Toward an Ontology-based model of key performance indicators for business process improvement

Emna Ammar El Hadj Amor[1], Sonia Ayachi Ghannouchi[2]
[1]ISITCom Hammam Sousse, RIADI Laboratory-ENSI Manouba, Tunisia
[2]ISG Sousse, RIADI Laboratory-ENSI Manouba, Tunisia

***Abstract*— The selection and the use of performance measurement have received considerable interest in recent years. In addition, it is crucial not only to track the process behavior and to derive the key performance indicators but also to understand all necessary concepts involved in the BP and incorporate domain knowledge of the field. The improvement of business processes is based on comprehensive measurement, data understanding, task design, and relevant result interpretation of organization's performance. In this context, in order to make relevant decisions and to provide a consistent understanding of the field, we present a new ontology based on a real business process to create semantic relationships between all terms. After that, we were based on data mining technique to extract the most important information from data measurement. An example of implementation of our proposed contribution as well as its validation on a real case study in the healthcare domain is presented.**

Keywords— Business Process (BP); Business Process Management (BPM); Performance measurement; Key Performance Indicators (KPI); Ontology; Data mining; health care process

I. INTRODUCTION

The improvement of business processes is a critical requirement for any organization. Indeed many organizations rely more and more on the development of appropriate Key Performance Indicators (KPIs) for evaluating the process performance. KPIs provide critical information to the organization for monitoring and predicting business performance in accordance with strategic objectives [1]. According to Parmenter [4] key performance indicators represent a set of measures focusing on those aspects of organizational performance that are the most critical for the current and future success of the organization.

Depending on the context, several kinds of KPIs can be developed, including quantitative or qualitative aspects, and different data source can be used for retrieving and calculating their appropriate values. In this perspective, BPM helps organizations to model, manage, and optimize these processes for a significant gain. BPM has become an effective mean for creating abstract representations of knowledge, providing formalized definitions of the different activities, evaluating the process executions and/or evolutions [2]. Usually, the data related to the business process execution useful for monitoring facilities. Monitoring KPIs helps increase activities that add more value and eliminate activities with less or no value to the enterprise, which in the end will provide a permanent value for the organization as a whole [3].

To monitor KPIs, organizations rely on reports and dashboards [4] presenting one or more KPIs together in order to help decision makers identify opportunities for improvements or re-engineering the desired process.

However, this practice presents several drawbacks. First, it provides partial information to decision makers, without taking into account which activity in the BP is involved and which inter-relationships between KPI is considered. Then, they are required to interpret and link these raw data back to their performance measurement and how it can affect other business tasks and other performance measurements.

Some authors [5, 6, 7, and 8] have pointed out the vagueness, imprecision of KPIs values that they are intended to represent, and the lack of an explicit representation of their semantics. According to [8], the major obstacles for effective design and management of PI monitoring systems are related to the facts that PIs are complex objects with an aggregate/compound nature. This often leads to unawareness of indicator semantics as well as of dependencies among indicators. So, the authors in [8], propose to enrich the data cube model with the formal description of the structure of an indicator given in terms of its algebraic formula and aggregation function. Similarly, in [9], Ortega establishes that, in practice, PPIs are defined informally usually in ad-hoc natural language, with its well-known problems or they are defined from an implementation perspective, hardly understandable. In order to solve this problem, the authors propose an approach to improve the definition of PPIs using templates and linguistic patterns.

The problem above is how to perform BPs in a more effective and efficient way without ignoring qualitative aspects. Hence, selecting KPIs requires a high cooperation between its quantitative indicators and their related qualitative indicators. For example, to improve the overall performance of the healthcare business process, the evaluation focus on interesting quantitative KPIs related to the behavior of the process and also on interesting qualitative KPIs related to patients experience in the emergency department (ED). So, we need to identify relationships between BP elements and KPIs, and relationships amongst them to obtain relevant information that can assist BP analysts during BP improvement.

In our work, we focus not only on gathering the data required to extract and calculate KPI values, but also to infer knowledge in order to answer questions such as: which are the business process activities related to the appropriate KPI? Because for example, each accomplished task directly or indirectly has an effect on the process duration and including this information is required.

Another important question is the following: how does this affect the other indicators? Since all existing KPIs can affect each other, therefore none of these KPIs should be ignored for the sake of another one. And the last point, how to discover valuable information hidden in the KPIs data and how to transform these data into valuable and useful knowledge.

The remainder of the paper is organized as follows. In section 2 the proposed solution is presented. Section 3 presents and discusses experimental results obtained in a real case study. A conclusion ends this paper in the last Section.

II. PROPOSED SOLUTION

According to literature, we can say that we have to know more about the real meaning and the implications of performance measurement value and to understand why an indicator is needed and what it is measuring and what decisions it supports.

Hereafter, we refer to Business Process Management approach which includes methods, techniques, and tools to support the design, enactment, management, and analysis of operational business processes involving humans, organizations, applications, documents and other sources of information [10].

Figure 1 gives an overview of our proposed solution.

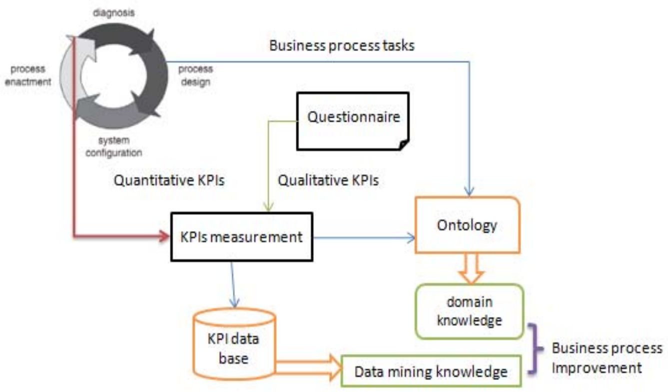

Fig. 1. Overview of our proposed solution

Starting from the last phase of BPM life cycle proposed by Van der Aalst [11], we consider both quantitative KPI (through execution-observation) and qualitative KPIs (through questionnaires). Process tasks and KPIs are related through an ontological model corresponding to their dependencies among each other (dependencies between KPIs and dependencies between KPIs and process tasks). Finally, analyzing the obtained values for both types of KPIs will allow validating or correcting.

Event data can be analyzed according to different perspectives [17]: (1) the control-flow perspective; (2) the organizational perspective; (3) the data perspective; and (4) the performance perspective. The control- flow perspective is concerned with the process behavior, namely the activities in the process and their order of execution. The organizational perspective focuses on the relationships between the users who performed the activities, such as whether they belong to the same or to different groups or organizational units. The performance perspective aims at detecting bottlenecks or calculating performance indicators, such as throughput times and sojourn times. The data perspective is related to the data objects that serve as input and output for the activities in a case.

To have a deeper knowledge about data involved in business process improvement, patients experience and their level of satisfaction seems very interesting. In order to implement this solution, we were based on jBPM log to extract all relevant measurements related to our process. jBPM is an open source BPMS. It features a robust management console and development tools, with user support during the business process life cycle: development, deployment, and versioning [12]. Our choice of jBPM is motivated by the availability of history logs of all process instances as a data source for measurement, the capability of storing logs and information persistence about each performed activity or occurred event that can be exploited. In parallel, the input received from the patients as results of the questionnaire about their satisfaction in the emergency department is formatted to be inserted into the PostgreSQL database.

PostgreSQL is an open source object-relational database system [13]. Essentially, this database contains all KPI values. First, we configured PostgreSQL database system to store the history log of jBPM historical data through its persistence module. Moreover, in order to formulate indicator queries correctly, we analyzed the relationship among history logs to see how the log tables are related to each other and to identify what information can be used for performance measurement purpose. Second, we translated all qualitative indicator values from the questionnaire into the new KPI table where we took into account the consistency of several quantitative and qualitative measurements based on process instance identifiers.

In order to incorporate domain knowledge, especially semantic relationships between concepts, we use protégé editor. Protégé is one of the ontology design tools [14] to construct domain ontology by storing them in several formats such as XML, RDF or OWL.

After that, we developed a Java project on NetBeans environment. The idea of this project is to create in one project all necessary information for effective decision making. This data will enable us to put together a comprehensive picture of all related components in the ontology.

Data mining is conducted through the use of well-known data mining algorithms to extract association rules from the given data. Based on KPI database, we use SIPINA the data mining tool where the Apriori algorithm for data mining was applied.

III. EXPERIMENTAL RESULTS

Multiple indicators are used in the organization at different scales and for different stakeholders. Indeed, KPIs development implicitly reflects functional requirements. So quantitative and qualitative measurement should be transmitted together with the data and evaluated regarding the target value. The lack of such information might cause undesirable effects. So we focus, in the first step, in tracking KPI values from the appropriate data sources. After that, in order to cover all missing links between quantitative and qualitative indicators and between indicators and activities in the BP, we propose a new ontology. The most significant advantage when using this ontology results in how to represent measurement and tasks without information loss into a semantically rich presentation. However, this is not sufficient and it is not adequate to just design the link between elements in the ontology. In fact, in some cases, we discovered that some quantitative indicators respect the target value but many patients are still unsatisfied. So we use data mining techniques to determine all corresponding rules in order to validate this link in the ontology or to discover a new link. The following subsections give more details.

A. KPIs data sources

This step is one of the most important steps because it is based on a deeper understanding of the requirements to formalize performance indicators. This analysis involves domain expert. Both qualitative and quantitative measurements are used in this study. First, we extracted the possible indicators related to the BP and built other indicators by means of diverse techniques such as inquiries, questionnaires or conversations. Second, we established a preliminary list of the possible indicators including quantitative and qualitative indicators. And third, we analyzed and validated the list of performance measures with experts.

In the illustrated case study, there are many measurements that could be used in the evaluation of healthcare process, but after discussion with experts, we arranged a list of relevant indicators. From this analysis, 32 quantitative indicators including aggregated indicators and specific indicators were discussed with experts. And 19 qualitative indicators were selected. But, for sake of consistency representation and reasoning, we took into account only 14 quantitative indicators. The main general reason for this selection is the necessity to focus only on basic KPIs (For example, the KPI Process Instance execution time is a basic metric which takes as input one Process Instance). We discard some measurement because they are aggregated KPIs which are built on basic KPIs (For example, the min of total time spent in the ED by all patients (all process instances) in all activities). Additionally, in order to provide a convenient way for business analysts to analyze metrics, we take into consideration only the basic metrics because the number of individuals (e.g. patients, medical staff) of the same kind (e.g. a set of Process Instances) to be analyzed varies over time and then the overall perception of some aspects of interest changes (e.g. the satisfaction will be evaluated for each individual).

The quantitative KPIs are derived from a close look at the activities involved in the BP, such as time duration in each activity and waiting time before any activity designed in the healthcare process. For qualitative indicators, a qualitative inquiry is necessary to get close enough to the patients and capture the level of their satisfaction and to examine the quality of care provided to patients attending the ED. Hence, at the end of the health process, each patient was invited to provide feedback about his/her level of satisfaction.

After that, the basic idea is in one hand, to extract relevant information from the event logs produced during the BPM execution and on the other hand, to extract relevant information from qualitative inquiry. By addressing patient queries, the qualitative aspect will be in a better position to improve satisfaction toward the healthcare processes.

Quantitative KPIs evaluation derived from business process monitoring solutions. For this reason, we were essentially based on the availability of jBPMlog to extract all relevant data where 100 instances have been created through the execution of the healthcare process. From the available data in jBPM logs, we conclude that not all KPI can be directly retrieved from these logs. So we created another data table named “KPIs real values” which contains all real values related to our research.

Qualitative Indicators reflect the qualitative aspects of care in the ED, such as staff attitudes towards patients, but also, they concern some quantitative aspects such as paramedical staff availability and the overall waiting time before treatment by paramedical personnel. This is due to the fact that for example sometimes the nurse is available but the patients who are in less urgent are not given sufficient attention by nursing staff. So, in order to provide a higher level of quality and consistency with quantitative measurement and with the processes in place, we record the patient’s level of satisfaction with those indicators. The aim of this questionnaire is to determine why patients are unsatisfied and what we can do to make it better.

In order to fulfill the qualitative database, we were based on the responses of patients, and consequently, a set of qualitative results is recorded by using a Likert scale. For each question, we invite the patient to use a 5-point Likert scale (Very satisfied; satisfied; neither satisfied nor dissatisfied; dissatisfied; very dissatisfied) to indicate his/her level of satisfaction. After that, we created a new table in the database with the respective responses of the patient as the values of KPIs. Since the exactly followed paths in the process are different, the number of questions asked to each specific patient varies and then some columns in the database have a null value.

At this stage, record all real values of indicators of both quantitative and qualitative aspects offer a better comprehension of the situation. But in order to give more sense to ensure that the health care process is fulfilling the expectations of domain experts and patients, this data needs to be checked and evaluated by the organization to detect if KPIs value is reaching the desired results. As a result, we create another table titled “KPIs estimated values” which contains all estimated values related to the patient instances. This table contains two possible values (Ok codes the fact that the KPI value is accepted by ED and Not ok corresponds to the fact that the KPI value is not acceptable).

As aforesaid, this table contains 100 records which can statistically be considered as significant. These records were obtained through the observation of various cases involved in the healthcare process in the emergency department. They are consequently considered as the basis for continuous process optimization of our context of ED.

B. KPIs ontology

This ontology provides all necessary information for effective decision making. Hence, it will allow obtaining some guidelines related to the choice and use of appropriate indicators. More precisely, first, it helps the decision maker to monitor both quantitative and qualitative performance measurements in terms of business process. Second, the ontology operates at a sufficiently detailed level such that it provides a comprehensive basis for assessing the relations between all elements of the business process.

This ontology is defined as a set of terms used to describe a given domain and derives inferences from it [15]. More precisely, it is modeled as an OWL ontology which is composed of individuals, properties, and classes. The tool used for constructing this ontology is protégé editor. We present in figure 2 our ontology for the representation of KPIs and process activities.

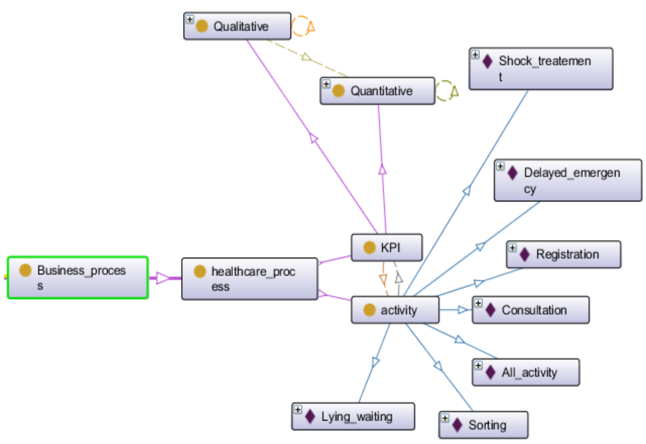

Fig. 2. main classes of our ontology

Thereafter, the inputs in this phase are process tasks and qualitative and quantitative indicators, thus facilitating their future improvement. In our ontology, individuals, represent all KPIs and all activities in which we are interested in our use case. The main benefit from this ontology is to offer the opportunity to continuously improve the BP by manipulating relationships and later to identify the reason of bottlenecks.

So in order to make more informed decisions, we collect as much data and information as possible about the possible relationships between KPIs and between KPIs and process activities. To represent, on one hand, the relationships between an activity and the attached KPI, and on another hand, the relationships between indicators, we define a set of object properties. Therefore, we create three object properties related to our KPIs. For example, the property “related_with” might link the possible individuals from the qualitative class (KPI related to a patient satisfaction) to the related individuals in the quantitative class. Also, in order to represent the relationships between indicators in the same category, the owl model implements the properties “Depend_quantitative” and “Depend_qualitative” which indicate the links between quantitative/ qualitative indicators, their relationships and mainly represent the need to share data between them. Table 1 gives an overview of some qualitative indicators linked by “Depend_qualitative” object property.

TABLE I. INDICATORS LINKED BY “DEPEND_QUALITATIVE” OBJECT PROPERTY

Depend_qualitative object property	
Qualitative indicators	***Qualitative indicators***
KPI8 the quality of care for patients by medical staff	KPI2 interest and attention brought by medical staff (doctor)
KPI7 the quality of care for patients by paramedical staff	KPI1 interest and attention brought by paramedical staff (nurses...)
KPI9 The clarity of information	KPI1 interest and attention brought by paramedical staff (nurses...) KPI2 interest and attention brought by medical staff (doctor)

Furthermore, properties can have inverses and they may be transitive or symmetric. For example, the inverse of “has_activity” is “has_indicator”. Those definitions of relationships and the relative individuals need to be highlighted and modeled in order to aid the monitoring of KPIs that contribute to performance improvement. As well, we envision a collection of Datatype properties to describe relationships between individuals and data values. Figure 3 depicts an example of “has activity”, “Depend_qualitative” and “related_with” object properties and “HasNameKpi” data property related to qualitative indicators.

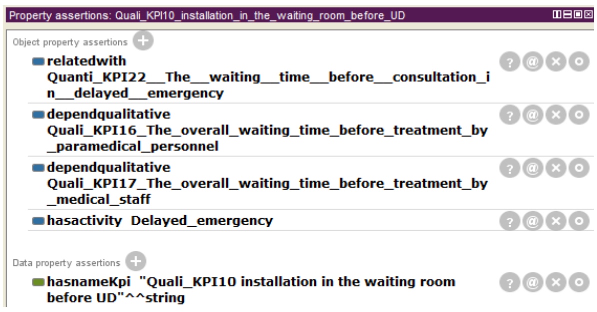

Fig. 3. Example of property assertions for Quali_KPI10 (installation in the waiting room case of Not urgent patient)

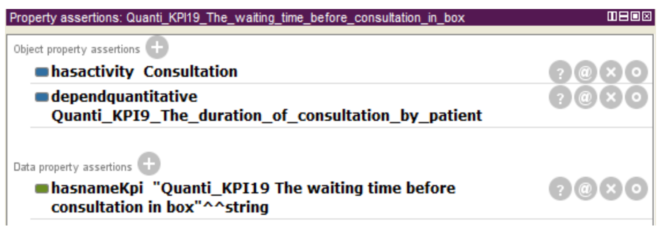

Fig. 4. Example of property assertions for Quanti_KP1I9 (The waiting time before consultation in box)

Figure 4 depicts another example of quantitative indicator "has activity" Consultation and "Depend_quantitative" with another quantitative indicator. In order to exploit the results of this ontology and to support the decision, we decided to develop a java project via the loading KPI OWL file with Jena. The Jena-OWL API is an open-source Java library for the Web Ontology Language OWL and RDF(S). And then we use SPARQL to query semantically the relationships established in owl model.

Further, the outputs of the SPARQL queries will be analyzed based on end users' views. An example of our result after running this project is shown in figure 5.

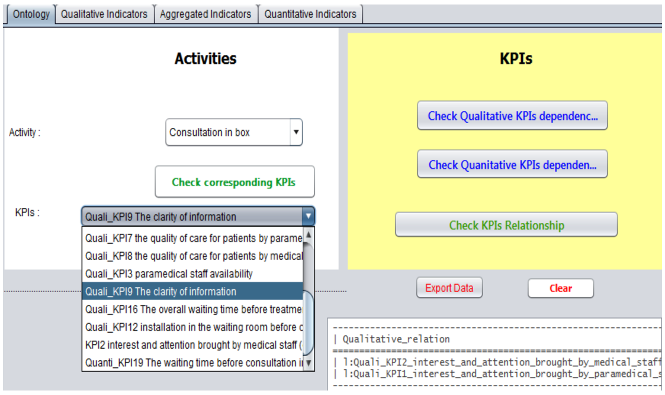

Fig. 5. Example of execution result using SPARQL query

Figure 5 displays some execution result from the proposed ontology. For example, after the user selects the activity in the BP, he clicks in "check corresponding KPIs". And after that, all appropriate indicators are displayed in the combo box list. If he wants to know which is the possible link between indicators, he selects the indicator in question and in the left there are several options. The result of each button corresponds to the main object properties established in our KPI ontology. This graphical interface helps to find relationships between quantitative and qualitative indicators and also between indicators and business process tasks.

C. KPIs discovering rules

Furthermore, it is crucial to discover interesting relations between KPIs. The task of establishing KPI ontology and mining association rules consists of two main steps. The first involves modeling of all possible relationships between indicators and indicators of appropriate activities involved in the BP from the point of view of domain experts. The second step involves testing, validating and generating all missing or predefined relations based on association rules and filtering the obtained result by adopting the appropriate parameters of the algorithm.

In fact, to make decisions, it is important to analyze all the relevant row data of our performance measurement. For this purpose, we use data mining, especially association rule learning as a research method in this stage.

An association rule is defined as an implication expression of the form $X \Rightarrow Y$, where $X, Y \subseteq I$. Note that every association rule has a support and a confidence.

To illustrate this associating rule, we take a simple example, presented in table 2, of a database D that contains 10 transaction rows where I= {i1, i2... in}: a set of all the items (Quantitative and qualitative indicators). A transaction T is a set of items such that $T \subseteq I$, for example, transaction 1 contain all measurement recorded value related to process instance 1. An itemset is a set of items. For example, X = {Quanti_KPI8, Quanli_KPI1} is a 2-itemset.The support of an itemset X is the percentage of transactions in the transaction database D that contains X.

A frequent example of item set is: {Quanti_KPI8, Quali_KPI1}. The support of the rule {Quanti_KPI8⇒Quali_KPI1} = 4 where: The support (Quanti_KPI8) = 5 and the support (Quali_KPI1) = 7. Then by definition: Confidence(X=>Y) equals support(X, Y) / support(X). So the Confidence (Quanti_KPI8⇒Quali_KPI1) = 4/5 = 0.8 or 80%

TABLE II. AN EXAMPLE FOR TRANSACTION DATABASE *D*

Transactions	Items
T_1	Quanti_KPI8, Quanti_KPI9, Quali_KPI1
T_2	Quanti_KPI8, Quali_KPI1
T_3	Quanti_KPI8, Quali_KPI1, Quali_KPI12
T_4	Quanti_KPI8, Quanti_KPI10
T_5	Quanti_KPI10, Quali_KPI12
T_6	Quali_KPI1, Quali_KPI12
T_7	Quanti_KPI8, Quali_KPI1
T_8	Quali_KPI1, Quanti_KPI10
T_9	Quali_KPI1, Quanti_KPI9,
T_{10}	Quanti_KPI10, Quali_KPI12

Apriori algorithm is interested in finding all such rules having high enough support and confidence [16]. This analysis of frequent items aims to find all interesting rules that correlate one set of items with another set of items. Association rule mining algorithm needs to be configured before learning. So, we give in advance appropriate values for the parameters in order to obtain a good number of rules. The minimum threshold of support and confidence is 10% and 50%, respectively. In order to check a link between indicators established in our ontology, we select the appropriate indicators and we apply Apriori algorithm to predict the occurrence of an item based on the occurrences of other items in the transaction.

For example, we take into account "depend qualitative" link to check the relationship between the quality of care for patients by medical staff (quali_KPI8), the interest and attention brought by medical staff (quali_KPI2), and The clarity of information (quail_KPI9). Using Tangara data mining software and based on Estimated KPI value table, we import the dataset. We insert the DEFINE STATUS component in order to define the type of attributes: quali_KPI2 is the TARGET attribute, the others are the INPUT. After that, we can use The SPV ASSOC TREE component to extract association rules from the dataset. The procedure internally uses a search tree but the outputs are rules. We must set the quali_KPI2 to the value that we want to characterize. Figure 6 shows the association rules when we set quali_KPI2 VALUE = ok. And Figure 7 shows the association rules when we set quali_KPI2 VALUE = Not ok.

"quali_kpi2" is "ok" -- IF ...

N°	Antecedent	Length	Support	Confidence	Lift
1	quali_kpi8=ok	1	0,640 (0,00)	0,914 (0,00)	1,236 (0,00)
2	quali_kpi8=ok - quali_kpi9=ok	2	0,520 (0,00)	0,912 (0,00)	1,233 (0,00)
3	quali_kpi9=ok	1	0,560 (0,00)	0,903 (0,00)	1,221 (0,00)
4	quali_kpi9=Not ok	1	0,140 (0,00)	0,538 (0,00)	0,728 (0,00)

Fig. 6. Obtained rules if quali_KPI2=ok

This figure displays all rules can be used to see what other KPIs should be taken into account to promote a high satisfaction for quail_KPI2.

"quali_kpi2" is "Not ok" -- IF ...

N°	Antecedent	Length	Support	Confidence	Lift
1	quali_kpi9=Not ok - quali_kpi8=Not ok	2	0,100 (0,00)	0,714 (0,00)	3,759 (0,00)
2	quali_kpi8=Not ok	1	0,110 (0,00)	0,647 (0,00)	3,406 (0,00)

Fig. 7. Obtained rules if quali_KPI2= Not ok

For example for Rule 2 (N°2), here quali_KPI8 is an antecedent and the quali_KPI2 is the consequent. Antecedent is the element that is found in the database, and consequent is the element that is found in the combination with the first one. The itemset {quali_KPI8 quali_KPI2} has a support out of 0.11 since it occurs in 11% of all transactions. In the first rule the itemset {quali_KPI8, quali_KPI9, quali_KPI2} occurs in 10% of all transactions. The confidence for this rule is 0.71. This value measures how often patients are unsatisfied with quali_kpi2 and also are unsatisfied with quali_KPI8 and quali_KPI9.

IV. CONCLUSION

Up to this point, a KPI data table was built and the business process concept of the domain translated into the ontology. The two table real/ estimated value of indicators of both quantitative and qualitative aspect offers a better comprehension of the situation.

To make crucial business decisions, we exploited the knowledge of the domain for business process improvement by using Data mining knowledge and domain knowledge. The use of Data mining knowledge focuses on data mining algorithms, while the domain knowledge includes an understanding of KPIs relationships among other KPIs and business process tasks.

The contribution in this work suggests that KPIs knowledge as understood from patient experience in ED and interactions with other indicators are meaningful properties for BP improvement (for which our research questions are dedicated in order to explore them). As part of our future work, we plan to enrich our KPI ontology to discover new relationships between KPIs or between activities and KPIs.

References

[1] V. Andrikopoulos, S. Benbernou, M. Bitsaki, O. Danylevych, M. Hacid, W. van den Heuvel, D. Karastoyanova, B. Kratz, F. Leymann, M. Mancioppi, K. Mokhtari, C. Nikolaou, M. Papazoglou, and B. Wetzstein. Survey on Business Process Management., July 2008.

[2] J. v. Brocke and M. Rosemann, “Handbook on business process management 1: Introduction, methods, and information systems,2014.

[3] Vallabhaneni, S. Rao. Corporate management, governance, and ethics best practices. John Wiley & Sons, 2008.

[4] D. Parmenter. Key performance indicators: developing, implementing, and using winning KPIs.Wiley, 2009.

[5] G. Pitzos; M. Matsas and G. Chryssolouris. “Defining manufacturing performance indicators using semantic ontology representation”, in CIRP 3, 2012, Athens, Greece. Proceedings… Athens, Greece, v. 3, 2012, pp. 8-13.

[6] S. Opoku-Anokye and Y. Tang. “The design of a semantic-oriented organizational performance measurement system”, in 14th Internacional Conference on Informatics and Semiotics in Organisation (ICISO), Stockholm, Sweden, p. 45-49. 2013. Available from http://centaur.reading.ac .uk/31975 [Accessed 10th March 2014].

[7] Y. Shen; D. Ruan and E. Hermans. “Modeling qualitative data in data envelopment analysis for composite indicators”. International Journal of System Assurance Engineering, v. 2, n. 1, 2011, pp. 21-30.

[8] Claudia Diamantini, Domenico Potena, and Emanuele Storti. Extended drill-down operator: Digging into the structure of performance indicators. Concurrency and Computation: Practice and Experience, 2015.

[9] Ortega DEL-RÍO-ORTEGA, Adela, RESINAS, Manuel, DURÁN, Amador, et al. Using templates and linguistic patterns to define process performance indicators. Enterprise Information Systems, 2016, vol. 10, no 2, p. 159-192.

[10] Park, Jong Hyuk James, et al., eds. Information Technology Convergence, Secure and Trust Computing, and Data Management: ITCS 2012 & STA 2012. Vol. 180. Springer Science & Business Media, 2012.

[11] van der Aalst, Wil MP. "VIEWPOINT Business process management: a personal view." Business Process Management Journal 10.2 (2004): 135.

[12] Simone Fiorini, Arun V Gopalakrishnan, Mastering jBPM6, Packt Publishing, 25 June 2015.

[13] Obe, Regina O., and Leo S. Hsu. PostgreSQL: Up and Running: A Practical Introduction to the Advanced Open Source Database. " O'Reilly Media, Inc.", 2014.

[14] Horridge, Matthew, et al. "A Practical Guide To Building OWL Ontologies Using The Protégé-OWL Plugin and CO-ODE Tools Edition 1.0." University of Manchester 2004.

[15] Usha Yadav, Gagandeep Singh Narula, Neelam Duhan, Vishal Jain and B. K.Murthy1, Development and Visualization of Domain Specific Ontology using Protege, Indian Journal of Science and Technology, Vol 9(16), April 2016.

[16] Alpaydin, Ethem. Introduction to machine learning. MIT press, 2014.

[17] Rebuge, Álvaro, and Diogo R. Ferreira. "Business process analysis in healthcare environments: A methodology based on process mining." Information systems 37.2 (2012): 99-116.