cover the whole phases of the KDD process (cf. 6 phases suggested by CRISP-DM standard in Figure 4).

A VIEWPOINT-BASED APPROACH IN KDD

The concept of viewpoint is a polysemic word introduced in knowledge representation since the 70s by Minsky (1975), especially for modeling and design of complex systems that are inherently multi-view. Since then, several proposals have focused on the meaning, representation, interpretation and confrontation of viewpoints, sometimes with different names (e.g., perspective, context, opinion, view, etc.).

The multi-view approach has been used in various fields: in object-oriented methods and languages VBOOM (View Based Object Oriented Method) (Kriouile, 1995) and VBOOL (View-Based Object Oriented Language) (Marcaillou, 1995), in view programming (Mili et al., 2000), in the O2VIEWS object databases (Abiteboul & Bonner, 1991), in software process (Finkelstein et al., 1990) and also in UML (Nassar, 2004).

The concept of viewpoint was also addressed implicitly or explicitly by a large community of knowledge engineering. Especially in object knowledge representation languages like KRL (Bobrow & Winograd, 1977) and its improvement LOOPS (Bobrow & Stefik, 1982), ROME (Carré et al., 1990) and its extension to Frames FROME (Dekker & Carré, 1992), and TROPES (Marino, 1993); in the context of multi-expertise modeling in designing complex systems such as spacecrafts (Trousse, 1998) and in developing multi-expert knowledge bases with C-VISTA (Rivière & Dieng, 2002), and MVP (Bach, 2006).

According to the works and areas mentioned above, the definition of viewpoint notion varies. For example Rivière and Dieng (2002) have classified definitions of the concept viewpoint into two categories which correspond to two possible interpretations: *perspective* and *opinion*. *Perspective viewpoints* refer to different conceptual positions from which different experts examine an object (system, knowledge base, problem...). This allows one “to index consensual descriptions of the same object by different experts” (Rivière & Dieng, 2002), and also to access a subset of relevant information or knowledge from a given viewpoint. *Opinion viewpoints* refer to opinions given by experts on the object. Opinion viewpoints are closely related to the expert and take account of his experience, knowledge, task, etc.

Viewpoint in KDD

We propose, in this paper, to make explicit the notion of viewpoint in KDD. In fact KDD is a complex process (complex system) most often held by several experts (i.e. multi-view). The definition we propose for the notion of viewpoint in KDD is inspired from knowledge engineering (knowledge involved during a KDD analysis) and takes into account the two facets *perspective* and *opinion* defined by Rivière and Dieng (2002).

Definition: *A viewpoint in KDD is an interface allowing (1) access to a subset of domain knowledge (analyzed domain knowledge and analyst domain knowledge) and leading the analyst to achieve his goals, (2) capture the logic of reasoning and trace of major decisions made by the analyst during a KDD process (i.e. capture the semantics of the process).*

Thus the viewpoint of the analyst allows him to filter the relevant expert domain knowledge according to his vision on analyzed data, on application domain, on the domain of KDD (tasks, methods, algorithms, tools...), and according to his objective.

The specification of an analyst viewpoint is based on the instantiation of a set of generic criteria that may fall into three components: *analyzed domain*, *analyst domain*, and *context of analysis* (see Figure 3).

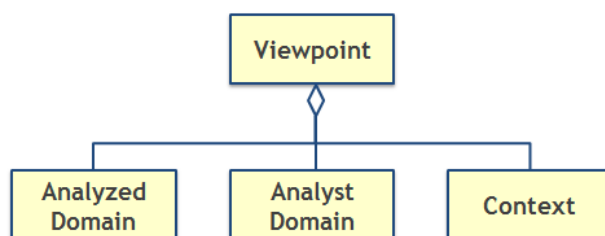


Figure 3. Components of a viewpoint in KDD.

To identify these three components of viewpoint, we were inspired by viewpoints in second-order cybernetics: observed object, observer and context of observation (Ben Ahmed, 2005). This is a systemic approach of a multi-view KDD process that we analyze from three

viewpoints: analyzed domain viewpoint (observed object), analyst domain viewpoint (observer), and context of analysis.

To identify generic criteria that characterize a viewpoint in KDD, we were based on the CRISP-DM standard CRISP-DM. This point is discussed in the next sub-section.

Generic criteria of viewpoint in KDD

Our objective in this paper is to formalize the definition of viewpoint in KDD. For this, we will first identify a set of generic criteria that characterize a viewpoint in KDD. Once instantiated, these criteria define an analyst viewpoint. This viewpoint will guide the execution of the KDD process, and then keep trace of reasoning and major decisions made by the analyst.

The criteria of a viewpoint are called *generic* “if they are not instantiated” (Rivière & Dieng, 2002) (e.g. *KDD_Task* is a generic criterion), they are called *specific* criteria if they are instantiated (e.g. *KDD_Task* = “*Description*” is a specific criterion). The criteria are generic also if they are independent of the application domain, and of the data mining tools and techniques. For this reason, we have found useful to elicitate these generic criteria based on CRISP-DM standard. Genericity of criteria will be ensured by the level of abstraction and description of KDD process according to CRISP-DM.

In fact, CRISP-DM (Cross-Industry Standard Process for Data Mining) (Chapman et al., 1999) is a *process model* and a methodology that describes commonly used approaches to conduct a data mining project. CRISP-DM process model focuses on the life cycle of a KDD project; it does not rely on a particular application domain, data mining technique or tool. CRISP-DM methodology is described in terms of a hierarchical set of tasks organized at four levels of abstraction: phase, generic task, specialized task, and process instance (from generic to specific). At the top level, the life cycle of a KDD project is organized into six phases as depicted in Figure 4.

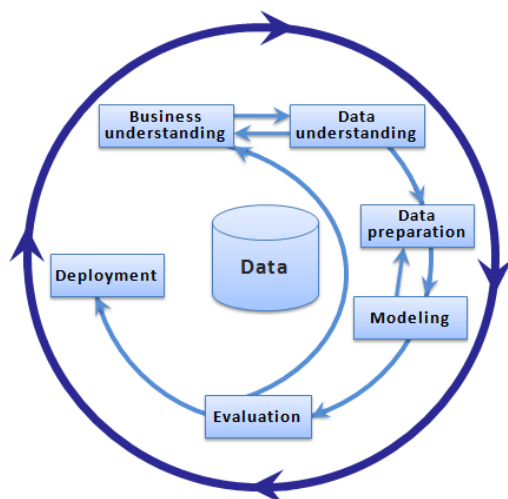


Figure 4. CRISP-DM reference model (Chapman et al., 1999).

In each phase of the process, the analyst must perform a number of tasks. According to the knowledge required for each task, we can distinguish: tasks that require analyzed domain knowledge, tasks that rely on analyst domain knowledge, his skills and expertise, and tasks that rely on business objectives and criteria for validating the results of the process (models).

Based on these tasks, we have identified the set of generic criteria that we draw in Table 1, and formalize in the next section.

Table 1. Selected generic criteria of a viewpoint in KDD based on CRISP-DM.

	Analyzed domain viewpoint	Analyst domain viewpoint	Context viewpoint
Business understanding	Business area Project resources +Data +Knowledge +personnel +computing Requirements, Assumptions, Constraints Risks – Contingencies Terminology	KDD Task KDD success criteria +Model accuracy +Model performance +Model complexity Initial project plan Assessment of tools and techniques	Business objective Business success criteria Costs and Benefits
Data understanding	Data locations Datasets (list) Sources Problems Data description + Data format + Data quantity - number of fields - number of records - number of classes + Attributes list (names) + attribute data types Data quality + has missing values + has inconsistent values + has outlier values	Collect methods Selection criteria Data exploration +Basic statistical	
Data preparation	Dataset Data selection method Data selection criteria Derived Attributes Merged data Reformatted data	Data cleaning method +missing handling +outliers handling +duplicates handling +errors handling Data Transformation +Method +Parameters	
Modeling	Modeling assumptions +Data format +Data quality Test Design +Options +Training Dataset +Test Dataset +Validation Dataset	Modeling Technique Parameter settings +Parameter, +Value Model (Final Model Location) Model description +Estimated Model Accuracy Model Assessments +Model Quality -DM success criteria achieved	
Evaluation		Approved Model Improvement suggestions List of possible actions Decision	Business success criteria achieved
Deployment		Deployment plan Monitoring plan Maintenance plan Final report Final presentation Experience documentation	

KNOWLEDGE MODEL FOR MULTI-VIEW KDD PROCESS

According to CommonKADS project (Schreiber et al., 2000) a knowledge model is “a specification of the information and knowledge structures and functions involved in a knowledge-intensive task”. It gives an implementation-independent description of knowledge involved in a task. A knowledge model is composed of three types of knowledge that are subject of separate models with specific modeling primitives: domain knowledge, tasks and methods (Charlet et al., 2000).

Our knowledge model integrating the viewpoint notion (see Figure 5) consists of four hierarchical sub-models structured in domain knowledge and strategic knowledge according to Aussenac-Gilles et al. (1996). Domain level describes the domain concepts and their relationships. The strategic level is based on the domain level and expresses how a task will be achieved.

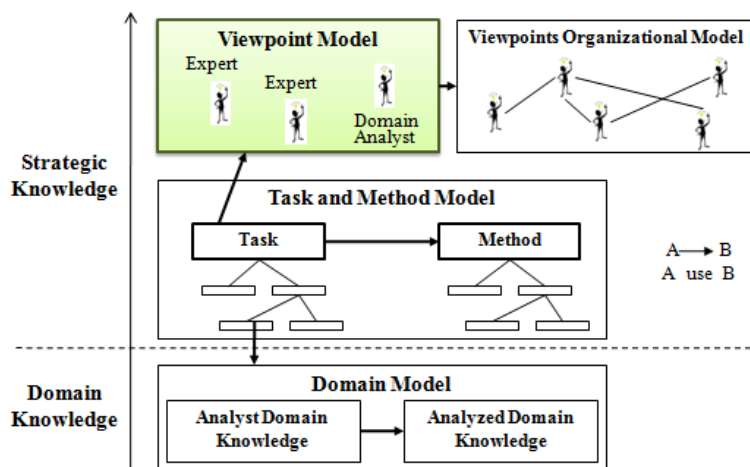


Figure 5. Knowledge Model for a multi-view KDD process.

In this paper, we will focus much more on the description and formalization of the viewpoint model in figure 5.

Domain Model

Domain knowledge is knowledge about application domain that is necessary to execute methods. The domain model provides conceptualization of studied domain concepts and also the various relations between them. It describes the analyzed domain knowledge in terms of manipulated data and the analyst domain knowledge in terms of tasks performed, methods selection, configuration and composition. In our context of KDD, we consider domain model as an ontology used to index manipulated data and their attributes.

Task and Method Model

This model describes the KDD process in terms of tasks and methods. Tasks are performed by methods. A task is a description of what must be done in the application in terms of goals and sub-goals. The methods describe how a goal can be achieved in terms of a series of operations and an order of execution.

We have formalized this model as a generic semi-formal ontology OntoECD (Zemmouri et al., 2009). OntoECD conceptualizes methods and functions of KDD process regardless of the application domain and the structure of data to improve reusability.

Viewpoint Model

Viewpoint model is a conceptualization of the generic criteria introduced in section 2. These criteria are independent of the task and the application domain. They allow modeling the vision of the analyst on manipulated data, the objective of analysis, and part of the expertise required for decisions made during the analysis.

Modeling viewpoint in KDD will promote coordination and understanding between different experts in a multi-view analysis. It also promotes the reuse of analysis according to a given viewpoint (this is possible thanks to annotations of KDD process that we introduced in (Behja et al., 2010)).

To formalize the viewpoint model, we have chosen OWL (Web Ontology Language) due to its expressiveness compared with XML Schema and RDF(S), and to its representational and inferring capacity. Figures 6, 7 and 8 draw a subset of classes' hierarchy and relations of viewpoint in KDD.

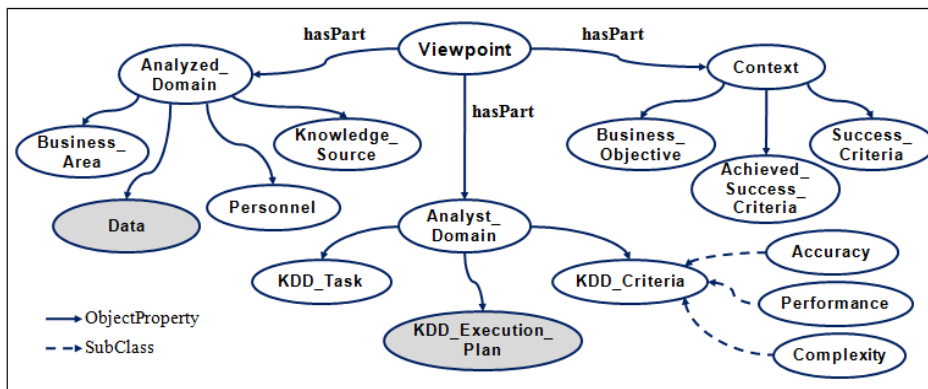


Figure 6. Class hierarchy of viewpoint in KDD.

A viewpoint in KDD is composed of three parts (see Figure 6): the *analyst domain viewpoint* which describes actions and decisions taken by the analyst in terms of KDD task (verification, description, or prediction), selected KDD methods, and criteria for validating methods. The *analyzed domain viewpoint* describes the analyst vision on analyzed domain in terms of selected data and relevant attributes for the analysis. The *context viewpoint* describes business objectives of the project and criteria for validating the results (models).

The main part of the analyst domain viewpoint is actions and decisions made to construct a KDD execution plan (see Figure 7). A KDD execution plan consists of a set of selected methods (data preparation and modeling according to CRISP-DM). Each method has some constraints (pre- and post-conditions) and a set of parameters. Modeling methods have models as output. The analyst has to select, configure and execute methods according to the KDD task, and available data.

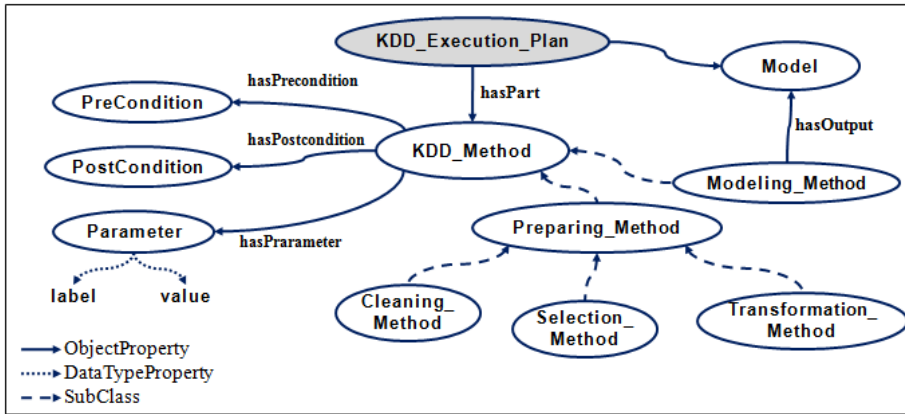


Figure 7. Class hierarchy for analyst domain viewpoint.

The main part of analyzed domain viewpoint (see Figure 8) is the description of analyzed data and the effects of executed methods on data. Analyzed data have some properties: format, source, quality (missing values, errors...), and quantity (number of attributes, number of rows...). Each selected and executed method by the analyst has a transformational effect on data and their properties.

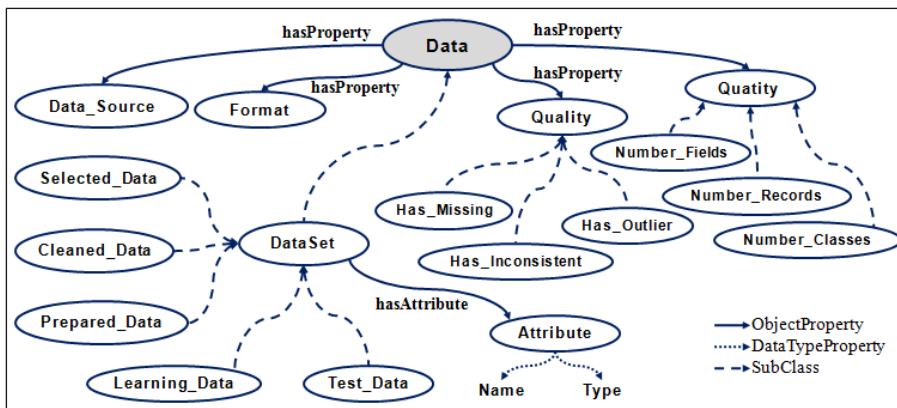


Figure 8. Class hierarchy for analyzed domain viewpoint.

The instantiation of these classes characterizing a viewpoint in KDD is considered as an analyst viewpoint's definition. This viewpoint will guide the execution of the process and allow keeping trace of decisions made by the analyst.

Viewpoints Organizational Model

In a multi-view analysis, it is important to emphasize the interaction and interdependence between various analyses according to different viewpoints. The viewpoints organizational model (see Figure 5) requires the identification of a set of relations between viewpoints like: equivalence, exclusion, complementarity, composition, subsumption. These relations are under study and development.

CONCLUSION

In this paper we have presented a formalization of the concept of viewpoint in KDD that integrates both analyzed domain knowledge and analyst domain knowledge. This multi-view approach of KDD will promote coordination and understanding between different experts in a

multi-view analysis, and also the reuse of analysis according to a given viewpoint. For example there will be some complementarities between the viewpoint “evaluation” of a teacher, and the viewpoint “reliability” of the administrator (see Figure 2). Also a teacher will reuse the analysis of another teacher (i.e. one can use the same KDD execution plan of another to evaluate his/her course materials).

Figure 9 below shows the current version of a platform that we are developing to support multi-view analysis. It is able to interact with a user (a teacher in this example) to define his viewpoint by instantiating some of the generic criteria defined in this paper (context, KDD_Task, Selected_Data, relevant attributes) and then guide him in the selection, configuration and execution of KDD methods based on the collection of algorithms Weka³. Our system is supported by the Jena⁴ Toolkit and its SPARQL query engine. It is architected as a Web Services Architecture (Zemmouri et al., 2010).

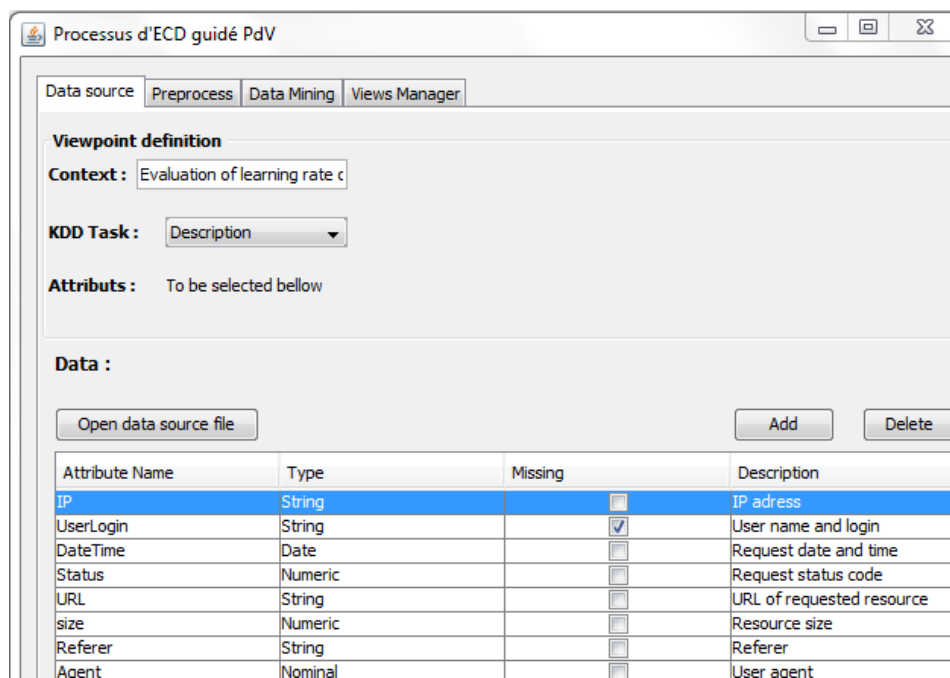


Figure 9. Definition of a teacher viewpoint to analyze data (HTTP log files) of an e-learning platform.

As future works, we plan to develop the organizational model for defining relationships between viewpoints, and allowing the reasoning on viewpoints. We also plan to develop the aspect of assessment of methods and models that are not treated in depth in this paper.

REFERENCES

- Abiteboul, S., & Bonner, A. (1991). Objects and views. In *Proceedings of SIGMOD International Conference on Management of Data* (pp. 238-247). ACM Press. doi:10.1145/119995.115830.
- Aussenac-Gilles, N., Laublet, P., & Reynaud, C. (1996). *Acquisition et ingénierie des connaissances : tendances actuelles*. Toulouse, France : Cépaduès.
- Bach, T. L. (2006). *Construction d'un web sémantique multi-points de vue*. Unpublished doctoral dissertation, École des Mines de Paris à Sophia Antipolis, France.
- Behja, H., Trousse, B., & Marzak, A. (2005). Prise en compte des points de vue pour l'annotation d'un processus d'Extraction de Connaissances à partir de Données. In S. Pinson & N. Vincent (Eds.), *Revue des Nouvelles Technologies de l'Information (RNTI-E-3) : Vol. 1*. (pp. 245-256). Cépaduès-Éditions.
- Behja, H., Zemmouri, E., & Marzak, A. (2010). Viewpoint-based annotations for knowledge discovery in databases. In *Proceedings of IEEE International Conference on Machine and Web Intelligence ICMWI* (pp. 299-302), Algiers, Algeria. doi:10.1109/ICMWI.2010.5647881
- Ben Ahmed, W. (2005). *Safe Next : une approche systémique pour l'extraction de connaissances de données*. Unpublished doctoral dissertation, Ecole Centrale de Paris, France.
- Bernstein, A., Provost, F., & Hill, S. (2005). Towards intelligent assistance for a data mining process: an ontology based approach for cost-sensitive classification. *IEEE Transactions on Knowledge and Data Engineering*, 17(4), 503-518. doi:10.1109/TKDE.2005.67.
- Bobrow, D. G., & Stefik, M. J. (1982). LOOPS: data and object oriented Programming for Interlisp. *European AI Conference*, Orsay, France.
- Bobrow, D. G., & Winograd, T. (1977). An overview of KRL, a knowledge representation language. *Cognitive Science*, 1(1).
- Cannatro, M., and Comito, C., (2003). A data mining ontology for grid programming. In *Proceedings of 1st International Workshop on Semantics in Peer-to-Peer and Grid Computing, in conjunction with WWW2003* (pp. 113-134).
- Carré, B., Dekker, L., & Geib, J. M. (1990). Multiple and evolutive représentation in the ROME language. In *Proceedings of TOOLS'90* (pp. 101-109).
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (1999). *CRISP-DM 1.0 Step-by-step data mining guide*. (Tech. Rep.) CRISM-DM consortium. Retrieved from, <http://www.crispdm.org>
- Dekker, L., & Carré, B. (1992). Multiple and dynamic representation of frames with points of view in FROME. In *Proceedings of Représentation Par Objets* (pp. 97-111). La Grande Motte.

Diamantini, C., Potena, D., & Storti, E. (2009). Ontology-driven KDD process composition. In N. M. Adams, C. Robardet, A. Siebes, & J.-F. Boulicaut (Eds.), *Lecture Notes in Computer Science: Vol. 5772, Advances in Intelligent Data Analysis VIII* (pp. 285-296). Springer Verlag. doi:10.1007/978-3-642-03915-7.

Euler, T. (2005). Publishing operational models of data mining case studies. In *Proceedings of Workshop on Data Mining Case Studies at the 5th IEEE ICDM* (pp. 99-106), Houston, Texas, USA.

Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, 39(11), 27-34.

Finkelstein, A., Kramer, J., & Goedicke, M. (1990). Viewpoint oriented software development. In *Proceedings of International Workshop on Software Engineering and its Applications*, Toulouse, France.

Hilario, M., Nguyen, P., Do, H., Woznica, A., & Kalousis, A., (2011). Ontology-based meta-mining of knowledge discovery workflows. In N. Jankowski, W. Duchs, & K. Grabczewski (Eds.), *Studies in Computational Intelligence, Vol. 358/2011, Meta-Learning in Computational Intelligence*, (pp. 273-315), Springer Verlag. doi: 10.1007/978-3-642-20980-2_9.

Kietz, J-U., Serban, F., & Bernstein, A. (2010). eProPlan: A tool to model automatic generation of data mining workflows. In P. Brazdil, A. Bernstein, & J-U. Kietz (Eds.), *Proceedings of the 3rd Planning to Learn Workshop (WS9) at ECAI 2010*.

Kriouile, A. (1995). *VBOOM : une méthode orientée objet d'analyse et de conception par points de vue*. Unpublished doctoral dissertation, University Mohammed V, Rabat, Morocco.

Marcaillou-Ebersold, S. (1995). *Intégration de la notion de points de vue dans la modélisation par objets : le langage VBOOL*. Unpublished doctoral dissertation, University of Toulouse, France.

Marino, O. (1993). *Raisonnement classificatoire dans une représentation objets multi-points de vue*. Unpublished doctoral dissertation, University of Grenoble, France.

Mili, H., Dargham, J., & Mili, A. (2000). Views: A framework for feature-based development and distribution of OO applications. In *Proceedings of the 33rd Hawaii International Conference on System Sciences Vol. 8* (pp.8049), Honolulu, HI. doi:10.1109/HICSS.2000.927011

Minsky, M. (1975). A framework for representing knowledge. In Patrick Henry Winston (Ed.), *The Psychology of Computer Vision*. McGraw-Hill, New York, U.S.A.

Morik, K., & Scholz, M. (2004). The MiningMart approach to knowledge discovery in databases. In N. Zhong & J. Liu (Eds.), *Intelligent Technologies for Information Analysis* (pp. 47-65). Springer Verlag.

Nassar, M. (2004). *VUML : une extension UML orientée point de vue*. Unpublished doctoral dissertation, ENSIAS, Rabat, Morocco.

Rivière, M., & Dieng-Kuntz, R. (2002). A Viewpoint model for cooperative Building of an Ontology. In U. Priss, D. Corbett, & G. Angelova(Eds.), *Lecture Notes in Computer Science: Vol. 2393, Conceptual Structures: Integration and Interfaces* (pp. 220-234). Springer Verlag. doi: 10.1007/3-540-45483-7_17.

Schreiber, A., Akkermans, J. M., Anjewierden, A., Hoog, R., Shadbolt, N. Van de Velde, W., & Wielinga, B. J. (2000). *Knowledge engineering and management: The CommonKADS methodology*. MIT Press, Cambridge, Massachusetts, London.

Trousse, B. (1998). Viewpoint management for cooperative design. In *Proceedings of the the IEEE Computational Engineering in Systems Applications (CESA'98)*, Nabeul-Hammamet, Tunisia.

Zemmouri, E., Behja, H., & Marzak, A. (2009). OntoECD : une ontologie pour le processus d'ECD. In *Proceedings of the WOTIC Workshop*, Agadir, Morocco, (pp 24-25).

Zemmouri, E., Behja, H., & Marzak, A. (2010). Architecture web service d'un processus d'ECD multiutilisateur. In *Proceedings of IEEE NGNS'10*, Marakech, Morocco (pp. 171-177).

EL Moukhtar Zemmouri received the Engineer degree in Computer Science and Telecommunication from "Institut National des Postes et Télécommunications" (INPT, Rabat, Morocco). He is currently an engineer at "Ecole Nationale Supérieure d'Arts et Métiers" (ENSAM, Meknes, Morocco), where he is preparing his PhD in Computer Science at the LM2I Laboratory. His research interests knowledge management in a multi-view KDD process (Knowledge Discovery in Databases) to support users with a special focus on multi-view analysis of data from e-learning systems. Prior to joining ENSAM in 2005, he worked at STMicroelectronics (French-Italian electronics and semiconductor manufacturer) and Maroc Telecom (telecommunication company in Morocco).

Hicham Behja received his Master degree in Computer Science from "Ecole Nationale Supérieure d'Informatique et Analyse des Systèmes" (ENSIAS) Rabat, Morocco in 1996, his first PhD in computer science from Faculty of Sciences University Mohammed V Agdal Rabat Morocco in 1999. He received his second PhD in Computer Science from Faculty of science Ben M'Sik University Hassan II Mohammedia in 2009. Since 1997, he has been an associate professor at "Ecole Nationale Supérieure d'Arts et Métiers" (ENSAM, Meknes, Morocco). His research interest includes Data mining, Web mining, semantic Web, knowledge management and e-learning.

Abdelaziz Marzak hold a PhD in Computer Science in 2000 (Definition of a measure for the viewpoint approach and the inheritance relationship: validation of the maintainability factor), and a PhD in Computer Science in 1997 (Design and implementation of the tool VBTOOL: Feature Analysis / Design), both from Mohammed V University at Rabat Morocco. He is currently a Professor of Computer Science at the Faculty of Science Ben M'Sik Casablanca, where he is responsible for PARS project, and DOME project (project proposal in collaboration with the CNRST of Morocco). He is the leader of the team "Engineering of Information Systems". He has several publications on Web and databases.

Brigitte Trousse is a research scientist at INRIA Sophia Antipolis and from 2003 the scientific leader of the AxIS Project-team there. She received her PhD in Artificial Intelligence and Computer aided Design from the University of Nice-Sophia Antipolis (1989). Before joining INRIA and her Ph-D, she held a permanent Professor position of Mathematics in a High School in Nice (1980-1984). Her research interests at Inria include Artificial Intelligence in complex design, Knowledge Engineering and Knowledge Discovery from Data, Semantic Usage Mining and more recently User Experience Measurement and ICT-SHS multi-disciplinary research for supporting living labs development and user-centered open innovation ICT based projects.

Endnotes

¹ OWL, W3C Consortium, <http://www.w3.org/TR/owl-ref/>

² e-LICO Project, <http://www.e-lico.eu/>

³ Weka - A Data Mining Software in Java, <http://www.cs.waikato.ac.nz/ml/weka/>

⁴ Jena – A Semantic Web Framework for Java, <http://jena.sourceforge.net/>