Ontology for Performance Measurement Indicators’ Comparison

Vanderlei Freitas Junior¹, Victoria Uren², Christopher Brewster², Alexandre Leopoldo Gonçalves³
¹Catarinense Federal Institute of Education, Science and Technology, Brazil; ²Aston University, United Kingdom; ³Federal University of Santa Catarina, Brazil
junior@ifc-sombrio.edu.br, v.uren@aston.ac.uk, C.A.BREWSTER@aston.ac.uk, a.l.goncalves@ufsc.br

ABSTRACT

Indicators are widely used by organizations as a way of evaluating, measuring and classifying organizational performance. As part of performance evaluation systems, indicators are often shared or compared across internal sectors or with other organizations. However, indicators can be vague and imprecise, and also can lack semantics, making comparisons with other indicators difficult. Thus, this paper presents a knowledge model based on an ontology that may be used to represent indicators semantically and generically, dealing with the imprecision and vagueness, and thus facilitating better comparison. Semantic technologies are shown to be suitable for this solution, so that it could be able to represent complex data involved in indicators comparison.

KEYWORDS

Performance measurement; semantic indicators; vagueness; imprecision; ontology; comparison.

1 INTRODUCTION

Performance measures have been recognized as important tools for organizational development [1]. Velimirovic, Velimirovic and Stankovic [2] state that continuous monitoring of performance measures has been the basis for improving organizational performance, and qualitative and quantitative expression of selected results. In other words, the authors state that performance measures enable organizations to effectively express their success through numbers.

The performance, in this context, can be understood as “sum of all processes that lead to a potential or future sequence of outcomes and results” [3]. Kaplan and Norton [4] indicate that performance can be expressed “only as a set of parameters or indicators that are complementary, and sometimes contradictory, and that describe the process through which the various types of outcome and results are achieved” [5], [4].

Thus, performance measures enable organizations to carry out a diagnosis of strengths and weaknesses, and to determine alignment with their goals by the analysis of indicators, which plays a central role in assessment processes.

Some authors, however, have pointed out difficulties in the use of indicators, since they lack semantics, and may be vague or imprecise in the statement of values that they are intended to represent [6], [7], [8], [9], [10], [11].

Pintzos, Matsas and Chryssolouris [6] have pointed out the difficulties and challenges posed by the lack of semantics in indicators, in particular the difficulty of comparing indicators and results internally, between sectors or departments of the organization itself, or between external agents. We will explore specific examples of this issue in the usage scenario presented in Section 4.

So, despite their importance in performance measurements, indicators have vagueness and imprecision, and lack semantics and these insufficiencies hinder their comparison. We will demonstrate that other information, for example, attributes and relationships between indicators, about these indicators is needed in order to allow their full comprehension and more accurate comparison.

Knowledge Engineering is important in order to face these challenges, which, according to Schreiber [12], “has evolved into the development of theory, methods and tools for developing knowledge-intensive applications. In other words, it provides guidance about when and how to apply particular knowledge representation techniques for solving particular problems”. Therefore, the purpose of this paper is to present a model to address these challenges, that can represent indicators in a semantic way. This model has to deal with vagueness, imprecision and temporality, to
allow improved comparison between performance measurement indicators.

An important consideration about the use of semantic technologies to address the challenges of this study comes from the fact that a main aspect from our research problem is the need to compare two indicators. To make any comparison, it is important to know if the indicators are equivalent, if they have the same background.

Therefore, in this paper we present a new approach to performance measurement representation, based on an ontology, which increases semantic representation of indicators, in particular specifying background information about measurements, and thereby reducing vagueness and imprecision. This in turn supports more accurate comparison. Our proposal is general and reusable, so a wide range of domains could be represented by the proposed model.

The remainder of this paper is structured as follows. Section 2 provides some preliminaries on the fundamental theoretical aspects underlying this paper. In Section 3 we present our model. Section 4 demonstrates the application of the model in a usage scenario. Some conclusions are drawn and ideas for future research are proposed in Section 5.

2. PERFORMANCE MEASUREMENT

The daily work of organizations requires their efforts to face the challenges ahead. The success of organizations, then, is connected directly to the capability to assess its reality and the variables that affect it internally and externally. In this respect, it follows that the practice of evaluation, in the broadest sense, is part of human nature and underpins decision-making tasks [1][13].

According to Neely, Platts and Gregory [14], performance measurement may be defined as a process of quantifying the efficiency and effectiveness of an action, which leads to performance.

So, performance measures have wide application in modern organizations, being understood as “qualitative, and quantitative expression of some results by chosen indicators. Performance measurement enable to effective organizations to express their success by numbers” (sic) [2]. In other words, it has the goal to translate organizational results into numbers and concepts that best express the satisfaction of the goals and objectives set by the company. Indicators play a central role in organizational performance assessment processes [7]. These authors also emphasize the role of indicators, stating that organizational performance measurement constitutes a development of measurable indicators that can be used to assess the progress toward a particular goal, as well as in achieving these goals.

2 SEMANTIC REPRESENTATION AND COMPARISON OF INDICATORS

As noted in the introduction, Pintzos, Matsas and Chryssolouris [6] have pointed out the difficulties caused by the lack of semantics in indicators. For these authors, the main characteristics that describe the context of an indicator are its name, calculation formula and measurement units. These attributes could situate an indicator, giving some meaning, allowing, as consequence, its comparison.

On the other hand, Opoku-Anokye and Tang [7] recognize that the task of setting indicators for organizational performance evaluation is multidisciplinary, since it is developed from various sources. This diversity often puts performance measurement primarily in the control of the functional units of organizations. Thus, the technical measures, models and approaches are developed with the same approach, i.e., based on functional and divisional structures of organizations, rather than semantics, business processes and product life cycle.

Pintzos, Matsas and Chryssolouris [6] present another limitation involving the lack of semantics in indicator development and performance measurement. According to the authors, the decision-making process in industry is based on performance requirements, which specify the values of the relevant production attributes. These values are expressed by internal performance indicators. The indicators may express the same metrics and use the same information, but have different definitions and calculation methods. Just to give an example, the “cost” indicator is often a monetary value and means an amount of resource needed to create some product or service. However,
the method of cost calculation could vary even in different sectors of the same company. This situation hinders performance comparisons, not only between different companies, but within the organization itself. In this context, the authors claim that semantic modeling is able to provide a broad overview of the data needed for a company to operate [6], facilitating the understanding and standardization of indicators.

Various tools and frameworks for performance measurement are available to guide planning, indicators definition, data collection and results monitoring. Among them, one of the best known and most referenced in the literature is the Balanced Scorecard (BSC). Developed by Kaplan and Norton [4], the BSC is a decision support tool at the level of strategic management, which supports improved satisfaction of an organization’s strategic objectives, starting from appropriate planning and selection of indicators.

Despite the robustness of these tools, the literature indicates some difficulties to be addressed by academic research. Bobillo et al. [11] point out that, some variables or indicators are associated with some vagueness and imprecision, for example where it is more natural to refer to their values through linguistic expressions (e.g. low, medium, high) instead of numerical values. Furthermore, the authors also point out the difficulty caused by the lack of an explicit representation of their semantics.

This situation hinders performance comparison not only between different companies, but within the organization itself. In this context, Pintzos, Matsas and Chryssolouris [6] claim that semantic modeling is able to provide a broad overview of the data needed for a company to operate, facilitating the understanding and standardization of indicators.

It becomes clear that there exists a need for a generic semantic indicators model that would be sufficient to attenuate the vagueness and imprecision of indicator values by supplying background information on measurement conditions, allowing a more accurate comparison between indicators.

Bobillo et al. [11] also point out that the variables involved in planning and performance measurement may have inaccuracies and vagueness, which can interfere in results analysis and interpretation. On the other hand, they also claim that the data collected by the performance indicators and variables in the main methodologies are lacking in semantics, damaging the interpretation and analysis of these data. This difficulty makes the process of comparison between these indicators fragile.

The indicators, in summary, are widely used in organizations to assess institutional processes. However, the phenomena of imprecision and vagueness, and the lack of semantics, render their interpretation and comparison difficult.

3 SEMANTIC TECHNOLOGIES AND THE PROPOSED MODEL

Our proposed model is based on semantic technologies. These technologies were selected for three main reasons outlined below.

First, according to Belhadef, Eutamene and Kholadi[15], “the concept of ontology is a concept that is not always easy to characterize. Indeed, it is used in different contexts: philosophy, linguistics, intelligence (AI), and each one's have its particular definition” (sic). But, in the context of this study, ontologies are considered as a well-established technology to represent knowledge in a specific domain. According to Gomez-Perez [16], an ontology can be used as a knowledge base skeleton, where inference processes or reasoning are executed. The use of ontologies for domain knowledge representation allows improvement in the information extraction process and the exchange of knowledge, which is one of the major motivations of this study. Gobin [17] reports that an ontology definition is based on the idea of conceptualization, i.e., a simplified version of the real world needing to be represented by providing “a shared and common understanding of a domain that can be communicated across people and application systems”.

The second aspect of the use of semantic technology is the easy comprehension of the model. An ontology is used to represent consensual knowledge in a specific domain, and this aspect makes them easy to understand. Even users without prior knowledge of the technology or the domain could understand the meaning of all concepts expressed by the ontology. Gomez-Perez [16]
reports that clarity and objectivity are principles that have proved useful in development of ontologies. It “means that the ontology should provide the meaning of defined terms by providing objective definitions and also natural language documentation”.

The third reason for the use of semantic technologies is the opportunities for reuse of existing components. The process of ontology design takes into account the possibility of the reutilization of some parts of other ontologies. Guarino [18] confirm that “an important benefit of using an ontology at development time is that it enables the developer to practice a "higher" level of reuse than is usually the case in software engineering (i.e. knowledge reuse instead of software reuse). Moreover, it enables the developer to reuse and share application domain knowledge using a common vocabulary across heterogeneous software platforms.”

An important consideration about the use of semantic technologies to address the challenges of this study comes from the fact that a major aspect from our research problem is the need for comparison between two indicators. To make any comparison, it is important to know if the indicators are equivalent, if they have the same background, if they are using the same attributes, the same formulae etc.

Furthermore, Tang [19] indicate two advantages of the ontologies over databases, as a well-known technology to represent and organize data. The author emphasizes that “as a comparison, database technology does not readily illustrate relationships among data entities. Another advantage of this approach over that of the database is that complex relationships between and among classes may not be so easily defined” in a database.

Thus, taking into account the scientific literature presented, we can identify the necessity to create a knowledge model for generic representation of indicators, considering the treatment of background information in order to allow the comparison between indicators with improved accuracy.

Knowledge Engineering is important in order to face the challenges described in this paper. It aims to provide systems capable of explicitly representing and storing the knowledge of the organization, considering all the systemic organizational context of knowledge intensive tasks [12].

Knowledge Engineering, therefore, provides instruments for knowledge-based systems modeling and developing that are able to explicitly formalize and represent knowledge for knowledge intensive tasks.

Taking into account all these tools, we highlight, within the limits of this work, the ontologies.

Ontologies are explicit specifications of the resource types and the possible relationships between them. In addition, they may include specific instances of concepts in the ontology [20].

An ontology can also be defined as a set of terms hierarchically ordered to represent a specific domain. It can be used as a knowledge base skeleton where inference processes are executed (reasoning). So, the use of an ontology allows the knowledge engineer to define a domain, allowing improvement in the information extraction process and the exchange of knowledge [21].

Another definition widely accepted by ontology engineers [17] is authored by Gruber [22] who argues that ontology is an explicit specification of a conceptualization. Conceptualization is an abstract model of the world to be represented and that representation must explicitly specify the concepts, properties and relations. Borst [23] modified this definition, stating that an ontology is a formal specification of a shared conceptualization. He emphasized that there must exist a model in the specification of the ontology and the conceptualization should be done in such way to allow sharing. The expressed knowledge should be common sense and not particular to the person writing. Studer, Benjamins and Fensel [24] complement the definition asserting that an ontology is a formal, explicit specification of a shared conceptualization, reinforcing the requirement that the specification be explicit.

To Gobin [17], this definition is based on the idea of conceptualization, i.e., a simplified version of the real world represented by providing a common and shared vision of a domain that can be communicated between people and systems.

To Dillon and Simmons [25], ontologies support sharing a common understanding of the
information structure among people or software agents, but not only that, they also make it possible to reuse the given knowledge domain, and explain the assumptions, separating the operational knowledge from the domain knowledge, in addition to permit the analysis of domain knowledge. The ontologies engineering, in turn, is defined by the same authors as a highly collaborative process, since an ontology developed fairly accurately will be useless if it will not be accepted by domain experts, who should be directly involved in its development.

Bobillo et al. [11] state that ontologies allow data to be enriched with semantics, permitting automatic checking of data consistency, and giving an easier way to maintain the knowledge base and reuse of components.

Thus, assuming ontologies as tools for computational representation of specific knowledge of a domain, giving it meaning, ontologies can be used to provide semantics to indicators modeling.

In order to address the representation of performance measurement indicators, considering imprecision, vagueness, temporality and relationship between other indicators, with the aim of improving comparison of indicators, we propose a model based on ontology, as shown in Figure 1.

According to the literature, we can say that we have to know more about a given indicator, to understand its real meaning and the implications of its value. More complete knowledge about the origins of the data involved in the calculation of the value of an indicator could create the basis for a more accurate comparison between indicators, improving comprehension of its meaning and allowing the decision on what indicators could be compared and used for a decision-making process.

We also believe that we should know, at least, the following information about an indicator in order to facilitate improved comprehension and comparison:

- Entity interested in an indicator.
- Variables the indicator is related to.
- The description, the value and the importance (weight) of the indicator.
- What kind of criteria are used in a given domain to assess the indicator.
- What kind of formula or mathematical calculus is executed to calculate the indicator value.
- What other attributes compose the indicator value.
- Which point in time the indicator is related to.

**Figure 1.** Proposed model
• What relations between other indicators need to be known.
All of these needs are represented by classes and properties in the proposed model, shown in Figure 1.

The ontology consists of six classes.
The “Entity” class is related to the “Variable” class by the "hasVariable" property in order to represent the relationship between an entity and the variable to be analyzed, which in turn, comprises one or more indicators.
The “Variable” class is related to the “Indicator” class through the "hasIndicator" property in order to represent the relationship between an indicator and the assessed variable referred to by this indicator.
The “Indicator” class, in its turn, has three properties. The first, "hasDescription" has the objective of providing a text description of the indicator. The second, called "hasValue" indicates the final value of the indicator calculation. Finally, the property "hasWeight", seeks to establish the importance of this indicator in an instantiated set of indicators by the ontology.

Also, in order to represent the relationships between indicators, the model implements the properties "givesDataTo" and "takesDataFrom" which indicate the links between indicators, their relationships and mainly represent the need to share data between them.
The “Indicator” class relates to the other three classes, which seek to ensure their meaning: “Criteria”, “Formula” and “Attribute”.
The “Criteria” class stores the necessary information for classification of represented indicators, from which you can set the sufficiency of the calculated values.
The “Formula” class tries to show the calculation methods of an indicator, ranging from simple arithmetic expressions such as sum or average, to complex mathematical formulas needed to calculate the numerical value represented by the indicator. This class is linked to the class indicator for the property “hasFormula”.
The “Indicator” class is also related to the “Attribute” class through the property “hasAttribute”. This class has the objective of representing involved features and variables in the calculation of the modeled representation indicator value. These attributes are therefore part of the process of semantic enrichment of indicators. They allow the description of elements that complement and ensure meaning, and give the necessary data to determine indicator values. The “Attribute” class has two properties: “hasDescription” and “hasValue”. The first is responsible for describing the feature to be modeled, and the second for assigning a value.

In the next section we will see an usage scenario of our proposed ontology.

4 USAGE SCENARIO

An example could be given to demonstrate the use of the proposed model in order to represent indicators, dealing with imprecision, vagueness and the relations between indicators, which shows how comparison is supported. A given company with operations across the country, needs to analyze its fleet fuel consumption, and needs this to be done car by car.

A report will show the fuel consumption of each car, telling that a car “A”, located in the North of the country, had an average fuel consumption of 12 km/l. And a car “B”, located in the South of the country, had a consumption of 13 km/l.

Now, in order to analyse this data, we have to make comparisons, and the question we have is: are both car equally economical? Is it possible to compare the consumption of car “A” and car “B”?

We have to know some more details about these indicators in order to make a comparison between these two values. All this information is represented in the ontology shown in Figure 2.

According to the specialists [26], a car’s fuel consumption could be influenced by a lot of factors, like seasonal features, (e.g. temperature, wind), and the use of some components of the car, like air conditioning.

To establish an accurate comparison, variables must be defined by domain specialists or according to company interests. The ontology allows these adjustments, adapting to each situation. Thus, all of these variables must be analyzed to determine if both indicators are equivalent and could be fairly compared.

Relations and information about the indicators are shown in the Figure 2.
In our case of comparison, we have two indicators that could be assessed as the same, but each one has its own particularities, as we can see in Figure 3.

The car “A” average consumption and the car “B” average consumption indicators has seven attributes that must be considered in order to be assessed: consumption, distance, location, average time of air conditioning use, average temperature and average wind speed and period.

Apparently, car “B” is more economical than car A, because it has travelled 130 km instead of the 120 km of car A, using 10 l of fuel. But we have to consider other variables involved in the representation of these indicators.

Car “A” has travelled 120 km using 10 l of fuel. It is located in North of the country, where the average temperature is around 35ºC, which demands 8 hours of use of air conditioning per day. The wind, in this part of the country, is around 10 km per hour.

On the other hand, car “B” has travelled 130 km using the same 10 l of fuel in a very different environment. This car is located in the South of the country, where the average temperature is 15ºC. In such temperatures, cars do not need to use air conditioning. The wind speed is around 20 km per hour in this part of the country at this period of the year.

So, as we can see, some factors that could increase a car’s fuel consumption are affecting the indicators related to car “A”, and apparently making it less economical than car “B”. However, analysing all variables we could say that, despite of all the factors, car “A” is just 1 km/l, on average, less economical than car “B”.

Dealing with all these variables, a lot of other calculus could be performed to achieve an accurate comparison, considering, for example, what is the impact of the use of air conditioning in the car consumption, or even what is the influence of the wind or temperature on the performance of

Figure 2. Car consumption ontology
the engine, and how the company cost is increasing because of these variables.

Figure 4 shows another analyses that could be performed over our model. As we may see, some attributes were considered higher than in Car “B”, others were considered equal or lower. Two attributes can be considered as just information in order to increase the semantic of the indicator. Analysing this attributes, we can say that in North of this given country, the temperature is higher in July then it’s in the South, what demands more time of air conditioning. In this case, as the cars have used the same amount of fuel, this difference of temperature and air conditioning can be reflected in lower covered distance by car A.

This usage scenario has demonstrated the importance of considering a set of information when making a comparison between indicators. We have demonstrated that ontologies can represent the background, increasing semantics and exactly situating the indicator in its domain, reducing vagueness and giving improved comprehension of the background of the measurements.

5 CONCLUSIONS AND FUTURE RESEARCH

In this paper, we proposed a semantic model to represent performance measurement indicators, based on an ontology, which addresses vagueness, and the lack of semantics concerning the background of measurements, in order to allow improved comparison between indicators.

The semantic technologies have been shown to be able to represent all the complex data involved in the comparison of indicators.

In the usage scenario proposed, we demonstrated that in a huge country like Brazil, with different seasonal conditions in different regions (or even in a comparison between the consumption of cars of the same company in different countries in Europe) the task of comparing indicators is non-trivial. It is important to take into account factors involved in the use and fuel consumption of the car. In the same country, or region, the same season may have different weather conditions, temperature or wind, for example. Such variables could affect fuel consumption, e.g. where a higher temperature could increase the regular consumption due to use of air conditioning. Background information is therefore needed about an indicator to be able to make comparisons, and semantic technologies are ready to address this sort of challenge.

The main contribution of this study is a new approach to indicators representation in order to support comparison. The model has the potential to allow computer-based comparison between two or more indicators. An interesting direction of further
research would be to implement a knowledge-based system to simulate the use of the model in a real and computational situation.

As future research we propose two main aspects of the model: time and imprecision. We know that time plays a central role in indicators comparison, because a given indicator represents a specific point in time, or even a specific period of time. So, a model that could expand the representation of time, allowing different granularities, could be more effective in order to give a complete semantic representation of indicators.

On the other hand, dealing with imprecision could be a good way to improve representation of indicators. In our usage scenario, some questions could be answered, like: what is the economical level of car “A” related to car “B”? What is the impact of the variables in the consumptions of the car? Some linguistic variables could be associated to this analysis, like “high”, “low”, “better”. The specific literature about fuzzy logic offer some answers to this questions, so a model that could represent this imprecision could improve the indicators’ representation and increase the performance of the reasoning.

6 ACKNOWLEDGEMENTS

This work was partially supported by the Ministry of Education of Brazil. Vanderlei Freitas Junior holds a PDSE scholarship from CAPES Foundation, Ministry of Education of Brazil, Brasília - DF 70.040-020, Brazil.

7 REFERENCES


