Ontology-driven approach for KPI meta-modelling, selection and reasoning

María del Mar Roldán-García, José García-Nieto , Alejandro Maté , Juan Trujillo , José F. Aldana-Montes

A B S T R A C T

Keywords:

 Ontology
KPI Modelling
Semantics
Reasoning
Knowledge extraction
Water management

A key challenge in current Business Analytics (BA) is the selection of suitable indicators for business objectives. This requires the exploration of business data through data-driven approaches, while modelling business strategies together with domain experts in order to represent domain knowledge. In particular, Key Performance Indicators (KPIs) allow human experts to properly model ambiguous enterprise goals by means of quantitative variables with numeric ranges and clear thresholds. Besides business-related domains, the usefulness of KPIs has been shown in multiple domains, such as: Education, Healthcare and Agriculture. However, finding accurate KPIs for a given strategic goal still remains a complex task, specially due to the discrepancy between domain as-sumptions and data facts. In this regard, the semantic web emerges as a powerful technology for knowledge representation and data modeling through explicit representation formats and standards such as RDF(S) and OWL. By using this technology, the semantic annotation of indicators of business objectives would enrich the strategic model obtained. With this motivation, an ontology-driven approach is proposed to formally concep-tualize essential elements of indicators, covering: performance, results, measures, goals and relationships of a given business strategy. In this way, all the data involved in the selection and analysis of KPIs are then integrated and stored in common repositories, hence enabling sophisticated querying and reasoning for semantic validation. The proposed semantic model is evaluated on a real-world case study on water management. A series of data analysis and reasoning tasks are conducted to show how the ontological model is able to detect semantic conflicts in actual correlations of selected indicators.

1. Introduction

Modern Business Analytics (BA) demand reference models to enhance interoperability and common virtual spaces for advanced data consolidation and analysis. In this environment, the selection of suitable indicators for business objectives is a current challenge (Tenneson & Brocklehurst, 2018), which requires the exploration of business data through data-driven approaches, while modelling business strategies together with domain experts in order to represent domain knowledge. In particular, Key Performance Indicators (KPIs) allow human experts to properly model ambiguous enterprise goals by means of quantitative variables with numeric ranges and clear thresholds. Besides business-related domains, the usefulness of KPIs has been shown in multiple domains, such as: Education (Maté, de Gregorio, Cámara, & Trujillo, 2014), Healthcare (Barone, Topaloglou, & Mylopoulos, 2012)

and Agriculture (Benke & Tomkins, 2017).

Nevertheless, finding precise KPIs for a given strategic goal is still a complex task, since there is a general lack of conceptualizations and data-driven approaches to capture the differences between performance (lead) and result (lag) indicators. In this regard, the semantic web emerges as a powerful technology for knowledge representation and data modeling through explicit representation formats, ontologies, vocabularies and standards, such as RDF(S) and OWL. Specifically, ontologies describe concepts, relationships, classes, individuals, formal logic axioms and objects of a particular domain (Gruber, 1995). The objects refer to entities and events (concepts) in the real-world, and their relations represent the semantic links between these entities. Using this technology, the semantic annotation of indicators of business objectives would enrich the strategic model obtained.

In this sense, a series of studies have been appearing in the last years,

in which ontological approaches are suggested to enhance data warehouse by means of semantic representations (Bellatreche, Khouri, & Berkani, 2013; Huang, Chou, & Seng, 2007; Nebot & Berlanga, 2012). These studies are mostly aimed at bridging the gap between the strategical definition of indicators and the design of data cubes, supporting the design of data marts and OLAP analysis. Another interesting proposal is KPIOnto (Diamantini, Potena, & Storti, 2016), which is designed to represent the formulas used to calculate KPIs, hence allowing to make the algebraic relationships among indicators explicit. However, in these works, all indicators are treated as KPIs disregarding whether they refer to actual performance (KPIs) or measured results (Key Result Indicators – KRIs) (Horkoff et al., 2014; Maté, Trujillo, & Mylopoulos, 2016). The distinction between the concepts of KPI and KRI (Parmenter, 2015) is crucial to avoid mistakes and enable the process of finding an adequate KPIs to focus on for an enterprise objective. Therefore, the generation of standardized models to support formal analysis and exploration in this direction is a challenge in current Business Analytics (Simoni, 2018).

This motivates us to propose an ontology-driven approach to formally conceptualize essential elements of indicators, covering: performance, results, measures, goals and relationships of a given business strategy. In this way, all the data involved in the selection of KPIs are then integrated and stored in common repositories, hence enabling sophisticated querying and reasoning for semantic validation.

The proposed semantic model is evaluated on a real-world case study on water management, which comprises an RDF[2] (Resource Description Framework) repository that follows the ontology scheme. This repository can be queried by high level algorithms using SPARQL. The goal is to properly feed post processing procedures capable of guiding the design of accurate KPIs. As a proof-of-concept, a series of data analysis and reasoning tasks are conducted to show how the ontological model is able to detect semantic conflicts in actual correlations of selected indicators.

The main contributions of this study can be outlined as follows:

- The proposed ontology, called KPIOWL, has been designed and implemented in OWL 2 for the representation and consolidation of domain knowledge for the elicitation and selection of Key Performance Indicators. It considers a large and complemented set of concepts, attributes and relationships that have been taken from business intelligence field.
- A semantic approach has been implemented for the semantic annotation of all the involved concepts and measures from the data sources, as well as those process and components required. The concepts are integrated following the ontology structure and stored in a common RDF repository.
- The semantic model is evaluated in the context of a real-world use case of water supply network management company. A series of tests and reasoning tasks are conducted to explore objectives and measures when monitoring water supply networks in the Mediterranean region of Alicante (Spain). Reported results allow us to support domain experts in decision making process.

The rest of this paper is structured as follows. In Section 2, background concepts and literature overview are presented. In Section 3, the metamodel for elicitation and selection of business indicators is explained. Section 4 describes the proposed semantic model, with design and implementation details. Section 5 presents the use case for testing and validation. In Section 6, comments concerning theoretical and managerial implications are included. Conclusions and future work are drawn in Section 7.

2. Background and related work

In this section, the main background concepts in the field of Semantic Web and Key Performance Indicators are briefly explained for the sake of a better understanding of this paper. A set of most important related works in the state of the art are also revised in order to point out the main differences with regards to the proposed approach.

2.1. Background concepts

- Ontology. In accordance with Noy, McGuinness, et al. (2001) and Guarino et al. (1998), an ontology provides a formal representation of the real world. It defines an explicit description of concepts in a domain of discourse (classes or concepts), properties of each concept describing various features and attributes of the concept (properties) and restrictions on properties. Ontologies are part of the W3C standard stack of the Semantic Web.[3] An ontology together with a set of individual instances of classes constitutes a knowledge base and offer services to facilitate interoperability across multiple, heterogeneous systems and databases.
- RDF. Resource Description Framework (McBride, 2004) is a W3C recommendation that defines a language for describing resources on the web. RDF describes resources in terms of triples, consisting of a subject, predicate and object. RDF Schema (RDFS) (Staab & Studer, 2013) describes vocabularies used in RDF descriptions.
- OWL. The Ontology Web Language is used to define ontologies on the Web, which extends RDF and RDFS, but adding a vocabulary. From a formal description, OWL is equivalent to a very expressive description logic DL, where an ontology corresponds to a Tbox (Gruber et al., 1993). This equivalence allows the language to exploit description logic research results. OWL provides two sublanguages: OWL Lite for simple applications, and OWL DL, which represents the subset of language equivalents to description logic, those reasoning mechanisms of which are quite complex. OWL-DL is syntactic description that gives maximum expressiveness while retaining computational completeness and decidability (Mcguinness, Van Harmelen, et al., 2004). The complete language is called OWL Full. In October 2007, a new W3C working group was created to extend OWL with several new features, as proposed in the OWL 1.1 member submission. This new version is called OWL 2 (Group, 2019). W3C announced the new version on the 27th October 2009. OWL 2 also defines three new profiles, OWL 2 EL, OWL 2 QL and OWL 2 RL (Group, 2019), and a new syntax (OWL 2 Manchester Syntax). In addition, some of the restrictions applicable to OWL DL have been relaxed; as a result, the set of RDF Graphs that can be handled by description logic reasoners is slightly larger in OWL 2.

  In this sense, OWL-DL is syntactic description that gives maximum expressiveness while retaining computational completeness and decidability (Mcguinness, Van Harmelen, et al., 2004). For all these reasons, we use OWL 2 to define our proposed ontology, since it is a popular expressive language that adds several new features to OWL (first version), including increased expressive power for properties, extended support for datatypes, simple meta-modeling capabilities, extended annotation capabilities, and keys. Like OWL, OWL 2 specifies a precise mapping from ontology structures to RDF graphs. Conversely, OWL 2 also benefits from an explicitly specified mapping from RDF graphs back to ontology structures.
- SPARQL is a query language for easy access to RDF stores. It is the query language recommended by W3C (Harris, Seaborne, & Prud'hommeaux, 2013) to work with RDF graphs (Prud, Seaborne, et al., 2006), then supporting queries and web data sources identified by URIs.

M. SWRL. The Semantic Web Rule Language provides the OWL-based ontologies with procedural knowledge, which compensates for some of the limitations of ontology inference, particularly in iden-tifying semantic relationships between individuals (Horrocks, Patel-Schneider, Bechhofer, & Tsarkov, 2005). SWRL uses the typical logic expression "*Antecedent ⇒ Consequent*" to represent semantic rules. Both antecedent (rule body) and consequent (rule head) can be conjunctions of one or more atoms written as "$atom_1 \wedge atom_2 \wedge \cdots \wedge atom_n$". Each atom is attached to one or more parameters represented by a question mark and a variable (e.g., ? *x*). The most common uses of SWRL include transferring characteristics and inferring the existence of new individuals (Grosof & Poon, 2004).[4]

2.2. Related work

Since the last decade, there have been appearing a series of studies in which ontological approaches are proposed to take part in different steps of Business Intelligence processes (Berkani, Bellatreche, & Benatallah, 2016; Gómez et al., 2017; Nebot & Berlanga, 2012; Sell, Cabral, Motta, Domingue, & Pacheco, 2005). In particular, in the field of data warehouse, there exists a number of works (Bellatreche et al., 2013; Huang et al., 2007; Nebot & Berlanga, 2012) in which semantic representations are mostly designed for bridging the gap between the strategical definition of indicators and the design of data cubes, supporting the design of data marts and OLAP analysis. In this regard, the RDF Data Cube vocabulary (QB) (Cyganiak & Reynolds, 2014) is a W3C recommendation for the publication of multi-dimensional data on the web, such as statistics. It defines the dimensions, attributes and measures used in the dataset and it builds upon existing RDF vocabularies (for example, SKOS (Miles & Brickley, 2005), SCOVO (Hausenblas, Halb, Raimond, Feigenbaum, & Ayers, 2009), Dublin Core (Group, 2017), FOAF (Brickley & Miller, 2000), etc.). The Data Cube vocabulary is compatible with SDMX (Statistical Data and Metadata eXchange) (Community, 2019), an ISO standard for exchanging and sharing statistical data and meta-data among organizations. Nevertheless, the RDF Data Cube vocabulary shows a series of limitations to fully support the multidimensional model operations. An extension of QB to overcome its limitations is QB4OLAP (Etcheverry, Vaisman, & Zimányi, 2014), which allows the implementation of OLAP operations, such as: rollup, slice, dice, and drill-across by means of SPARQL queries.

From a bottom-up perspective, early formal specification languages (Popova & Treur, 2005) and ontologies (del Río-Ortega, Resinas, & Ruiz-Cortés, 2010) are also used for the annotation of structural definition of indicators to enhance business goals representation (del Mar Roldán García, García-Nieto, & Aldana-Montes, 2016). An example of this can be found in Kehlenbeck and Breitner (2009), where based on MathML markup language, an ontology is defined to explicitly define formula of indicators, hence allowing the automatic linking to specific data warehouse elements. In del Río-Ortega et al. (2010), an ontology is designed for the definition of Process Performance Indicators (PPIs) that explicitly defines the relationships between the indicators and the elements, also enabling the analysis of PPIs at design-time. Another interesting proposal is KPIOnto (Diamantini, Genga, Potena, & Storti, 2014), which is designed to represent the formulas used to calculate KPIs, hence allowing to make the algebraic relationships among indicators explicit. An extension of KPIOnto is conducted by means of SemPI (Diamantini et al., 2016), a semantic framework for representing Key Performance Indicators that supports the construction and maintenance of a minimal and consistent dictionary.

Recently, a KPI-based OWL-Q ontology has been proposed in Kritikos, Plexousakis, and Woitch (2018), which enables formally and fully specifying how KPIs can be measured over Business Process as a Service

(BPaaS) hierarchy components. Authors of this last work argued that via introducing KPI metric hierarchies that span the whole BPaaS hierarchy, the measurably of KPIs is guaranteed.

Despite these advances, a common drawback can be identified in these previous works. Even though they enable the calculus of KPIs, they do not conduct advanced semantic analysis to connect underlying data with business objectives, indicators and, most importantly, business rules and domain knowledge. The lack of distinction between actual performance (KPIs) and results obtained (KRIs), makes difficult to evaluate the suitability of indicator and strategic models built from a business perspective. Therefore, in order to achieve deeper strategic analysis it is required not only to provide the ability to calculate KPIs together with an ontological foundation, but also that this ontological foundation is aimed to test domain knowledge and assumptions against data facts.

The KPIOWL ontology proposed here is designed for filling this gap, e.g., for covering representation and consolidation of domain knowledge for the elicitation and selection of KPIs, including KRIs, measures, and further data analysis, as crucial concepts in this process. The aim is not to substitute previous ontologies, but to complement them by incorporating new elements, attributes and relationships leading to future ontology alignments (e.g. with KPIOnto) and the generation of Linked Open Data extensive models. In this way, the proposal in this paper leads to a framework that enables advanced analysis and reasoning over business strategies, which was not previously possible.

3. Elicitation and selection of business indicators

In Business Intelligence, existing modeling languages (Horkoff et al., 2014; Silva Souza, Mazón, Garrigós, Trujillo, & Mylopoulos, 2012) include a large set of concepts that are required for analyzing different aspects of the business strategy, such as dependencies across organizations, external influences, or the business mission and vision. However, these concepts fall out of the scope when modeling KPIs for the elicitation process, and do not provide the expressiveness required for the indicator analysis.

In order to keep the analysis focused, we follow a reduced metamodel proposed in Maté et al. (2016) that includes only those concepts required for applying our methodology, and can be integrated as an extension for any of the existing modeling languages. A design overview of this metamodel, created using the Ecore framework[5] within the Eclipse platform, is shown in Fig. 1, whose main concepts are detailed as follows:

- **Goals:** Desired state of affairs that represent business objectives. For example in water management we can define the goal "*Minimize water lost*". Goals are the basic blocks of the strategic model, and we do not need to make a distinction of whether they are strategic, operational or tactical for the analysis. They are the most common concept included in strategic modeling languages (Horkoff et al., 2014; Silva Souza et al., 2012).
- **Relationships:** They allow domain experts and analysts to express the expected relationships between goals to be achieved. Relationships can either be contributions, where one goal affects positively or negatively the achievement of another goal, or decompositions, where a goal is further refined into more detailed, finer-grained goals. For example, "*Minimize leaks*" could be decomposed into "*Minimize transportation network leaks*" and "*Minimize distribution network leaks*". The expressiveness of relationships varies across strategic modeling languages. In our language, relationships have the evidence property, which captures the results from the data analysis showing whether the relationship is supported by the data or not. The evidence property is evaluated in practice by using a

---

[4] https://www.w3.org/Submission/SWRL/.

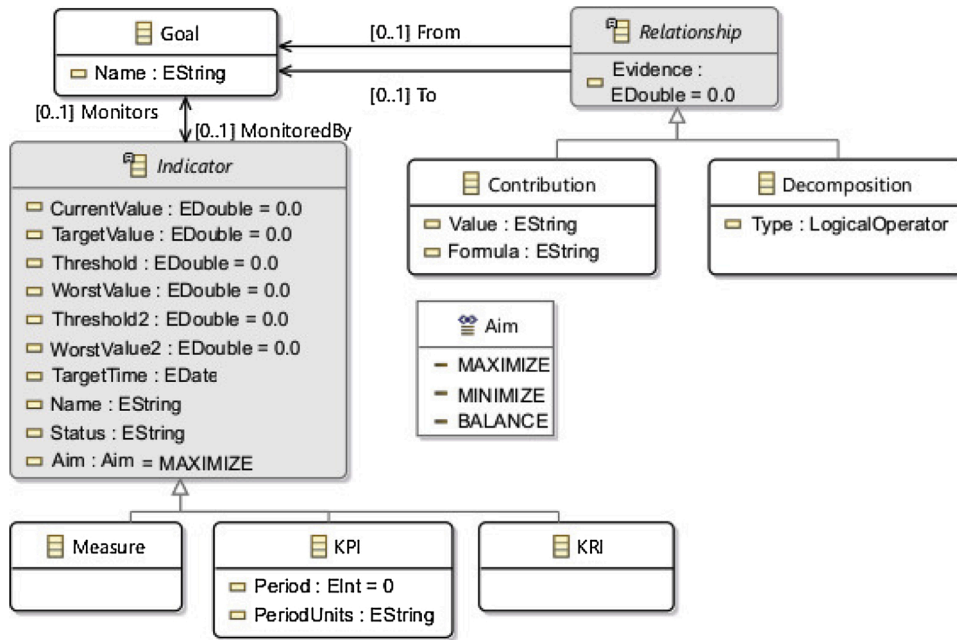[5] https://www.eclipse.org/modeling/emf/.

Fig. 1. Metamodel with the concepts and relationships for out modeling language.

combination of sentinels and statistical techniques (namely correlation and cross-correlation when there is insufficient information for building more complex ARIMA models) (Maté et al., 2016).

- **Indicators:** They measure the satisfaction of goals. They translate business objectives into measures that can be monitored, such as "*Decrease water lost by 10%*". In order to make indicators from our model compatible with existing modeling proposals (Horkoff et al., 2014; Silva Souza et al., 2012) all indicators can have a formula, a current value, a target value, a threshold, a worst value, and a target time. Furthermore, they also have a status, which provides information on the status of the indicator with respect to the data available. In addition, we have added a new property to the indicator construct in the metamodel – aim – in order to improve the reasoning capabilities. Aim denotes whether the indicator is trying to be maximized, minimized, or stabilized. This property is derived from the target and threshold values. If the target value attribute is below the threshold attribute, then, the indicator is a minimization indicator. Conversely, if the target value is above the threshold value, then the indicator is a maximization indicator. Finally, if the second threshold value is not null, and the target value is in-between, then it is a stabilization indicator, which essentially condenses a maximization and minimization indicator into a single one. Any other combination would mean an inconsistent indicator definition that should be revised.

Indicators in our language are further specialized into three types for capturing the nature of the analysis carried out in this methodology, which not found in other modeling languages:

- **Measures** are the simplest form of indicators. They represent known formulas for measuring business activities, with no known targets or thresholds. Due to the absence of any clear criteria, they cannot be used to make any statements with regards to goal satisfaction. For example, given the "*N of water connections*" measure, we cannot argue whether the associated objective has been fulfiled or not. Nevertheless, measures are the most common representation of data from enterprises, and they serve a fundamental role, being candidates to define KPIs and KRIs.
- **Key Result Indicators** are indicators which evaluate the actual satisfaction of a goal and have a set time to meet its target. For

example, "*Decrease water lost by 10%*" is a KRI. It provides information about the results of the business objective "*Minimize water lost*" and must be achieved within the lifespan of the current business plan. Other examples of KRIs we can find in different sectors are "*Increase sales by 3%*", "*Decrease average inpatient stay by 10%*", or "*Increase overnight stays by 15% during summer*". KRIs differentiate from the concept of KPIs in two critical aspects. First, KRIs cannot be affected directly by the company. They must be improved through changes in company's processes that are *expected* to improve a company's results. A water management company cannot simply decrease water lost, they have to try to minimize it by fixing open leaks and the health of the water network. In a similar way, a company can increase the number of workers at an industrial plant to directly affect the "*Production rate*". However, they cannot do anything to directly increase their sales. In turn, this means that any KRI unrelated to KPIs in the model is essentially a goal that is not being actively pursued. Second, they have no predictive power by themselves outside trend analysis. Their main objective is to evaluate the actual effectiveness that business initiatives eventually had.

- **Key Performance Indicators** are indicators that measure the performance of key activities and initiatives. They follow the intuitive idea that performing well will lead to obtaining good results. Like KRIs, KPIs have clear defined thresholds. However, they may or may not have a target time, since they can be used in monitoring continuous tasks, thus having a period instead. For example, "*Time since last network maintenance under 2 years*" is a continuous task. Other examples of KPIs are "*# of daily complaints unattended*", "*Average medical consultation time*", or "*Total food costs*". KPIs are important for the semantic analysis due to the ability of the company to affect them directly through business rules and policies. Due to their nature, if KRIs change, it is likely that the set of KPIs to be monitored also changes. The great advantage of KPIs is that they can provide information ahead of time about underlying problems or the expected results of the company. However, the information they provide w.r.t. company goals is not always accurate, as KPIs only measure a subset of factors that influence the outcome of a KRI.

This metamodel enables the construction of strategic and semantic models focused on indicators in collaboration with domain knowledge experts. The process for building the initial strategic model is

approached in a top-bottom fashion, from the identification of the main goals pursued by the organization, to the definition of candidate measures, KRIs, KPIs and their relationships. In case any candidate KRI, KPI, or measure is not related to any goal, it is listed and included into the model with no relationship to the rest of elements.

4. Semantic model

After the strategic modeling step, a candidate model is obtained with varying degrees of completion depending on the knowledge available about the business and its measurement. The indicators in the model are then used as input to the second step in this process, the data analysis, in order to test their suitability according to existing data. At this point, the complete process is semantically modeled to allow the representation and consolidation of all data involved in the selection of KPIs, hence enabling sophisticated querying and reasoning for semantic validation.

The semantic model proposed here is driven by an OWL ontology that covers all the concepts and relationships concerning the KPI elicitation and selection process. To this end, the standard Ontology 101 development process (Noy and McGuinness, 2001) has been followed, which main steps are:

1. Determine the domain and scope of the ontology. As a starting point, the ontology definition is based on the KPI selection and elicitation meta-model proposed by Maté et al. (2016).
2. Consider reusing existing ontologies. KPIOnto (Diamantini et al., 2016) has been considered for reusing, but it focus on KPI calculation modeling, without annotation of elementary concepts for our model, such as: KRIs, measures, contribution and decomposition. Therefore, the proposed ontology here (KPIOWL) has been designed from scratch, although incorporating properties and classes that could be used for a future ontology alignment with KPIOnto (and others).
3. Enumerate important terms in the ontology. Important terms in the ontology were extracted from KPI selection meta-model (Maté et al., 2016). Examples of such terms are: goal, relationship, indicator, measure, etc.
4. Define classes, properties, slots, facets and instances. These terms are described in Tables 2–4 . Instances (individuals in OWL) correspond to the specific indicators in the business strategical domain (in this article, water management). Individuals are obtained by mapping the dataset of sensorized data (in water management supply network) to RDF in accordance with the ontology.

As a result, the KPIOWL ontology is detailed next.

4.1. KPIOWL ontology

One of the main goals in this study is to formalize the metamodel described in Section 3. This metamodel has been extended to include relationships between indicators. Therefore, we opted to design an OWL 2 ontology to describe goals, relationships, indicators, measures, KPIs and KRIs. The result is a simple, although complete ontology called "KPIOWL", which allows to capture all the information required in business KPI selection and elicitation processes. The proposed ontology consists of 8 classes (groups of individuals sharing the same attributes), 11 object properties (binary relationships between individuals), 12 data properties (individual attributes), 128 logical axioms and 69 declaration axioms. The complete ontology is available in the GitHub repository.[6]

Fig. 2 shows the main classes in the hierarchy starting from the top class *Thing* (⊤), to the main classes, subclasses, and including some of their most interesting object and data properties. These main classes are: *Goal*, *Relationships* and *Indicator*. Each of these classes incorporates a set of properties and conditions in order to be conceptualized. An individual

---

that satisfies those properties is considered a member of that class. For a formal definition of them, we use next the OWL-DL syntax as described in Table 1, where a summarized logic syntax is represented (left-hand column) with regards to the corresponding OWL-DL equivalent (right).

– **Goal**. Class modeling the metamodel goals, i.e. the business objectives. Goals are related one to each other. Types of "*Relationships*" between *Goals* are subclases *Contribution* and *Decomposition* (see Section 3 and Maté et al., 2016 for further descriptions). The contribution relationship can be positive (direct) or negative (reverse). A main data property in this class annotates that each goal has a name. Table 2 shows the properties of the class *Goal* in OWL-DL notation.

– **Relationship**. It models relationships between goals. *Relationship* has two subclasses, *Contribution* and *Decomposition*. Each of these subclasses has different properties. Therefore, *Relationship* must be defined as a class with two subclases, and specifies the corresponding properties for each subclass. It is not defined as an object property with two subproperties, because OWL does not allow to define attributes for a property. The object property *hasRelationship* connects a goal with the corresponding type of relationship. The object property *withGoal* connects the relationship with the target goal if this relationship is *Decomposition*. The object properties *goalDirect* and *goalReverse* connect the relationship with the target goal if the relationship is *Contribution*. For example, in KPI modeling for a water management environment, a graphical representation of the decomposition relationship between the two goals "*maintain a healthy distribution network*" and "*renovate supply network*" is shown in Fig. 3. As data properties, each relationship has an *evidence*, a contribution has a *value* and a *formula* and a decomposition has a *type*, which is a logical operator. Table 3 displays the formal definitions of *Relationship*, *Contribution* and *Decomposition* object and data properties.

– **Indicator**. This class has three subclasses, which correspond to the types of indicators, i.e. *Measure*, *KRI* and *KPI*. Indicator class has a current value, a target value to be surpassed by the indicator, a threshold that separates acceptable from struggling performance, a worst value below which the indicator is considered to be failing, a target time, a name, a status and its aim (see data properties in Table 4). Indicators monitor goals. Indicators are also related to each other, so in a similar way to goals, this relationship can be positive (*indicatorDirect*) or negative (*indicatorReverse*). In addition, two object properties between indicators are defined in the ontology to capture whether a relationship between indicators is inconsistent or redundant, according to the information provided by knowledge domain experts (see object properties in Table 4) throughout the business intelligence modeling process.

KPIOWL is formally defined using OWL 2 DL, which facilitates its alignment with other OWL ontologies. For example, OWL provides owl:EquivalentClass and owl:subClass axioms to align classes in different ontologies. To integrate classes, properties and restrictions defined in a different ontology, OWL provides the owl:imports statement. Therefore, KPIOWL could be alignment with KPIOnto by asserting KPIOWL:Indicator owl:EquivalentClass KPIONTO:Indicator. To include the KPIONTO:Indicator properties in KPIOWL, KPIOnto must be imported.

4.2. Reasoning rule framework

Given the definitions above, a series of semantic rules are built on top of the KPIOWL ontology to deduce new information from the existing knowledge. These rules are formulated in SWRL (see Section 2.1) language to infer new relationships and/or to detect inconsistent relationships between indicators.

First, a set of rules are defined to infer relationships between indicators from the model. According to the model, an indicator i1 is considered to have a relationship with another indicator i2 if there is a relationship between the objectives monitored by each of them. The relationships between indicators capture the effect that i1 has over the value of i2. Formally, a sample of these rules (R1, R2 and R3) is as
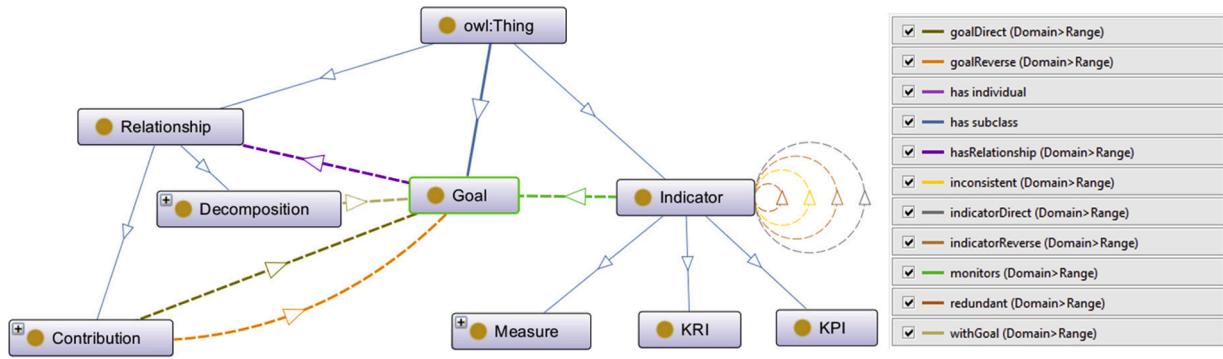
Fig. 2. KPIOWL ontology general scheme with main classes and properties. It has been generated by using OntoGraph plugging of Protégé tool for ontology modeling. Legend indicates relationships between classes by means of data and object properties.

Table 1
Basic OWL-DL semantic syntax used to formally define the proposed ontology.

Descriptions	Abstract syntax	DL syntax
Operators	intersection($C_1, C_2, …, C_n$)	$C_1 \sqcap C_2 \sqcap \cdots \sqcap C_n$
	union($C_1, C_2, …, C_n$)	$C_1 \sqcup C_2 \sqcup \cdots \sqcup C_n$
Restrictions	for at least 1 value V from C	$\exists V.C$
	for all values V from C	$\forall V.C$
	R is Symmetric	$R \equiv R^-$
Axioms	A partial($C_1, C_2, …, C_n$)	$A \sqsubseteq C_1 \sqcap C_2 \cdots \sqcap C_n$
	A complete($C_1, C_2, …, C_n$)	$A \equiv C_1 \sqcap C_2 \cdots \sqcup C_n$

Table 2
Goal: object and data properties.

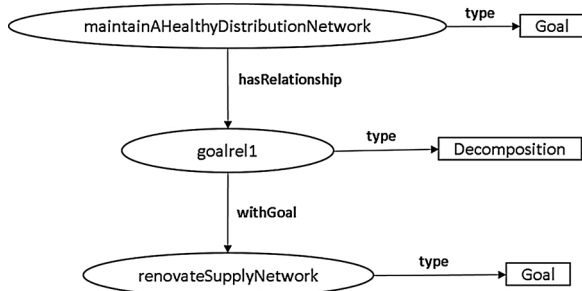| Object properties	Description logic
hasRelationship	$\exists$ hasRelationship.Thing $\sqsubseteq$ Goal
	$\top \sqsubseteq \forall$ hasRelationship.Relationship
monitoredBy	$\exists$ monitoredBy.Thing $\sqsubseteq$ Goal
	$\top \sqsubseteq \forall$ monitoredBy.Indicator
name	$\exists$ name.Datatype Literal $\sqsubseteq$ Goal
	$\top \sqsubseteq \forall$ name.Datatype string

Fig. 3. Example of a decomposition relationship between two goals.

follows:

```
R1: aim(? i1,"maximize")^monitors(? i1, ? o1)
^hasRelationship(? o1, ? gr1),
^Contribution(? gr1),
^goalDirect(? gr1, ? o2),
^aim(? i2,"maximize")^monitors(? i2, ? o2)
-> indicatorDirect(? i1, ? i2)

R2: aim(? i1,"maximize")^monitors(? i1, ? o1)
^hasRelationship(? o1, ? gr1),
^Contribution(? gr1),
^goalDirect(? gr1, ? o2),
```

Table 3
Relationship, Contribution and Decomposition: object and data properties.

Object properties	Description logic
withGoal	$\exists$ withGoal.Thing $\sqsubseteq$ Decomposition
	$\top \sqsubseteq \forall$ withGoal.Goal
goalDirect	$\exists$ goalDirect.Thing $\sqsubseteq$ Contribution
	$\top \sqsubseteq \forall$ goalDirect.Goal
goalReverse	$\exists$ goalReverse.Thing $\sqsubseteq$ Contribution
	$\top \sqsubseteq \forall$ goalReverse.Goal
evidence	$\exists$ evidence.Datatype Literal $\sqsubseteq$ Relationship
	$\top \sqsubseteq \forall$ evidence.Datatye double
value	$\exists$ value.Datatype Literal $\sqsubseteq$ Contribution
	$\top \sqsubseteq \forall$ value.Datatype string
formula	$\exists$ formula.Datatype Literal $\sqsubseteq$ Contribution
	$\top \sqsubseteq \forall$ formula.Datatype string
type	$\exists$ type.Datatype Literal $\sqsubseteq$ Decomposition
	$\exists$ type.Datatype Literal $\sqsubseteq$ Decomposition

Table 4
Indicator: object and data properties

Object properties	Description logic
monitors	$\exists$ monitors.Thing $\sqsubseteq$ Indicator
	$\top \sqsubseteq \forall$ monitors.Goal
indicatorDirect	$\exists$ indicatorDirect.Thing $\sqsubseteq$ Indicator
	$\top \sqsubseteq \forall$ indicatorDirect.Indicator)
indicatorReverse	$\exists$ indicatorReverse.Thing $\sqsubseteq$ Indicator
	$\top \sqsubseteq \forall$ indicatorReverse.Indicator)
inconsistent	$\exists$ inconsisten.Thing $\sqsubseteq$ Indicator
	$\top \sqsubseteq \forall$ inconsistent.Indicator)
redundant	$\exists$ redundant.Thing $\sqsubseteq$ Indicator
	$\top \sqsubseteq \forall$ redundant.Indicator)
name	$\exists$ name.Datatype Literal $\sqsubseteq$ Indicator
	$\top \sqsubseteq \forall$ name.Datatype string
currentValue	$\exists$ currentValue Datatype Literal $\sqsubseteq$ Indicator
	$\top \sqsubseteq \forall$ currentValue.Datatype double
status	$\exists$ status Datatype Literal $\sqsubseteq$ Indicator
	$\top \sqsubseteq \forall$ status.Datatype string
targetTime	$\exists$ targetTime Datatype Literal $\sqsubseteq$ Indicator
	$\top \sqsubseteq \forall$ targetTime.Datatype dateTime
targetValue	$\exists$ targetValue Datatype Literal $\sqsubseteq$ Indicator
	$\top \sqsubseteq \forall$ targetValue.Datatype double
threshold	$\exists$ threshold Datatype Literal $\sqsubseteq$ Indicator
	$\top \sqsubseteq \forall$ threshold.Datatype double
worstValue	$\exists$ worstValue Datatype Literal $\sqsubseteq$ Indicator
	$\top \sqsubseteq \forall$ worstValue.Datatype double
aim	$\exists$ aim Datatype Literal $\sqsubseteq$ Indicator
	$\top \sqsubseteq \forall$ aim.{"maintain", "maximize", "minimize"}

(continued)

```
^aim(? i2,"minimize")^monitors(? i2, ? o2)
```

```
-> indicatorReverse(? i1, ? i2)

R3: aim(? i1,"minimize")^monitors(? i1, ? o1)
^hasRelationship(? o1, ? gr1),
^Contribution(? gr1),
^goalDirect(? gr1, ? o2),
^aim(? i2,"minimize")^monitors(? i2, ? o2)
-> indicatorDirect(? i1, ? i2)
```

Aside from these rules that establish relationships between indicators, our semantic model is also able to infer inconsistency and redundancy between indicators in the model using both, its structure and data available.

A second rule is also defined next (R4) to indicate that, if two different indicators monitor the same goal, then there is a possible redundant relationship between these two indicators.

```
R4: monitors(? i2, ? o)^monitors(? i1, ? o)
-> redundant(? i1, ? i2)
```

In addition to SWRL rules, it is worth noting that *indicatorDirect* and *indicatorReverse* object properties are defined as OWL transitive properties. This allows the automatic reasoner to infer that, if an indicator i1 has a direct/reverse relationship with an indicator i2, and i2 has a direct/reverse relationship with an indicator i3, then there is a possible direct/reverse relationship between i1 and i3.

Finally, more complex rules could be added, which application depends on the available data, as the indicator's target value and current value are needed. An example of such a complex rule could be as follows:

When a goal (parent) is decomposed into several goals (children), and all of the children indicators are satisfied while the parent indicator is not, then the indicators are inconsistent. An indicator is satisfied if its current value is greater than its target value and its aim is maximize, or if its current value is less than its target value and its aim is minimize. An example of this kind of rules is defined in R5:

```
R5: monitors(? i1, ? o1)^monitors(? i2, ? o2)
^monitors(? i3, ? o3)^hasRelationship(? o1,? gr1)
^Decomposition(? gr1)^withGoal(? gr1, ? o2)
^withGoal(? gr1, ? o3)^aim(i1, "maximize")
^currentValue(? i1,? cv1)^targetValue(? i1,? tv1)
```

```
^lessThan(? cv1,? tv1)^aim(i2, "maximize")
^currentValue(? i2,? cv2)^targetValue(? i2,? tv2)
^greatherThan(? cv2,? tv2)^aim(i3, "maximize")
^currentValue(? i3,? cv3)^targetValue(? i3,? tv3)
^greatherThan(? cv3,? tv3)
-> inconsistent(? i1,? i2)^inconsistent(? i1,? i3)
```

4.3. Overall approach

As commented in Section 1, the selection of suitable indicators for business objectives entails the exploration of the specific business strategy together with domain experts, while providing data-driven insights whenever conformation or additional information is required. As argued in Maté et al. (2016), an iterative methodology that alternates conceptual modeling with data analysis will refine the KPI selection process. In this regard, the proposed semantic model allows domain experts to annotate business information and perform reasoning tasks to validate the strategic model obtained.

Fig. 4 illustrates an overview of the proposed approach, which consists on a two-fold strategy comprising: semantic model and KPI selection/elicitation meta-model. On top of this approach, domain experts are able to construct a strategic model in collaboration with the analyst, as well as to express their knowledge by means of semantic annotation, according to the KPIOWL scheme. This is performed in a first step (1), which is aimed at establishing the main business objectives pursued (which can be clearly related to a set of result indicators or performance indicators) and exploring other existing indicators, measures and objectives. In a second step (2), data involving indicators and measures are used as input for the data analysis (correlations) to detect potential or hidden relationships between indicators and to establish performance levels for measures. Step three (3) entails a series of reasoning task for semantic validation of those relationships between indicators extracted in the previous step. After this, the findings from the data analysis and the semantic validation are mapped back into the strategic model in step (4). Finally, a new cycle starts (step 5) until a stable strategic model is obtained (i.e. the model does not suffer a variation with respect to the previous iteration) or domain experts are satisfied with the current model.

In this approach, KPIOWL is the ontological scheme driving the whole process. It is the terminological box (TBox) that defines the
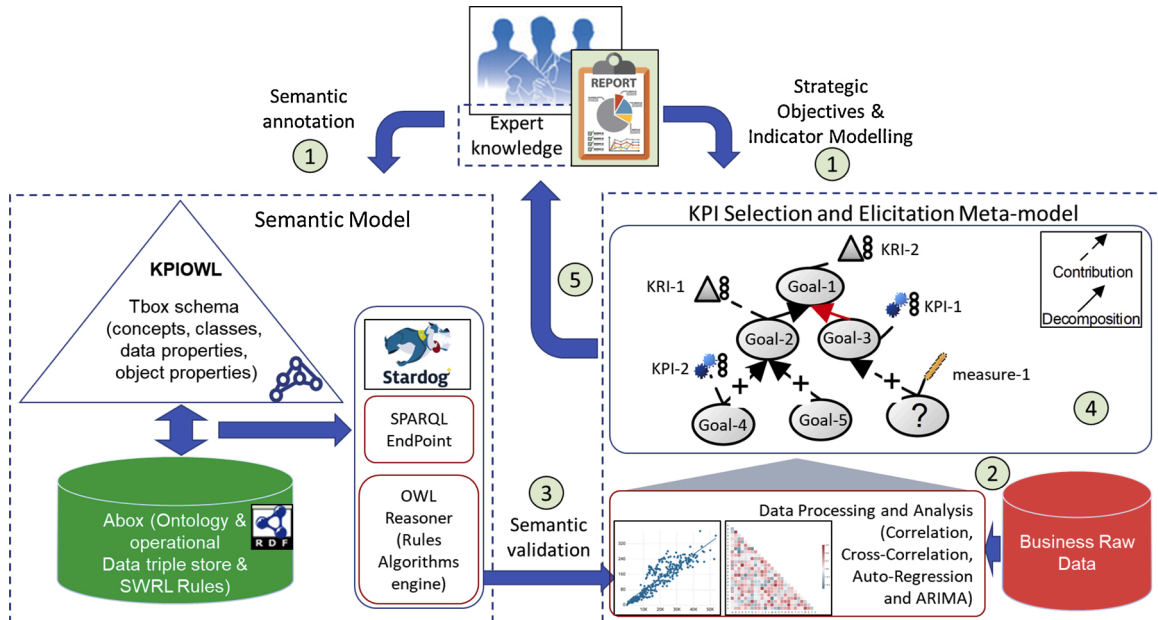

Fig. 4. General scheme of the semantic model driven by KPIOWL for business KPI selection and elicitation.

vocabulary with concepts and properties in the KPI selection meta-model. At bottom-left, the Assertional Box (ABox) defines all the instances in the knowledge domain (in OWL 2 an instance is represented by an individual) involving strategic models focused on indicators, i.e. specific goals, measures, etc. These instances are stored in RDF triple format in a Stardog[7] repository, which is a commercial version of the Pellet OWL 2 reasoner (Sirin, Parsia, Grau, Kalyanpur, & Katz, 2007), but enhanced with persistence capabilities. Once the ontology (Tbox) has been loaded together with SWRL rules, a series of reasoning tasks are launched by using the Stardog OWL 2 reasoner to derive new information that is not explicitly expressed in the knowledge base.

5. Use case

For validation purposes, a complete case of study has been developed to show how the proposed semantic approach is used for driving and validating the meta-modelling process of KPI selection. To this end, a real-world application is worked here, which consists in monitoring a water supply management network in the Mediterranean area of Alicante (Spain). It is a critical application nowadays because of the generalized water scarcity in the Southeast of Spain (Morote, Olcina, Rico, & Hernández, 2019), so the generation of innovative solutions for monitoring Key Performance Indicators is a must for policy and decision makers. Water supply management companies focus on ensuring water supply to multiple zones. On the one hand, water provided requires an adequate quality for its target, whether urban zones or farms, and cuts in service must be kept to a minimum. On the other hand, the water supply network incurs into losses, and must be renovated once critical points are reached. However, finding the specific parts of the network that require renovation is a challenging task, and thus entire blocks of the network have to be renovated, which is costly. In order to support in this task, a number of measures are gathered by the water supply management company in our study, although they still have to be associated with criteria to make decisions.

In this application, the original data sources describe hydro-graphic zones, sub-zones and monitored values, including measures, such as water leakages or structural breaks. Additionally, a series of mapping functions have been developed to convert these monitored values to RDF by following the ontology scheme provided by KPIOWL. The resulting data is then stored in the RDF repository (as shown in Fig. 4), which is now ready for querying and reasoning tasks. In this regard, it is worth specifying that KPIOWL has been extended with domain specific subclasses and individuals to enhance the inference procedure in reasoning time. Concretely, three subclasses of the class *Measure* have been included as shown in Fig. 5, namely: *TransportationNetwork*, *DistributionNetwork* and *ConnectionNetwork*, to classify those measures monitored in each of these three different parts of the water supply network (e.g. transportation, distribution and connection, respectively). In addition, a couple of new SWRL rules have been included in the semantic model to detect possible inconsistencies between direction of indicators and types of networks, as follows:

```
R6: ConnectionNetwork(? i2)
^TransportationNetwork(? i1)
^indicatorDirect(? i1, ? i2)
-> inconsistent(? i1, ? i2)

R7: ConnectionNetwork(? i2)
^TransportationNetwork(? i1)
^indicatorReverse(? i1, ? i2)
-> inconsistent(? i1, ? i2)
```

In accordance with the KPI selection model, the highest level goal is "*to provide an efficient water supply*", which does not have any known

measure associated. In order to track this high level objective, it is further decomposed into minimizing water lost and improve network efficiency. In order to minimize water lost, intuitively the company wishes to minimize breakdowns and leaks, which are avoided by maintaining the supply network and renovating it when needed. However, renovating the supply network involves a costly process, and thus harms the reduction of maintenance costs. For the first iteration of the analysis, a number of 20 measures have been considered, which contain yearly readings for the period of 2008–2014 (6 data points) for 574 instances of the data. An additional measure is included to refer to "*water lost*" (not directly available), which is calculated from the difference of water supplied and water registered. In a second phase, a preliminary data analysis is carried out, which comprises normalization of data values within regions, and correlation between different measures. This last analysis shows a number of 22 correlations (Pearson) stronger than 0.5 threshold, which are further analyzed using linear regression to determine the predicting power of each factor. The resulting model is then updated as illustrated in Fig. 6 , where a series of potential relationships between result indicators have been identified.

At next iteration, three new relationships have been identified as non-interesting, since the measures involved are calculated in a similar fashion. These relationships relate the number of natural leaks with the number of leaks according to the size of the network (Measures 12–16, 13–17) and the water supplied to the water lost (20-water lost). On the other hand, other three relationships that cover the renewed lengths of the different networks (4–5), number of water connections with the number of leaks, (7–11), and hydraulic system performance with final water lost (19–20) are marked as of special interest. However, as can be observed in Fig. 7, an initially expected relationship between measure regarding to breakdowns in the network and water lost is not supported by the data analysis. This leads the expert to review either the way the main goal is monitored, i.e. how are breakdowns measured, or review the suitability of this relationship, i.e. "breakdowns not cause water loss" (no sense semantically). A further semantic validation can be then carried out by means of reasoning tasks within KPIOWL on the annotated data, which effectively shows inconsistencies in this regard. Specifically, following our model, the water supply company experts define that "breakdowns monitors water lost and minimize breakdowns monitors minimize water lost". Breakdowns and water lost are indicators while minimize breakdowns and minimize water lost are goals. Furthermore, a contribution relationship between the two goals is included, so the application of rule R3 (see Section 4.2) infers that "breakdowns have a direct relationship with water lost".

A second task consists in using the semantic reasoner to infer new semantic relationships, by applying transitivity. This experiment provides the following results, which complement the analysis carried out from measured data in the KPI meta-modeling phase. For example, after applying rule R5, the following new relationships are inferred:

```
- AvgNetworkAge
  is-related-to WaterLost
- NaturalLeaksDivLengthOfDistributionNetwork
  is-related-to WaterLost
- NaturalLeaksDivLengthOfTransportionNetwork
  is-related-to WaterLost
- NaturalLeaksDivWaterConnections
  is-related-to WaterLost
- WaterConnectionsDivNetworkLength
  is-related-to WaterLost
```

By applying the rest of semantic rules (R5 and R6), a series of new relationships are obtained involving "*breakdowns*":

```
- breakdowns
  is-related-to waterRegistered
- breakdowns
  is-related-to waterlostperkmandday
- breakdowns
  is-related-to waterlost
```
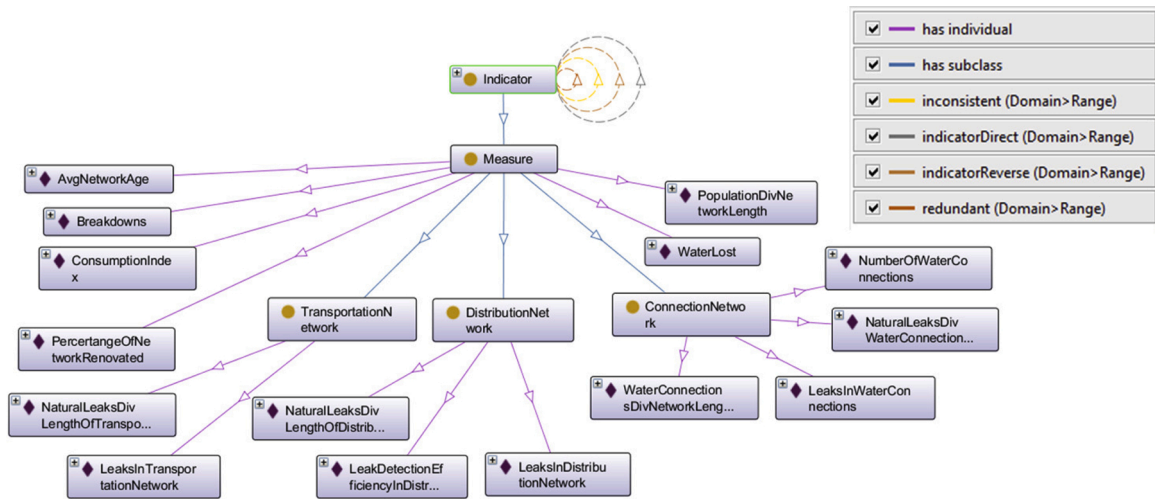
Fig. 5. Ontology overview of additional elements for the specific use case of water management.
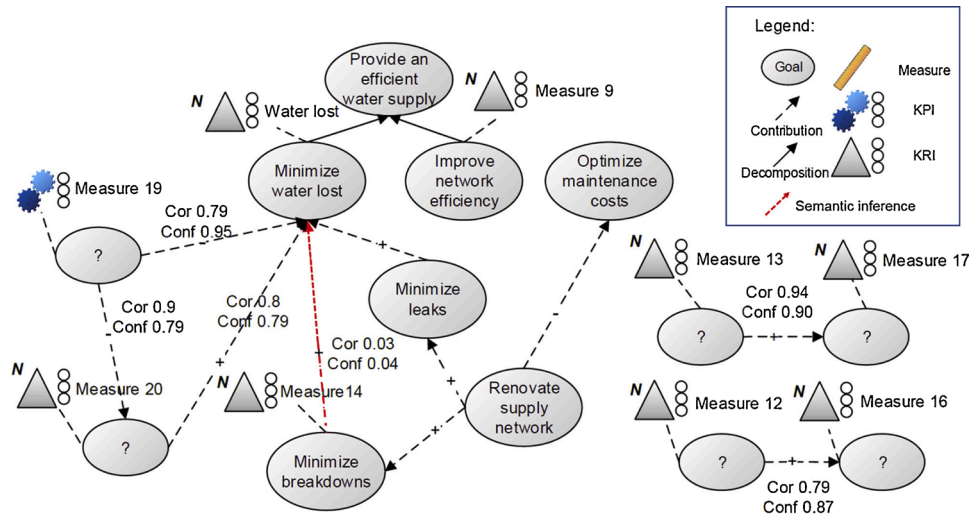

Fig. 6. Subset of the indicator model updated with data analysis results.

(continued)

- breakdowns
  is-related-to watersuply

Finally, by applying the rule to detect possible redundant relationships between indicators (R4), we obtain among others:

- Waterlost and waterRegistered
- waterlost and waterSupplied
- waterlost and waterlostperkmandday
- renovated network lenght and
  percentage of renovated network
- renovated network lenght and
  avg network renovated

A further step is to gather additional data and perform new analyses that lead the experts to obtain more insights, although it has been shown how the proposed semantic model has already helped in both, simplifying the indicator list, as well as enriching the strategic model, in the context of the water management case study.

6. Theoretical and managerial implications

The increasing amount of available data from multiple and heterogeneous sources in organizations entails an opportunity when feeding Business Intelligence and Analytic processes, although it also poses additional technical complexity derived from dealing with such diversity of data (Torres & Sidorova, 2019). The task of offering Information Management (IM) tools for supporting managers in monitoring organizations and in decision making processes, requires BI systems to work over standardized data processes through the whole data life cycle, since it is crucial to guarantee the conversion from raw data to valuable information about the business domain of knowledge.

In this sense, ontological approaches provide information management tools with a knowledge-based semantically enriched for the interpretation of unstructured and merged content (Mikroyannidis & Theodoulidis, 2010). However, the use of ontologies in IM tools is still preliminary and disregards important factors, such as: modeling concrete enterprise strategies driven by goals, statistical features of data and their relations with business indicators, linkage and integration with other external data ontologies, and especially, their implementation for practical use.

The KPIOWL semantic model proposed in this work is aimed at coping with these issues, so it goes beyond the conceptual modelling towards the actual implementation and use in a real-world scenario, although it can be exported to other different cases where monitored data are duly modelled in form of KRIs and KPIs in BI processes. In the
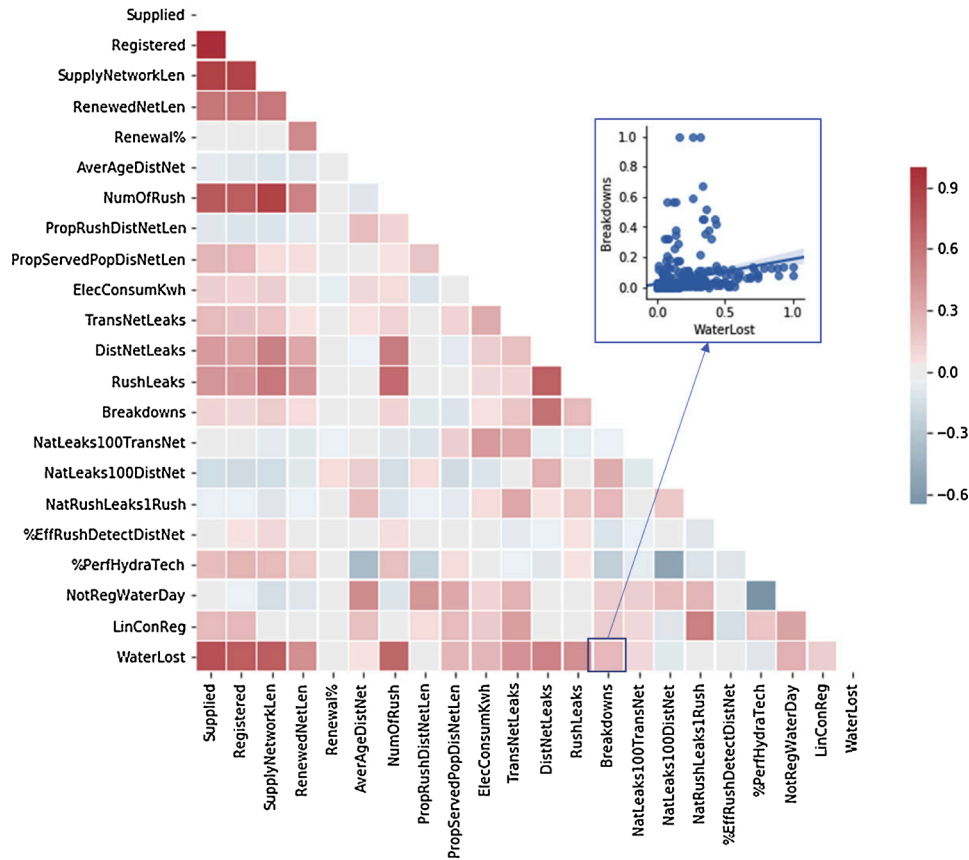
Fig. 7. Correlations computed between measures.

light of the experiences reported in this study, a series of implications can be extracted from theoretical and managerial points of view.

6.1. Theoretical implications

Our study offers two main theoretical implications to the efficient modeling of Business Intelligence processes and indicators for the generation of value from data. A first implication concerns the limited focus on the improved-decision making, without considering how knowledge domain is semantically contextualized, modelled and integrated. We contribute to this research by focusing specifically on the mechanism of selection of relevant KPIs, since it is an important initial task in BI processes to monitor performance to undertake the relevant corrective actions when necessary. Concretely, our model strategy emphasizes on the distinction between the concepts of KPI and KRI (Parmenter, 2015), which is crucial to avoid mistakes and enable the process of finding an adequate KPIs to focus on for an enterprise objective.

In this sense, KPIs are responsible of monitoring the performance of the crucial processes of the organizations taking into account the implemented corporate strategy, which requires to have a panoramic view of the company's processes and activities. Without a holistic and up-to-date vision of these processes, it is not possible to design a strategy tailored to the actual needs and potential of the organization, and therefore, to select the most appropriate KPIs to build a solid and valuable strategy. To this end, the adoption of ontology-driven approaches constitutes a promising line of work, since it allows semantic consolidation of heterogeneous sources of information and contextualization (Ogiela & Ogiela, 2014).

This entails the second implication of our proposal, which relies on the materialization of the semantic model (Tbox, Abox) on software architectures comprising RDF repository, SPARQL Endpoint and mapping methods for data gathering. In addition, the ontology alignment with other related proposals, such as QB4OLAP (Etcheverry et al., 2014) and KPIOnto (Diamantini et al., 2016), would complement KPIOWL allowing higher contextualization and richer strategic business modelling. This will likely contribute to create more detailed models and possibly extend the modeling language, where these complex relationships can be reflected explicitly in order to provide additional insights and ideas for domain experts. All this will allow the connection with external Open Data repositories and the generation of Linked Data with the possibility of performing federated querying and advanced reasoning (Mikroyannidis & Theodoulidis, 2010). The incorporation of Open Linked Data will enrich the semantic model with new perspectives of information, such as new data analysis with richer and more sophisticated algorithms, which could help in detecting more complex relationships between indicators. The use of reasoning procedures will lead the inference of new implicit information that could be incorporated to the knowledge base.

6.2. Managerial implications

In terms of managerial implications, a major issue often detected in current BI processes lies in the lack of machine-readable representation of such processes as a whole on a semantic level (Colomo-Palacios, García-Crespo, Soto-Acosta, Ruano-Mayoral, & Jiménez-López, 2010). Key Performance Indicators are core elements in BI for efficiently supporting decision making, so the formal modeling and data standardization in this kind of processes should be duly considered in practice. The present work investigates on the potential of ontologies in improving the selection of KPIs, within a corporate environment for business intelligence. The paper shows the use of KPIOWL, a framework that employs ontology management and evolution in the context of information management systems. The capabilities of the framework in facilitating information management and business intelligence are

evaluated through a real-life case study of urban management, specifically for water supply network monitoring.

The designed architecture and functionalities aim to create coherent semantic data layers for KPI modelling and descriptive analysis (correlations, ARIMA, etc.), defining the semantics for knowledge sharing and reasoning capabilities. Descriptive data analysis and insights identification from semantic querying and reasoning can transform and add value to an organization. This paper presents a knowledge management and engineering perspective (ontology based) for the application of KPI selection meta-model and insights at the organizational (corporate) workplace towards the development of the organizational BI processes.

In this regard, consistency preservation is another goal of ontological models. As done in related works such as Heraclitus II (Mikroyannidis & Theodoulidis, 2010), consistency is performed semantically by resolving inconsistencies that arise when the structure or semantics of an ontology become invalid because of a wrong assumption from data analytics (e.g. lack of causality in correlations). Semantic Knowledge-bases taking into account actual human expert's experience could led to detect, thanks to reasoning rules, semantic inconsistencies delivered from unbiased data analysis.

7. Conclusions

In this work, the ontology-driven approach KPIOWL is proposed to formally conceptualize essential elements of indicators, covering: performance, results, measures, goals and relationships of a given business strategy. In this way, all the data involved in the selection and analysis of KPIs are then annotated, integrated and stored in an RDF repository, hence enabling sophisticated querying and reasoning for semantic validation. The main objective is to enrich strategic business models for the efficient elicitation, assessment and selection of KPIs and KRIs.

The proposed semantic model is evaluated on a real-world case study on water management in the Mediterranean region of Alicante (Spain), where the elicitation and selection of correct indicators is mandatory for improving the supply network efficiency. A series of data analysis and reasoning tasks are conducted to show how the ontological model is able to detect semantic conflicts in actual (although wrong from a business perspective) correlations of selected indicators.

The semantic model elaborated here is also applicable to other domains where the selection of KPIs is involved. This motivates our main future line of research. In addition, ongoing work is focusing on the incorporation of Open Linked Data to enrich the semantic model with new perspectives of information, such as new data analysis with richer and more sophisticated algorithms, which could help in detecting more complex relationships between indicators. This will likely contribute to create more detailed models and possibly extend the modeling language, where these complex relationships can be reflected explicitly in order to provide additional insights and ideas for domain experts. In this regard, ontology alignment with other related proposals, such as KPIOnto (Diamantini et al., 2016), would complement KPIOWL allowing higher contextualization and richer strategic business modelling.

References

Barone, D., Topaloglou, T., & Mylopoulos, J. (2012). Business intelligence modeling in action: A hospital case study. In J. Ralyté, X. Franch, S. Brinkkemper, & S. Wrycza (Eds.), *Advanced information systems engineering* (pp. 502–517). Berlin, Heidelberg: Springer Berlin Heidelberg.

Bellatreche, L., Khouri, S., & Berkani, N. (2013). Semantic data warehouse design: From etl to deployment à la carte. In W. Meng, L. Feng, S. Bressan, W. Winiwarter, & W. Song (Eds.), *database systems for advanced applications* (pp. 64–83). Berlin, Heidelberg: Springer Berlin Heidelberg.

Benke, K., & Tomkins, B. (2017). Future food-production systems: vertical farming and controlled-environment agriculture. *Sustainability: Science, Practice and Policy, 13*(1), 13–26. https://doi.org/10.1080/15487733.2017.1394054.

Berkani, N., Bellatreche, L., & Benatallah, B. (2016). A value-added approach to design bi applications. In S. Madria, & T. Hara (Eds.), *Big data analytics and knowledge discovery* (pp. 361–375). Cham: Springer International Publishing.

Brickley, D., & Miller, L. (2000). *Friend of a friend (FOAF): an experimental linked information system* (Online). Accessed 23.07.19 http://www.foaf-project.org/.

Colomo-Palacios, R., García-Crespo, Ángel, Soto-Acosta, P., Ruano-Mayoral, M., & Jiménez-López, D. (2010). A case analysis of semantic technologies for R&D intermediation information management. *International Journal of Information Management, 30*(5), 465–469. https://doi.org/10.1016/j.ijinfomgt.2010.05.012.

Community, S. (2019). *Statistical data and metadata eXchange* (Online). Accessed 23.07.19 https://sdmx.org/.

Cyganiak, R., & Reynolds, D. (2014). *The RDF data cube vocabulary* (Online). Accessed 08.04.18 https://www.w3.org/TR/vocab-data-cube/.

del Mar Roldán García, M., García-Nieto, J., & Aldana-Montes, J. F. (2016). An ontology-based data integration approach for web analytics in e-commerce. *Expert Systems with Applications, 63*, 20–34. https://doi.org/10.1016/j.eswa.2016.06.034.

del Río-Ortega, A., Resinas, M., & Ruiz-Cortés, A. (2010). Defining process performance indicators: An ontological approach. In R. Meersman, T. Dillon, & P. Herrero (Eds.), *On the move to meaningful internet systems: OTM 2010* (pp. 555–572). Berlin, Heidelberg: Springer Berlin Heidelberg.

Diamantini, C., Genga, L., Potena, D., & Storti, E. (2014). Collaborative building of an ontology of key performance indicators. In R. Meersman, H. Panetto, T. Dillon, M. Missikoff, L. Liu, O. Pastor, A. Cuzzocrea, & T. Sellis (Eds.), *On the move to meaningful internet systems: OTM 2014 conferences* (pp. 148–165). Berlin, Heidelberg: Springer Berlin Heidelberg.

Diamantini, C., Potena, D., & Storti, E. (2016). Sempi: A semantic framework for the collaborative construction and maintenance of a shared dictionary of performance indicators. *Future Generation Computer Systems, 54*, 352–365. https://doi.org/10.1016/j.future.2015.04.011.

Etcheverry, L., Vaisman, A., & Zimányi, E. (2014). Modeling and querying data warehouses on the semantic web using QB4OLAP. In L. Bellatreche, & M. K. Mohania (Eds.), *Data warehousing and knowledge discovery* (pp. 45–56). Cham: Springer International Publishing.

Gómez, C., Ayala, C., Franch, X., López, L., Behutiye, W., & Martínez-Fernández, S. (2017). Towards an ontology for strategic decision making: The case of quality in rapid software development projects. In S. de Cesare, & U. Frank (Eds.), *Advances in conceptual modeling*. Cham: Springer International Publishing.

Grosof, B. N., & Poon, T. C. (2004). Sweetdeal: Representing agent contracts with exceptions using semantic web rules, ontologies, and process descriptions. *International Journal of Electronic Commerce, 8*(4), 61–97.

Group, T. D. C. (2017). *DCMI type vocabulary*. , (Online). Accessed 23.07.19 http://www.dublincore.org/documents/dcmi-type-vocabulary/.

Group, W. O. W. (2019). *OWL 2 web ontology language: Document overview* (Online). Accessed 05.04.18 http://www.w3.org/TR/owl2-overview/.

Gruber, T. R., et al. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition, 5*(2), 199–220.

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human–Computer Studies, 43*(5–6), 907–928.

Guarino, N., et al. (1998). Formal ontology and information systems. In *Proceedings of FOIS* (pp. 81–97).

Harris, S., Seaborne, A., & Prud'hommeaux, E. (2013). Sparql 1.1 query language. *W3C Recommendation, 21*(10).

Hausenblas, M., Halb, W., Raimond, Y., Feigenbaum, L., & Ayers, D. (2009). Scovo: Using statistics on the web of data. In L. Aroyo, P. Traverso, F. Ciravegna, P. Cimiano, T. Heath, E. Hyvönen, R. Mizoguchi, E. Oren, M. Sabou, & E. Simperl (Eds.), *The semantic web: Research and applications* (pp. 708–722). Berlin, Heidelberg: Springer Berlin Heidelberg.

Horkoff, J., Barone, D., Jiang, L., Yu, E., Amyot, D., Borgida, A., et al. (2014). Strategic business modeling: Representation and reasoning. *Software & Systems Modeling, 13*(3), 1015–1041.

Horrocks, I., Patel-Schneider, P. F., Bechhofer, S., & Tsarkov, D. (2005). OWL rules: A proposal and prototype implementation. *Web Semantics: Science, Services and Agents on the World Wide Web, 3*(1), 23–40.

Huang, S.-M., Chou, T.-H., & Seng, J.-L. (2007). Data warehouse enhancement: A semantic cube model approach. *Information Sciences, 177*(11), 2238–2254. https://doi.org/10.1016/j.ins.2006.12.022.

Kehlenbeck, M., & Breitner, M. H. (2009). Ontology-based exchange and immediate application of business calculation definitions for online analytical processing. In T. B. Pedersen, M. K. Mohania, & A. M. Tjoa (Eds.), *data warehousing and knowledge discovery* (pp. 298–311). Berlin, Heidelberg: Springer Berlin Heidelberg.

Kritikos, K., Plexousakis, D., & Woitch, R. (2018). A flexible semantic kpi measurement system. In D. Ferguson, V. M. Muñoz, J. Cardoso, M. Helfert, & C. Pahl (Eds.), *Cloud computing and service science* (pp. 237–261). Cham: Springer International Publishing.

Maté, A., de Gregorio, E., Cámara, J., & Trujillo, J. (2014). Improving massive open online courses analysis by applying modeling and text mining: A case study. In J. Parsons, & D. Chiu (Eds.), *Advances in conceptual modeling* (pp. 29–38). Cham: Springer International Publishing.

Maté, A., Trujillo, J., & Mylopoulos, J. (2016). Key performance indicator elicitation and selection through conceptual modelling. In I. Comyn-Wattiau, K. Tanaka, I.-Y. Song, S. Yamamoto, & M. Saeki (Eds.), *Conceptual modeling* (pp. 73–80). Cham: Springer International Publishing.

McBride, B. (2004). The resource description framework (RDF) and its vocabulary description language RDFs. In *Handbook on ontologies* (pp. 51–65).

McGuinness, D. L., Van Harmelen, F., et al. (2004). OWL web ontology language overview. *W3C Recommendation, 10*(10), 2004.

Mikroyannidis, A., & Theodoulidis, B. (2010). Ontology management and evolution for business intelligence. *International Journal of Information Management, 30*(6), 559–566. https://doi.org/10.1016/j.ijinfomgt.2009.10.002.

Miles, A., & Brickley, D. (2005). *SKOS core vocabulary specification* (Online). Accessed 23.07.19 https://www.w3.org/TR/swbp-skos-core-spec/.

Morote, L.-F., Olcina, J., Rico, A.-M., & Hernández, M. (2019). Water management in urban sprawl typologies in the city of Alicante (southern Spain): New trends and perception after the economic crisis? *Urban Science, 3*(1). https://doi.org/10.3390/urbansci3010007.

Nebot, V., & Berlanga, R. (2012). Building data warehouses with semantic web data. *Decision Support Systems, 52*(4), 853–868. https://doi.org/10.1016/j.dss.2011.11.009 (1) Decision Support Systems for Logistics and Supply Chain Management. (2) Business Intelligence and the Web.

Noy, N. F., & McGuinness, D. L. (2001). *Dontology development 101: A guide to creating your first ontology, Tech. rep.* Stanford University Knowledge Systems Laboratory. Technical Report KSL-01-05. http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html.

Noy, N. F., McGuinness, D. L., et al. (2001). *Ontology development 101: A guide to creating your first ontology*.

Ogiela, L., & Ogiela, M. R. (2014). Cognitive systems for intelligent business information management in cognitive economy. *International Journal of Information Management, 34*(6), 751–760. https://doi.org/10.1016/j.ijinfomgt.2014.08.001.

Parmenter, D. (2015). *Key Performance Indicators: Developing, implementing, and using winning KPIs.* New York, NY, USA: John Wiley & Sons, Inc.

Popova, V., & Treur, J. (2005). A specification language for organisational performance indicators. In M. Ali, & F. Esposito (Eds.), *Innovations in applied artificial intelligence* (pp. 667–677). Berlin, Heidelberg: Springer Berlin Heidelberg.

Prud, E., Seaborne, A., et al. (2006). Sparql query language for RDF. *W3C Recommendation*.

Sell, D., Cabral, L., Motta, E., Domingue, J., & Pacheco, R. (2005). Adding semantics to business intelligence. *16th international workshop on database and expert systems applications (DEXA'05*, 543–547. https://doi.org/10.1109/DEXA.2005.44.

Silva Souza, V. E., Mazón, J.-N., Garrigós, I., Trujillo, J., & Mylopoulos, J. (2012). Monitoring strategic goals in data warehouses with awareness requirements. *Proceedings of the 27th annual ACM symposium on applied computing, SAC'12* (pp. 1075–1082). New York, NY, USA: ACM. https://doi.org/10.1145/2245276.2231944.

Simoni, G. D. (2018). *How to use semantics to drive the business value of your data* (Online). Accessed 30.01.19 https://www.gartner.com/doc/3894095/use-semantics-drive-business-value.

Sirin, E., Parsia, B., Grau, B. C., Kalyanpur, A., & Katz, Y. (2007). Pellet: A practical OWL-DL reasoner. *Web Semantics: Science, Services and Agents on the WWW, 5*(2), 51–53.

Staab, S., & Studer, R. (2013). *Handbook on ontologies.* Springer Science & Business Media.

Tenneson, C., & Brocklehurst, G. (2018). *Digital business KPIs: Defining and measuring success for tech CEOs* (Online). Accessed 30.01.19 https://www.gartner.com/doc/3891236/digital-business-kpis-defining-measuring.

Torres, R., & Sidorova, A. (2019). Reconceptualizing information quality as effective use in the context of business intelligence and analytics. *International Journal of Information Management, 49*, 316–329. https://doi.org/10.1016/j.ijinfomgt.2019.05.028.